Exp. 1 Revised Analysis

Reviewer 2

There is another concern about the data analysis that I'd like to raise. The trial-level inclusion criterion of a minimum of 500 ms during the analysis window seems extremely liberal, particularly since the window is 3500 ms long.

However, a subsequent PNAS paper by Bergelson & Aslin (2017), also cited in the manuscript, used a more conservative criterion: 'Trials were excluded if infants did not look at either image for at least 1/3 of the target window'. This 1/3 criterion was also used by Bergelson & Swingley (2017). The 1/3 choice seems like a much more reasonable inclusion criterion than the 1/7 used in the current study, which seems like it could lead to a lot of noise.

It is also standard to exclude infants who do not provide data for at least half of the test trials (cf. Bergelson & Swingley, 2017). I understand that this is a very difficult-to-test sample, and that the authors want to retain as many infants in the sample as possible, but there is a tradeoff between data retention and interpretability of the effects. In particular, the 1/7 data retention criterion seems quite problematic.

Adopting (3133/3)ms looking-time threshhold for trial inclusion

```
## exclusions
cn_r2$exclusions <- cn_r2$keep
cn_r2$exclusions[cn_r2$keep=='keep'] <- ''
cn_r2$exclusions[cn_r2$exclusions %in% c(
  'infant not paying attention', 'infant crying')] <-
  'restive_infant'
cn_r2$exclusions[cn_r2$pre_looking_sum_ms == 0] <-
  'zero_pre_looking'
cn_r2$exclusions[cn_r2$post1_looking_sum_ms < (3133/3)] <-
  'under_third_window_looking'
```

```
cn_r2$exclusions[cn_r2$exclusions=='mom_interference'] <-
  'mom_error'
cn_r2$exclusions[cn_r2$exclusions=='discard'] <-
  ''
TOTAL_EXCLUDED_TRIALS <- nrow(cn_r2[cn_r2$exclusions %in%
                                    c('restive_infant', 'zero_pre_looking',
                                      'mom_error',
                                      'under_third_window_looking'),])
+ 3 # mom_errors flagged in Datavyu export script:
```

```
## [1] 3
```

```
# 'Invalid number of wordWindow cells' (because of mom not speaking prompt)'
```

```
cn_exclusions_tab <- cn_r2 %>%
  filter(exclusions!='trial_was_recycled',
         exclusions!='old_trial_type') %>%
  group_by(subject_id) %>%
  summarize(excluded_trials=sum(exclusions!='')) %>%
  ungroup() %>%
  summarize(children_w_excl_trials=sum(excluded_trials>0),
            min_excl_trials=min(excluded_trials),
            max_excl_trials=max(excluded_trials),
            mean_excl_trials=mean(excluded_trials))

xtable2kable(cn_exclusions_tab)
```

```
## # A tibble: 1 x 4
##   children_w_excl_trials min_excl_trials max_excl_trials mean_excl_trials
##                    <int>           <int>           <int>            <dbl>
## 1                     20               0              16             4.52
```

Excluded test trials came from 20 infants who had between 0 and 16 excluded trials; all infants provided usable data for at least half of the test trials. If we instead adopt the criteria that infants should provide data for at least half of the difference score calculations, then one subject ('J94252'), with 16 test trials, but only two calculable difference scores, should be excluded (see next section—'EXCLUSION CRITERIA 2'—for revised mean difference score analyses, which use this criterion).

```
cn_keeptrials_tab <- cn_r2 %>%
  group_by(subject_id) %>%
  summarize(good_sum=sum(exclusions==''),
            good_prop = good_sum/n(),
```

```
              good_prop_of_total = good_sum/32) %>%
    ungroup() %>%
    reframe(min_good_trials=min(good_sum),
             min_n=sum(good_sum==min_good_trials),
             max_good_trials=max(good_sum),
             max_n=sum(good_sum==max_good_trials),
             mean_good_trials=mean(good_sum),
             med = median(good_sum),
             mode_good_trials=DescTools::Mode(good_sum),
             mode_n=sum(good_sum==mode_good_trials),
             min_good_prop = min(good_prop_of_total),
             max_good_prop=max(good_prop_of_total),
             mean_good_prop=mean(good_prop_of_total)) %>%
    round(., 2)

xtable2kable(cn_keeptrials_tab)
```

```
32  ## # A tibble: 3 x 11
33  ##   min_good_trials min_n max_good_trials max_n mean_good_trials   med
34  ##             <dbl> <dbl>           <dbl> <dbl>            <dbl> <dbl>
35  ## 1              16     3              32     1             25.1    26
36  ## 2              16     3              32     1             25.1    26
37  ## 3              16     3              32     1             25.1    26
38  ## # i 5 more variables: mode_good_trials <dbl>, mode_n <dbl>,
39  ## #   min_good_prop <dbl>, max_good_prop <dbl>, mean_good_prop <dbl>
```

A total of 93 trials from 20 infants were excluded because the infant never looked to the displays during the pre-naming period ($n = 14$), looked for less than one-third of the analysis window ($n = 72$), was distracted or crying ($n = 3$), or because the caregiver erred or interfered ($n = 7$). All infants contributed at least half (16) of test trials ($range= c(16, 16, 16) - c(32, 32, 32)$, $M_{\text{trials}} = c(25.14, 25.14, 25.14)$, $Med_{\text{trials}} = c(26, 26, 26)$).

EXCLUSION CRITERIA 1

```
cn_fin <- cn_r2 %>% filter(exclusions=='')
write.csv(cn_fin,
          here('data/common_nouns_reviewer2_data.csv'))

cn_fin_pp_df <- cn_fin %>%
  distinct(subject_id, bebe_meses, mama_edad)

N <- length(unique(cn_fin_pp_df$subject_id))
```

```
MIN_AGE <- min(cn_fin_pp_df$bebe_meses)
MAX_AGE <- max(cn_fin_pp_df$bebe_meses)
MEAN_AGE <- mean(cn_fin_pp_df$bebe_meses)
SD_AGE <- sd(cn_fin_pp_df$bebe_meses)
```

46   We report on data from 21 infants ($5.33 - 15.43 mos$, $M_{age} = 10.96 mos$, $SD_{age} =$
47 $2.71 mos$).

```
MOT_M_AGE <- na.mean(cn_fin_pp_df$mama_edad)
MOT_MIN_AGE <- min(cn_fin_pp_df$mama_edad, na.rm=T)
MOT_MAX_AGE <- max(cn_fin_pp_df$mama_edad, na.rm=T)
MOT_SD_AGE <- sd(cn_fin_pp_df$mama_edad, na.rm=T)
```

48   Mothers ranged in age from 14 to $44 yrs$ ($M_{age} = 26.85 yrs$, $SD_{age} = 8.36 yrs$)

49 **Mean Difference Score Analysis**

```
## specific image pairs (which of two image-pairs for each noun-pair)
# want to know the pairs of values that go together for the B&S 2012 calculation,
# which means the specific instances of each noun-pair where, e.g., baby is on L
# and corn is on R
cn_fin$trial_name_short <- gsub('^[0-9]*[-]', '', cn_fin$old_trial_name)

cn_fin$stimulus_set <- ''
# baby left, corn right = 'A'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'babcor-baby-L', 'babcor-corn-R',
  'BABY-corn', 'baby-CORN',
  'corbab-corn-R', 'corbab-baby-L')] <- 'A'

# corn left, baby right = 'B'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'babcor-baby-R', 'babcor-corn-L',
  'corbab-baby-R', 'corbab-corn-L',
  'corn-BABY', 'CORN-baby')] <- 'B'

# car left, shoe right = 'C'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'carsho-car-L', 'carsho-shoe-R',
  'CAR-shoe', 'car-SHOE')] <- 'C'

# shoe left, car right = 'D'
```

```r
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'carsho-car-R', 'carsho-shoe-L',
  'shocar-shoe-L', 'shocar-car-R',
  'SHOE-car', 'shoe-CAR', 'carshow-car-R')] <- 'D'

# chayote left, cow right = 'E'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'chacow-chayote-L', 'chacow-cow-R',
  'chayote-COW', 'CHAYOTE-cow')] <- 'E'

# cow left, chayote right = 'F'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'chacow-chayote-R', 'chacow-cow-L',
  'cowcha-cow-L', 'cowcha-chayote-R',
  'COW-chayote', 'cow-CHAYOTE')] <- 'F'

# dog left, fire right = 'G'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'dogfir-dog-L', 'dogfir-fire-R',
  'DOG-fire', 'dog-FIRE',
  'firdog-dog-L', 'firdog-fire-R')] <- 'G'

# fire left, dog right = 'H'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'dogfir-dog-R', 'dogfir-fire-L',
  'fire-DOG', 'FIRE-dog',
  'firdog-dog-R', 'firdog-fire-L')] <- 'H'

# horse left, soda right = 'I'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'horsod-horse-L', 'horsod-soda-R',
  'HORSE-soda', 'horse-SODA',
  'sodhor-horse-L', 'sodhor-soda-R')] <- 'I'

# soda left, horse right = 'J'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'horsod-horse-R', 'horsod-soda-L',
  'SODA-horse', 'soda-HORSE',
  'sodhor-horse-R', 'sodhor-soda-L')] <- 'J'

# rabbit left, soup right = 'K'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'rabsou-rabbit-L', 'rabsou-soup-R',
  'RABBIT-soup', 'rabbit-SOUP',
```

```r
  'sourab-rabbit-L', 'sourab-soup-R')] <- 'K'

# soup left, rabbit right = 'L'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'rabsou-rabbit-R', 'rabsou-soup-L',
  'sourab-rabbit-R', 'sourab-soup-L',
  'SOUP-rabbit', 'soup-RABBIT')] <- 'L'

# sheep left, water right = 'M'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'shewat-sheep-L', 'shewat-water-R',
  'sheep-WATER', 'SHEEP-water',
  'watshe-sheep-L', 'watshe-water-R')] <- 'M'

# water left, sheep right = 'N'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'shewat-sheep-R', 'shewat-water-L',
  'watshe-water-L', 'watshe-sheep-R',
  'water-SHEEP', 'WATER-sheep')] <- 'N'

# chicken left, tortilla right = 'O'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'chitor-chicken-L', 'chitor-tortilla-R',
  'chicken-TORTILLA', 'CHICKEN-tortilla',
  'torchi-chicken-L', 'torchi-tortilla-R')] <- 'O'

# tortilla left, chicken right = 'P'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'chitor-chicken-R', 'chitor-tortilla-L',
  'torchi-tortilla-L', 'torchi-chicken-R',
  'tortilla-CHICKEN', 'TORTILLA-chicken')] <- 'P'

# car left, squash right = 'Q'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'carsqu-car-L', 'carsqu-squash-R')] <- 'Q'

# squash left, car right = 'R'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'carsqu-squash-L', 'carsqu-car-R')] <- 'R'

# cow left, spoon right = 'S'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'cowspo-cow-L', 'cowspo-spoon-R')] <- 'S'
```

```r
# spoon left, cow right = 'T'
cn_fin$stimulus_set[cn_fin$trial_name_short %in% c(
  'cowspo-spoon-L', 'cowspo-cow-R')] <- 'T'

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

cn_target_df <- cn_fin %>%
  dplyr::select(subject_id, target_noun, stimulus_set,
                target_stimulus,
                noun_pair, post1_target_prop,
                age_centered, bebe_meses, age_group) %>%
  mutate(merge_on_noun = target_stimulus)

cn_non_target_df <- cn_fin %>%
  dplyr::select(subject_id, non_target_noun, stimulus_set,
                non_target_stimulus,
                noun_pair, post1_nontarget_prop,
                age_centered, bebe_meses, age_group) %>%
  mutate(merge_on_noun=non_target_stimulus)

cn_target_nontarget_props_df <- merge(
  cn_target_df, cn_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, noun_pair, stimulus_set, merge_on_noun,
           bebe_meses, age_centered, age_group) %>%
  summarize(post1_target_prop = na.mean(post1_target_prop),
            post1_nontarget_prop = na.mean(post1_nontarget_prop))

cn_diffs_df <-  cn_target_nontarget_props_df %>%
  filter(noun_pair %in% noun_pairs) %>%
  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, noun_pair, bebe_meses, age_centered, age_group) %>%
  summarize(noun_pair_diff = na.mean(stim_pair_diff))

write.csv(cn_diffs_df, here('data/r_analysis_dfs',
                            'cn_bs2012_diffs.csv')
          )
```

```
num_scores_bypp_tab <- cn_diffs_df %>%
  filter(!is.na(noun_pair_diff)) %>%
  group_by(subject_id) %>%
  summarize(n_pairs=n()) %>%
  group_by(n_pairs) %>%
  summarize(n_subs = n())
```

We were able to calculate item-level scores for at least two noun-pairs for all subjects. There were eight subjects with calculable difference scores for all 8 noun-pairs, seven with calculable difference scores for 7/8, two with calculable scores for 6/8, one with calculable scores for 5/8, two with calculable scores for 4/8, and one with calculable scores for only 2/8.

```r
set.seed(36)

by_item_tab <- cn_diffs_df %>%
  group_by(noun_pair, subject_id) %>%
  summarize(subj_mean = na.mean(noun_pair_diff)) %>%
  ungroup() %>%
  group_by(noun_pair) %>%
  summarize(n = n(),
            M = na.mean(subj_mean),
            cil=M-ci.low(subj_mean),
            cih = M+ci.high(subj_mean),
            min=min(subj_mean, na.rm=T),
            max=max(subj_mean, na.rm=T))

ALL_ITEM_MEAN <- mean(by_item_tab$M)
ALL_ITEM_CIL <- ALL_ITEM_MEAN - ci.low(by_item_tab$M)
ALL_ITEM_CIH <- ALL_ITEM_MEAN + ci.high(by_item_tab$M)
ALL_ITEM_MIN <- min(by_item_tab$M)
ALL_ITEM_MAX <- max(by_item_tab$M)
N_ITEMS_POSITIVE <- sum(by_item_tab$M>0)
```

```r
write(apa_table(by_item_tab,
                caption='Experiment 1 Item Means'),
      here('supplement/tables/exp_1/cn_by_item_tab.tex'))

xtable2kable(by_item_tab)
```

Scores By Item.

```
## # A tibble: 8 x 7
##    noun_pair           n       M     cil   cih    min   max
##    <chr>           <int>   <dbl>   <dbl> <dbl>  <dbl> <dbl>
## 1 baby-corn           21 0.0341 -0.160  0.219 -0.794 1
## 2 car-shoe            19 0.146  -0.0816 0.364 -1     1
## 3 chayote-cow         19 0.113  -0.124  0.340 -1     1
## 4 chicken-tortilla    21 0.0401 -0.208  0.299 -0.901 1
## 5 dog-fire            21 0.217   0.0428 0.398 -0.450 1
## 6 horse-soda          21 0.232   0.0304 0.416 -0.563 0.768
## 7 rabbit-soup         21 0.127  -0.0767 0.321 -0.653 0.959
## 8 sheep-water         21 0.0517 -0.108  0.233 -0.606 0.723
```

```r
cn_item_wilcox <- wilcox.test(by_item_tab$M, mu=0,
                              alternative='two.sided')

CN_ITEM_P_WILCOX <- reportP(cn_item_wilcox$p.value)
```

```r
cn_item_binom <- binom.test(8, 8, p=.5)

CN_ITEM_P_BINOM <- reportP(cn_item_binom$p.value)
```

```r
set.seed(36)
# Bayes factor
ttestBF(by_item_tab$M)
```

67   ***Non-parametric tests.***

```
68  ## Bayes factor analysis
69  ## --------------
70  ## [1] Alt., r=0.707 : 17 ±0%
71  ##
72  ## Against denominator:
73  ##    Null, mu = 0
74  ## ---
75  ## Bayes factor type: BFoneSample, JZS
```

```r
# Standard deviation
stdev=sd(by_item_tab$M)
# Mean
mean_data=na.mean(by_item_tab$M)
# Effect size
CN_ITEM_D=abs(mean_data/stdev)
```

76   When implementing this new, more stringent exclusion criteria suggested by R2, 8/8
77 noun-pairs showed a positive mean difference score (*range*: $0.03 - 0.23$, $M = 0.12$ *95% CI*:
78 $[0.07, 0.17]$; $p < .01$, Wilcoxon test, $p < .01$, binomial test, $d = 1.57$).

```r
set.seed(36)

cn_diffs_df$noun_pair_label <- gsub('-', '-\n', cn_diffs_df$noun_pair)
```

```r
cn_diffs_df %>%
  group_by(noun_pair, noun_pair_label, subject_id) %>%
  summarize(subj_mean = na.mean(noun_pair_diff)) %>%
  ungroup() %>%
  group_by(noun_pair, noun_pair_label) %>%
  mutate(n_subjects = n(),
         item_mean = na.mean(subj_mean),
         min=min(subj_mean, na.rm=T),
         max=max(subj_mean, na.rm=T),
         ci.low=item_mean-ci.low(subj_mean),
         ci.high = item_mean+ci.high(subj_mean)) %>%
  ggplot(., aes(
    x=noun_pair_label, y=item_mean, color=noun_pair)) +
  geom_jitter(aes(y=subj_mean), alpha=.5) +
  geom_pointrange(aes(y=item_mean,
                      ymin=ci.low, ymax=ci.high,
                      fill=noun_pair), shape=23, size=0.5) +
  geom_hline(yintercept=0) +
  ylim(-1, 1) +
  sb.density.theme +
  theme(axis.text.x = element_text(
    angle=60, vjust=1, hjust=1, size=11),
        legend.position = 'none',
        axis.title = element_text(
          colour='gray30', size=11, face='bold')) +
  ylab('Increase in Target Looking') +
  xlab('Noun-Pair') +
  scale_discrete_manual(aesthetics=c('color', 'fill'),
                        values=noun_pair_fills)
```
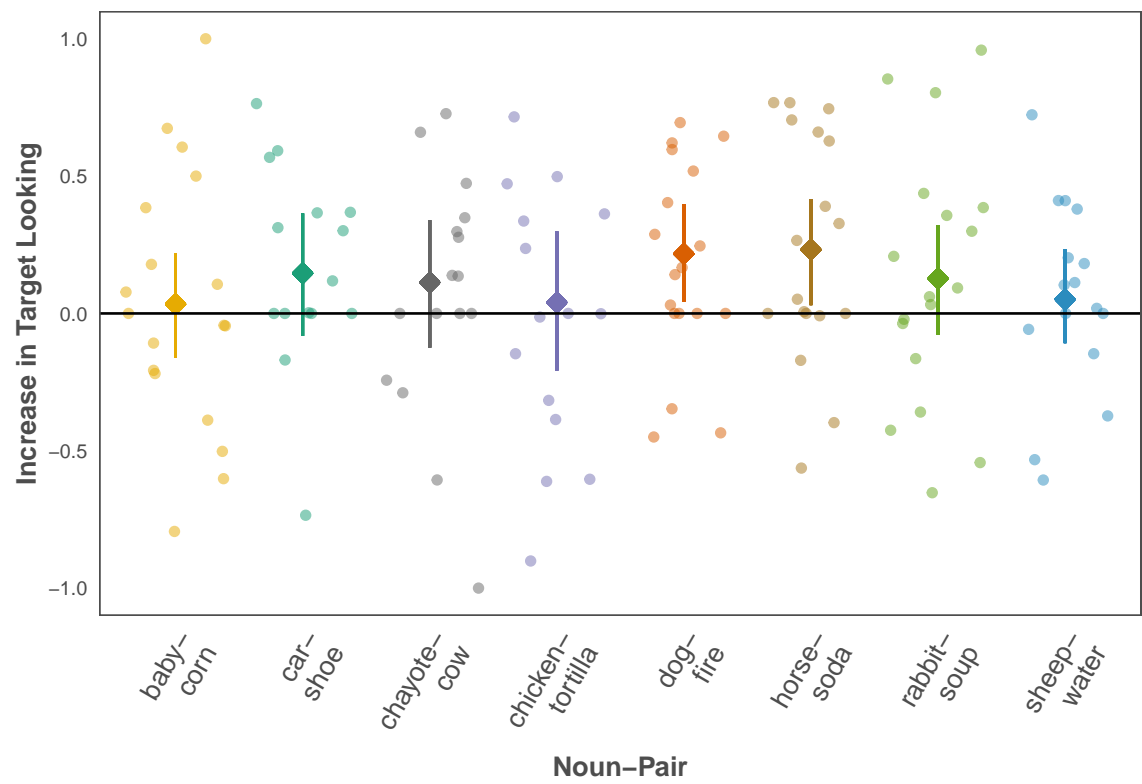
```r
sub_means_tab <- cn_diffs_df %>%
  filter(!is.na(noun_pair_diff)) %>%
  group_by(subject_id, bebe_meses) %>%
  summarize(n = n(),
    subj_mean = na.mean(noun_pair_diff))

positive_scorers_age_tab <- sub_means_tab %>%
  filter(subj_mean>0) %>%
  ungroup() %>%
  summarize(min=min(bebe_meses),
            max=max(bebe_meses),
            mean=mean(bebe_meses),
            sd=sd(bebe_meses),
            ci.low=mean-ci.low(bebe_meses),
            ci.high=mean+ci.high(bebe_meses))

N_SUBS_POSITIVE <- sum(sub_means_tab$subj_mean>0)
PS_MIN_AGE <- positive_scorers_age_tab$min
PS_MAX_AGE <- positive_scorers_age_tab$max
PS_MEAN_AGE <- positive_scorers_age_tab$mean
PS_CIL_AGE <- positive_scorers_age_tab$ci.low
PS_CIH_AGE <- positive_scorers_age_tab$ci.high
PS_SD_AGE <- positive_scorers_age_tab$sd
```

```r
set.seed(36)

 by_sub_overall_tab <- cn_diffs_df %>%
  group_by(subject_id) %>%
  summarize(subj_mean = na.mean(noun_pair_diff)) %>%
  ungroup() %>%
  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))

ALL_SUB_MEAN <- by_sub_overall_tab$group_subject_mean
ALL_SUB_CIL <- by_sub_overall_tab$ci.low
ALL_SUB_CIH <- by_sub_overall_tab$ci.high
ALL_SUB_MIN <- by_sub_overall_tab$min
ALL_SUB_MAX <- by_sub_overall_tab$max
```

80   **Scores By Subject.**

```r
cn_subj_wilcox <- wilcox.test(sub_means_tab$subj_mean, mu=0,
                              alternative='two.sided')

CN_SUBJ_P_WILCOX <- reportP(cn_subj_wilcox$p.value)
```

```r
cn_subj_binom <- binom.test(N_SUBS_POSITIVE, 21, p=.5)

CN_SUBJ_P_BINOM <- reportP(cn_subj_binom$p.value)
```

```r
set.seed(36)
# Bayes factor
ttestBF(sub_means_tab$subj_mean)
```

81   *Non-parametric tests.*

```
## Bayes factor analysis
## --------------
## [1] Alt., r=0.707 : 5.8 ±0%
##
## Against denominator:
##   Null, mu = 0
## ---
## Bayes factor type: BFoneSample, JZS
```

```r
ttestBF(sub_means_tab$subj_mean[sub_means_tab$bebe_meses<10])
```

```
## Bayes factor analysis
## --------------
## [1] Alt., r=0.707 : 0.42 ±0.01%
##
## Against denominator:
##   Null, mu = 0
## ---
## Bayes factor type: BFoneSample, JZS
```

```
ttestBF(sub_means_tab$subj_mean[sub_means_tab$bebe_meses>=10 &
                                sub_means_tab$bebe_meses<14])
```

```
## Bayes factor analysis
## --------------
## [1] Alt., r=0.707 : 2 ±0%
##
## Against denominator:
##    Null, mu = 0
## ---
## Bayes factor type: BFoneSample, JZS
```

```
ttestBF(sub_means_tab$subj_mean[sub_means_tab$bebe_meses>=14])
```

```
## Bayes factor analysis
## --------------
## [1] Alt., r=0.707 : 1.1 ±0.04%
##
## Against denominator:
##    Null, mu = 0
## ---
## Bayes factor type: BFoneSample, JZS
```

```
# Standard deviation
stdev=sd(sub_means_tab$subj_mean)
# Mean
mean_data=mean(sub_means_tab$subj_mean)
# Effect size
CN_SUBJ_D=abs(mean_data/stdev)
```

```
set.seed(36)

model0 <- lmer(noun_pair_diff ~ 0 + (1|subject_id), REML = FALSE,
            cn_diffs_df)
model1 <- lmer(noun_pair_diff ~ 1 + (1|subject_id), REML = FALSE,
            cn_diffs_df)

xtable2kable(summary(model1))
```

MLM Intercept.

```
115  ## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
116  ##    method [lmerModLmerTest]
117  ## Formula: noun_pair_diff ~ 1 + (1 | subject_id)
118  ##    Data: cn_diffs_df
119  ##
120  ##      AIC      BIC   logLik deviance df.resid
121  ##      176      185      -85      170      137
122  ##
123  ## Scaled residuals:
124  ##    Min     1Q Median     3Q    Max
125  ## -2.533 -0.435 -0.107  0.618  2.021
126  ##
127  ## Random effects:
128  ##  Groups     Name        Variance Std.Dev.
129  ##  subject_id (Intercept) 0.00301  0.0548
130  ##  Residual               0.19454  0.4411
131  ## Number of obs: 140, groups:  subject_id, 21
132  ##
133  ## Fixed effects:
134  ##             Estimate Std. Error      df t value Pr(>|t|)
135  ## (Intercept)   0.1218     0.0393 22.9081     3.1    0.005 **
136  ## ---
137  ## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
cn_b0_anova <- anova(model1, model0)
CN_B0_CHISQ <- cn_b0_anova$Chisq[2]
CN_B0_CHI_P <- reportP(cn_b0_anova$`Pr(>Chisq)`[2])

all_subs_intercept <- as.data.frame(cbind(b=fixef(model1),
                           ci.low=confint(model1)[3,1],
                           ci.high=confint(model1)[3,2]))

CN_B0_EST <- all_subs_intercept$b
CN_B0_CIL <- all_subs_intercept$ci.low
CN_B0_CIH <- all_subs_intercept$ci.high

CN_B0_TT_DF <- as.numeric(unlist(summary(model1)['coefficients'])[3])
CN_B0_TT_STAT <- as.numeric(unlist(summary(model1)['coefficients'])[4])
CN_B0_TT_P <- reportP(as.numeric(unlist(summary(model1)['coefficients'])[5]))

#noun pair not enough variance to warrant random intercept
write(texreg(model1),
      file=here('supplement/tables/exp_1/cn_lmer0_tab.tex'))
```

138   The mean of infants' subject-level scores was also positive (*range:* $-0.20 - 0.45$,

139 $M = 0.11$ *95% CI:* $[0.04, 0.18]$, $p < .01$, Wilcoxon; $p < .01$, *Cohen's d = 0.64*), with a
140 majority $(17/21)$ of infants $(M_{age} = 11.19mos, SD_{age} = 2.74mos)$ showing positive means
141 across noun-pairs.

142   A linear mixed effects model with random intercepts for subjects indicates that these
143 results were not driven by a few exceptional infants $(\beta_0 = 0.12$, *95% CI:* $[0.04, 0.20]$,
144 $t(22.91) = 3.10, p < .01; \chi^2(1) = 7.89, p < .01; d = 0.27)$.

```r
set.seed(36)

by_subj_age_tab <- cn_diffs_df %>%
  group_by(subject_id, bebe_meses, age_centered) %>%
  summarize(subj_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_mean-ci.low(noun_pair_diff),
            ci.high = subj_mean+ci.high(noun_pair_diff))

age_cor_test <- cor.test(by_subj_age_tab$age_centered,
                         by_subj_age_tab$subj_mean,
                         method='kendall')

AGE_CORR <- as.numeric(age_cor_test$estimate)
AGE_CORR_P <- reportP(as.numeric(age_cor_test$p.value))
AGE_CORR_Z <- as.numeric(age_cor_test$statistic)
```

145   ***Correlation with Age.***   Subject-level scores exhibited an insignificant positive cor-
146 relation with infants' age in months $(\tau = 0.25, p = 0.116)$.

```r
set.seed(36)

age_cor_test_no15mo <- cor.test(
  by_subj_age_tab[by_subj_age_tab$bebe_meses<15,]$bebe_meses,
  by_subj_age_tab[by_subj_age_tab$bebe_meses<15,]$subj_mean,
  method='kendall')

AGE_OLDEST_PP <- max(by_subj_age_tab$bebe_meses)
SCORE_OLDEST_PP <- op(by_subj_age_tab[
  by_subj_age_tab$bebe_meses==AGE_OLDEST_PP,]$subj_mean)
AGE_CORR_NO15MO <- as.numeric(age_cor_test_no15mo$estimate)
AGE_CORR_P_NO15MO <- reportP(age_cor_test_no15mo$p.value)
AGE_CORR_Z_NO15MO <- as.numeric(age_cor_test_no15mo$statistic)
```

147 Excluding the one (oldest) participant (15.43 months) who showed notably low ev-
148 idence of word recognition (*mean difference score*= 0.02), subject-level scores exhibited a
149 significant ($\alpha = 0.05$) positive correlation with infants' age in months ($\tau = 0.33$, $p < .05$).

```
age_lm  <- lm(subj_mean ~ 1 + age_centered, by_subj_age_tab)
summary(age_lm)
```

```
150 ##
151 ## Call:
152 ## lm(formula = subj_mean ~ 1 + age_centered, data = by_subj_age_tab)
153 ##
154 ## Residuals:
155 ##      Min       1Q   Median       3Q      Max
156 ## -0.24597 -0.14050  0.00887  0.09741  0.26630
157 ##
158 ## Coefficients:
159 ##               Estimate Std. Error t value Pr(>|t|)
160 ## (Intercept)    0.1170     0.0372    3.15   0.0053 **
161 ## age_centered   0.0226     0.0140    1.61   0.1241
162 ## ---
163 ## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
164 ##
165 ## Residual standard error: 0.17 on 19 degrees of freedom
166 ## Multiple R-squared:  0.12,   Adjusted R-squared:  0.0736
167 ## F-statistic: 2.59 on 1 and 19 DF,  p-value: 0.124
```

```
write(texreg(age_lm),
      file=here('supplement/tables/exp_1/cn_age_lm_tab.tex'))
```

168 **Results Section text.**   All (8/8) noun-pairs showed positive mean scores, suggest-
169 ing that the infants in our sample had knowledge of the nouns we tested, despite receiving
170 very little directed speech (*range:*$0.03 - 0.23$, $M = 0.12$, *95% bootstrapped CI:* [0.07, 0.17];
171 $p < .01$, Wilcoxon; $p < .01$, binomial test; $d = 1.57$).

172 The mean of infants' subject-level scores was also positive (*range:* $-0.20 - 0.45$,
173 $M = 0.11$, *95% CI:* [0.04, 0.18]; $p < .01$, Wilcoxon; $p < .01$, binomial; $d = 0.64$), with a
174 majority (17/21) of infants ($M_{\text{age}} = 11.19mos$, $SD_{\text{age}} = 2.74mos$) showing positive means
175 across noun-pairs (see Figure in Supplement).

176 A linear mixed effects model with random intercepts for subjects[1] indicates that
177 these results were not driven by a few exceptional infants ($\beta_0 = 0.12$, *95% CI:* [0.04, 0.20],
178 $t(22.91) = 3.10$, $p < .01$; $\chi^2(1) = 7.89$, $p < .01$; $d = 0.27$).

---
[1]A model which additionally included random intercepts for item failed to converge.

```r
cn_diffs_df %>%
  filter(!is.na(noun_pair_diff)) %>%
  group_by(subject_id, bebe_meses) %>%
  summarize(subj_mean = na.mean(noun_pair_diff),
            n_items = n()) %>%
  filter(!is.na(subj_mean)) %>%
ggplot(.) +
  geom_point(aes(x=bebe_meses, y=subj_mean),
             color=sheeppink, alpha=.95, shape=21, size=2.5,
             fill=sheeppink) +
  geom_linerange(aes(x=bebe_meses,
                     ymin=0, ymax=subj_mean),
                 color=sheeppink) +
  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)
```

```r
ggsave(here('supplement/plots/exp_1/pdfs', 'cn_meandiffs_bysub.pdf'),
       device='pdf', width=2.5, height=1.5, units='in', scale=2)
ggsave(here('supplement/plots/exp_1/pngs', 'cn_meandiffs_bysub.png'),
       device='png', width=2.5, height=1.5, units='in', scale=2)
```
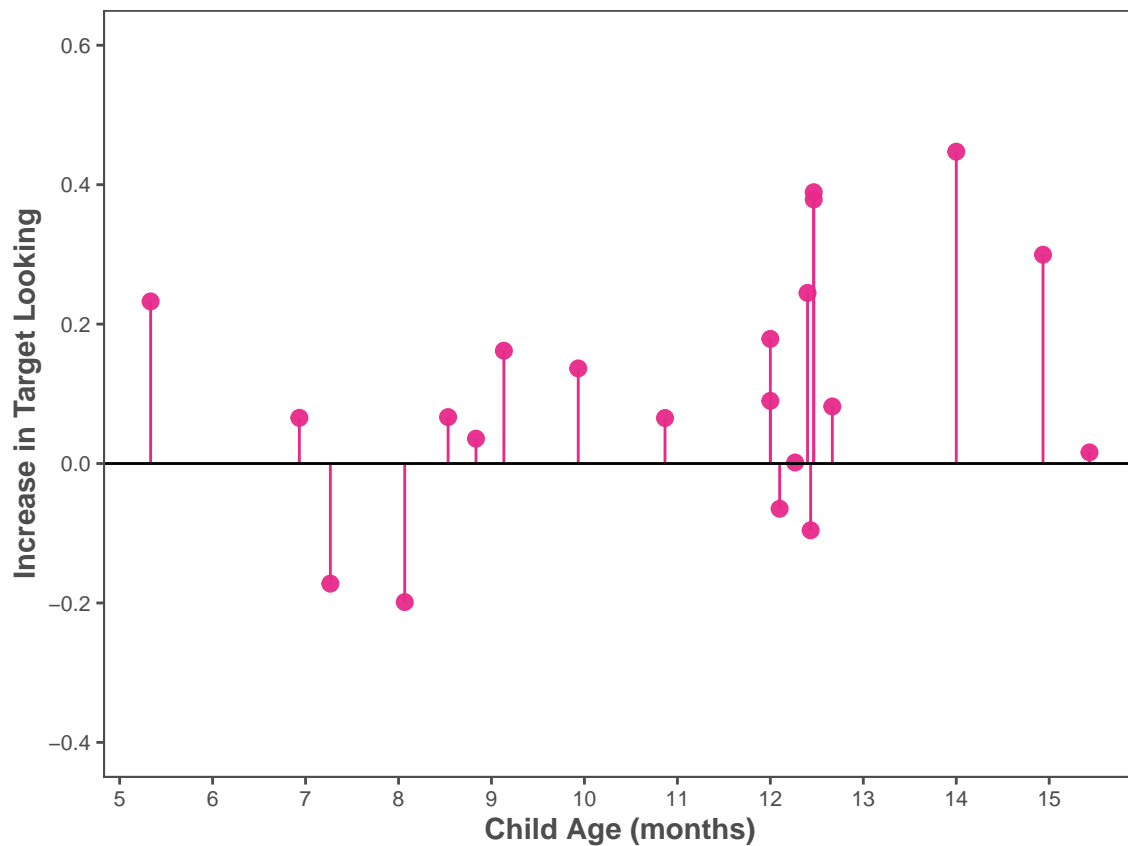
Figure 1

```
bysub_agegroup_means_df <-
  cn_diffs_df %>%
  group_by(subject_id, age_group) %>%
  summarize(subj_mean = na.mean(noun_pair_diff)) %>%
  group_by(age_group) %>%
  summarize(age_group_M=na.mean(subj_mean))

CN_SUBJ_MDIFF_6MOS <-
  bysub_agegroup_means_df$age_group_M[
    bysub_agegroup_means_df$age_group == '5-9 months']
CN_SUBJ_MDIFF_10MOS <-
  bysub_agegroup_means_df$age_group_M[
    bysub_agegroup_means_df$age_group == '10-13 months']
CN_SUBJ_MDIFF_14MOS <-
  bysub_agegroup_means_df$age_group_M[
    bysub_agegroup_means_df$age_group == '14-16 months']
```

```r
byitem_agegroup_means_df <- cn_diffs_df %>%
  filter(noun_pair %in% noun_pairs) %>%
  group_by(age_group, noun_pair) %>%
  summarize(item_mean = na.mean(noun_pair_diff)) %>%
  group_by(age_group) %>%
  summarize(n = n(),
            age_group_M = na.mean(item_mean),
            cil = age_group_M - ci.low(item_mean),
            cih = age_group_M + ci.high(item_mean),
            min = min(item_mean, na.rm=T),
            max = max(item_mean, na.rm=T))

CN_ITEM_MDIFF_6MOS <-
  byitem_agegroup_means_df$age_group_M[
    byitem_agegroup_means_df$age_group == '5-9 months']
CN_ITEM_MDIFF_10MOS <-
  byitem_agegroup_means_df$age_group_M[
    byitem_agegroup_means_df$age_group == '10-13 months']
CN_ITEM_MDIFF_14MOS <-
  byitem_agegroup_means_df$age_group_M[
    byitem_agegroup_means_df$age_group == '14-16 months']
```

179 **Comparison to B&S 2012.** The evidence that Tseltal infants possessed knowledge
180 of the words we probed is similarly strong to evidence that has been used to infer the
181 presence of word knowledge among U.S. infants. For example, Bergelson and Swingley
182 (2012) reported mean scores of 0.074 (over subjects) and 0.065 (over items) for 6- to 9-
183 month–olds, 0.055 (over subjects) and 0.059 (over items) for 10- to 13-month–olds, and 0.29
184 (over subjects) and 0.28 (over items) for 14-16–month-olds.

185 We found: mean scores of 0.04 (over subjects) and 0.03 (over items) for 6- to 9-
186 month–olds, 0.13 (over subjects) and 0.12 (over items) for 10- to 13-month–olds, and 0.25
187 (over subjects) and 0.26 (over items) for 14- to 16-month–olds.

EXCLUSION CRITERIA 2 (EC2)

189 **Replicating Mean Difference Analysis (EC2)**

190 *Including only infants affording calculation of at least half (4) of* **difference scores**
191 *(rather than contributing half of all trials)*

192 *Under this criterion, total sample includes:*

193 _- 8 infants with calculable difference scores for all 8 noun-pairs_
194 _- 7 infants with calculable difference scores for 7/8_
195 _- 2 infants with calculable scores for 6/8_

196 _- 1 infant with a calculable score for 5/8_
197 _- 2 infants with calculable scores for 4/8_

198   *- 1 EXCLUDED participant with calculable scores for only 2/8 (fewer than half of the*
199 *set of noun-pairs).*

200   *N=20*

```r
set.seed(36)

by_item_tab2 <- cn_diffs_df %>%
  filter(subject_id!='J94252') %>%
  group_by(noun_pair, subject_id) %>%
  summarize(subj_mean = na.mean(noun_pair_diff)) %>%
  ungroup() %>%
  group_by(noun_pair) %>%
  summarize(n_subjects = n(),
            item_mean = na.mean(subj_mean),
            median = median(subj_mean, na.rm=T),
            ci.low=item_mean-ci.low(subj_mean),
            ci.high = item_mean+ci.high(subj_mean),
            min=min(subj_mean, na.rm=T),
            max=max(subj_mean, na.rm=T))

ALL_ITEM_MEAN2 <- mean(by_item_tab2$item_mean)
ALL_ITEM_CIL2 <- ALL_ITEM_MEAN - ci.low(by_item_tab2$item_mean)
ALL_ITEM_CIH2 <- ALL_ITEM_MEAN + ci.high(by_item_tab2$item_mean)
ALL_ITEM_MIN2 <- min(by_item_tab2$item_mean)
ALL_ITEM_MAX2 <- max(by_item_tab2$item_mean)
N_ITEMS_POSITIVE2 <- sum(by_item_tab2$item_mean>0)
```

```r
set.seed(36)

by_item_tab2 <- cn_diffs_df %>%
  filter(subject_id!='J94252') %>%
  group_by(noun_pair, subject_id) %>%
  summarize(subj_mean = na.mean(noun_pair_diff)) %>%
  ungroup() %>%
  group_by(noun_pair) %>%
  summarize(n = n(),
            M = na.mean(subj_mean),
            cil=M-ci.low(subj_mean),
            cih = M+ci.high(subj_mean),
            min=min(subj_mean, na.rm=T),
            max=max(subj_mean, na.rm=T))

xtable2kable(by_item_tab2)
```

201   **Scores By Item (EC2).**

```
202  ## # A tibble: 8 x 7
203  ##   noun_pair          n      M     cil    cih     min    max
204  ##   <chr>          <int>  <dbl>   <dbl>  <dbl>   <dbl>  <dbl>
205  ## 1 baby-corn         20 0.0359 -0.183  0.251  -0.794  1
206  ## 2 car-shoe          18 0.146  -0.0860 0.359  -1      1
207  ## 3 chayote-cow       18 0.113  -0.108  0.335  -1      1
208  ## 4 chicken-tortilla  20 0.0401 -0.205  0.291  -0.901  1
209  ## 5 dog-fire          20 0.217   0.0374 0.394  -0.450  1
210  ## 6 horse-soda        20 0.269   0.0896 0.437  -0.563  0.768
211  ## 7 rabbit-soup       20 0.127  -0.0850 0.348  -0.653  0.959
212  ## 8 sheep-water       20 0.0517 -0.111  0.214  -0.606  0.723
```

```r
cn_item_wilcox2 <- wilcox.test(by_item_tab2$M, mu=0,
                         alternative='two.sided')

CN_ITEM_P_WILCOX2 <- reportP(cn_item_wilcox2$p.value)
```

```r
cn_item_binom2 <- binom.test(N_ITEMS_POSITIVE2, 8, p=.5)

CN_ITEM_P_BINOM2 <- reportP(cn_item_binom2$p.value)
```

```
set.seed(36)
# Bayes factor
ttestBF(by_item_tab2$M)
```

213   *Non-parametric Tests (EC2).*

```
214   ## Bayes factor analysis
215   ## --------------
216   ## [1] Alt., r=0.707 : 13 ±0%
217   ##
218   ## Against denominator:
219   ##   Null, mu = 0
220   ## ---
221   ## Bayes factor type: BFoneSample, JZS
```

```
# Standard deviation
stdev=sd(by_item_tab2$M)
# Mean
mean_data=mean(by_item_tab2$M)
# Effect size
CN_ITEM_D2=abs(mean_data/stdev)
```

222   When implementing this new, even more stringent exclusion criteria (leading us to
223 drop an entire participant's data), 8/8 noun-pairs showed a positive mean difference score
224 $(0.04 - 0.27$, $M = 0.12$ *95% CI:* [0.06, 0.17]; $p < .01$, Wilcoxon test; $p < .01$, binomial test,
225 $d = 1.47$).

```r
sub_means_tab2 <- cn_diffs_df %>%
  filter(!is.na(noun_pair_diff),
         subject_id != 'J94252') %>%
  group_by(subject_id, bebe_meses) %>%
  summarize(n = n(),
    subj_mean = na.mean(noun_pair_diff))

positive_scorers_age_tab2 <- sub_means_tab2 %>%
  filter(subj_mean>0) %>%
  ungroup() %>%
  summarize(min=min(bebe_meses),
            max=max(bebe_meses),
            mean=mean(bebe_meses),
            sd=sd(bebe_meses),
            ci.low=mean-ci.low(bebe_meses),
            ci.high=mean+ci.high(bebe_meses))

N2 <- length(unique(sub_means_tab2$subject_id))

N_SUBS_POSITIVE2 <- sum(sub_means_tab2$subj_mean>0)
PS_MIN_AGE2 <- positive_scorers_age_tab2$min
PS_MAX_AGE2 <- positive_scorers_age_tab2$max
PS_MEAN_AGE2 <- positive_scorers_age_tab2$mean
PS_CIL_AGE2 <- positive_scorers_age_tab2$ci.low
PS_CIH_AGE2 <- positive_scorers_age_tab2$ci.high
PS_SD_AGE2 <- positive_scorers_age_tab2$sd
```

Scores By Subject (EC2). 17/20 subjects showed a positive mean difference score (*range*: $-10.10yrs$, $M_{age} = 11.19$ *95% CI:* $[9.90, 12.33]$, $SD_{age} = 2.74$).

```r
set.seed(36)

 by_sub_overall_tab2 <- cn_diffs_df %>%
  filter(subject_id != 'J94252') %>%
  group_by(subject_id) %>%
  summarize(subj_mean = na.mean(noun_pair_diff)) %>%
  ungroup() %>%
  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))
```

```
ALL_SUB_MEAN2 <- by_sub_overall_tab2$group_subject_mean
ALL_SUB_CIL2 <- by_sub_overall_tab2$ci.low
ALL_SUB_CIH2 <- by_sub_overall_tab2$ci.high
ALL_SUB_MIN2 <- by_sub_overall_tab2$min
ALL_SUB_MAX2 <- by_sub_overall_tab2$max
```

```r
set.seed(36)

#noun pair not enough variance to warrant random intercept
model0_2 <- lmer(noun_pair_diff ~ 0 + (1|subject_id), REML = FALSE,
             cn_diffs_df[cn_diffs_df$subject_id!='J94252',])
model1_2 <- lmer(noun_pair_diff ~ 1 + (1|subject_id), REML = FALSE,
               cn_diffs_df[cn_diffs_df$subject_id!='J94252',])

xtable2kable(summary(model1_2))
```

228   *MLM Intercept (EC2).*

```
229 ## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
230 ##   method [lmerModLmerTest]
231 ## Formula: noun_pair_diff ~ 1 + (1 | subject_id)
232 ##    Data: cn_diffs_df[cn_diffs_df$subject_id != "J94252", ]
233 ##
234 ##      AIC      BIC   logLik deviance df.resid
235 ##      174      183      -84      168      135
236 ##
237 ## Scaled residuals:
238 ##     Min      1Q  Median      3Q     Max
239 ## -2.5378 -0.4097 -0.0874  0.6331  2.0047
240 ##
241 ## Random effects:
242 ##  Groups     Name        Variance Std.Dev.
243 ##  subject_id (Intercept) 0.00287  0.0535
244 ##  Residual               0.19546  0.4421
245 ## Number of obs: 138, groups:  subject_id, 20
246 ##
247 ## Fixed effects:
248 ##             Estimate Std. Error      df t value Pr(>|t|)
249 ## (Intercept)   0.1269     0.0395 22.3564    3.21    0.004 **
250 ## ---
251 ## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
cn_b0_anova_2 <- anova(model1_2, model0_2)
CN_B0_CHISQ2 <- cn_b0_anova_2$Chisq[2]
CN_B0_CHI_P2 <- reportP(cn_b0_anova_2$`Pr(>Chisq)`[2])

all_subs_intercept2 <- as.data.frame(cbind(b=fixef(model1_2),
                              ci.low=confint(model1_2)[3,1],
                              ci.high=confint(model1_2)[3,2]))

CN_B0_EST2 <- all_subs_intercept2$b
CN_B0_CIL2 <- all_subs_intercept2$ci.low
CN_B0_CIH2 <- all_subs_intercept2$ci.high

CN_B0_TT_DF2 <- as.numeric(unlist(summary(model1_2)['coefficients']))[3])
CN_B0_TT_STAT2 <- as.numeric(unlist(summary(model1_2)['coefficients']))[4])
CN_B0_TT_P2 <- reportP(as.numeric(unlist(summary(model1_2)['coefficients']))[5]))
```

```r
cn_subj_wilcox2 <- wilcox.test(sub_means_tab2$subj_mean, mu=0,
                              alternative='two.sided')

CN_SUBJ_P_WILCOX2 <- reportP(cn_subj_wilcox2$p.value)
```

```r
cn_subj_binom2 <- binom.test(N_SUBS_POSITIVE2, N2, p=0.5)

CN_SUBJ_P_BINOM2 <- reportP(cn_subj_binom2$p.value)
```

```r
set.seed(36)
# Bayes factor
ttestBF(sub_means_tab2$subj_mean)
```

252   *Non-parametric tests (EC2).*

253   ## Bayes factor analysis
254   ## --------------
255   ## [1] Alt., r=0.707 : 16 ±0%
256   ##
257   ## Against denominator:
258   ##   Null, mu = 0
259   ## ---
260   ## Bayes factor type: BFoneSample, JZS

```r
# Standard deviation
stdev=sd(sub_means_tab2$subj_mean)
# Mean
mean_data=mean(sub_means_tab2$subj_mean)
# Effect size
CN_SUBJ_D2=abs(mean_data/stdev)
```

261   Using this alternative interpretation of the reviewer's suggested exclusion criteria, the
262 mean across all 20 subjects was positive (*range: $-0.17 - 0.45$, M = 0.13 95% CI:* [0.06,
263 0.20]; $p < .01$, Wilcoxon; $p < .01$, binomial, $d = 0.77$).

264   A linear mixed effects model with random intercepts for subjects suggests that these
265 results were reliable across infants ($\beta_0 = 0.13$, *95% CI:* [0.05, 0.21], $t(22.36) = 3.21$, $p < .01$;
266 $\chi^2(1) = 8.45$, $p < .01$; $d = 0.28$).

```r
set.seed(36)

by_subj_age_tab2 <- cn_diffs_df %>%
  filter(subject_id != 'J94252') %>%
  group_by(subject_id, bebe_meses, age_centered) %>%
  summarize(subj_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_mean-ci.low(noun_pair_diff),
            ci.high = subj_mean+ci.high(noun_pair_diff))

age_cor_test2 <- cor.test(by_subj_age_tab2$age_centered,
                          by_subj_age_tab2$subj_mean,
                          method='kendall')

AGE_CORR2 <- as.numeric(age_cor_test2$estimate)
AGE_CORR_P2 <- as.numeric(age_cor_test2$p.value)
AGE_CORR_Z2 <- as.numeric(age_cor_test2$statistic)
```

267   ***Replicating Correlation with Age (EC2).***   Using this alternate exclusion crite-
268 ria, children's age in months exhibited an insignificant positive correlation with their mean
269 difference scores across items ($\tau = 0.20$, $p = 0.22$).

```r
age_lm2  <- lm(subj_mean ~ 1 + age_centered, by_subj_age_tab2)
summary(age_lm2)
```

270   ##

```
271  ## Call:
272  ## lm(formula = subj_mean ~ 1 + age_centered, data = by_subj_age_tab2)
273  ##
274  ## Residuals:
275  ##       Min        1Q    Median        3Q       Max
276  ## -0.24652  -0.09133  -0.00413   0.09738   0.26970
277  ##
278  ## Coefficients:
279  ##               Estimate Std. Error t value Pr(>|t|)
280  ## (Intercept)     0.1290     0.0365    3.53   0.0024 **
281  ## age_centered    0.0171     0.0139    1.23   0.2335
282  ## ---
283  ## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
284  ##
285  ## Residual standard error: 0.16 on 18 degrees of freedom
286  ## Multiple R-squared:  0.0779, Adjusted R-squared:  0.0266
287  ## F-statistic: 1.52 on 1 and 18 DF,  p-value: 0.234
```

288   **Pre/Post Looking Logit Model**

289   *Dueling exclusion criteria are not relevant here, as even participant 'J94252' con-*
290 *tributed at least half (16) of all test trials.*

```
cn_supp_pre <- cn_fin %>%
  select(pre_target_sum_ms, pre_nontarget_sum_ms,
         subject_id, bebe_meses, age_centered, noun_pair)

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

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

cn_supp_post <- cn_fin %>%
  select(post1_target_sum_ms, post1_nontarget_sum_ms,
         subject_id, bebe_meses, age_centered, noun_pair)

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

cn_supp_post <- cn_supp_post %>%
  select(subject_id, bebe_meses, age_centered,
```

```
          noun_pair, phase, target_bins, nontarget_bins)

cn_supp_stacked_r2 <- rbind(cn_supp_pre, cn_supp_post)

cn_supp_stacked_r2$phase <- as.factor(cn_supp_stacked_r2$phase)
cn_supp_stacked_r2$phase <- relevel(cn_supp_stacked_r2$phase, ref='pre-naming')

cn_supp_stacked_r2 %>%
  group_by(subject_id, bebe_meses, age_centered, noun_pair, phase) %>%
  reframe(target_bins = target_bins,
          nontarget_bins = nontarget_bins) %>%
write.csv(., here('data/r_analysis_dfs',
                  'cn_prepost_target_looking.csv')
         )
```

```
set.seed(36)
cn_supp_glmer_null <- glmer(cbind(target_bins, nontarget_bins) ~
                        1 + (1|subject_id) + (1|noun_pair),
                      family=binomial, cn_supp_stacked_r2)

cn_supp_glmer <- glmer(cbind(target_bins, nontarget_bins) ~
                        phase + (1|subject_id) + (1|noun_pair),
                      family=binomial, cn_supp_stacked_r2)

summary(cn_supp_glmer)
```

```
291  ## Generalized linear mixed model fit by maximum likelihood (Laplace
292  ##   Approximation) [glmerMod]
293  ##  Family: binomial  ( logit )
294  ## Formula: cbind(target_bins, nontarget_bins) ~ phase + (1 | subject_id) +
295  ##     (1 | noun_pair)
296  ##    Data: cn_supp_stacked_r2
297  ##
298  ##      AIC      BIC   logLik deviance df.resid
299  ##   114954   114973   -57473   114946     1052
300  ##
301  ## Scaled residuals:
302  ##     Min      1Q  Median      3Q     Max
303  ## -20.196  -7.622   0.822   7.850  19.954
304  ##
305  ## Random effects:
306  ##  Groups     Name        Variance Std.Dev.
307  ##  subject_id (Intercept) 0.0327   0.181
308  ##  noun_pair  (Intercept) 0.0234   0.153
```

```
309  ## Number of obs: 1056, groups:  subject_id, 21; noun_pair, 8
310  ##
311  ## Fixed effects:
312  ##                   Estimate Std. Error z value Pr(>|z|)
313  ## (Intercept)        0.02640    0.06726    0.39     0.69
314  ## phasepost-naming  0.24900    0.00983   25.32   <2e-16 ***
315  ## ---
316  ## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
317  ##
318  ## Correlation of Fixed Effects:
319  ##             (Intr)
320  ## phspst-nmng -0.059
```

```r
cn_supp_glmer_tab <- as.data.frame(
  cbind('OR'=op(exp(fixef(cn_supp_glmer))),
        'CIL'=op(exp(confint(cn_supp_glmer))[3:4,1]),
        'CIH'=op(exp(confint(cn_supp_glmer))[3:4,2])))

POSTNAMING_OR <- cn_supp_glmer_tab$OR[2]
POSTNAMING_CIL <- cn_supp_glmer_tab$CIL[2]
POSTNAMING_CIH <- cn_supp_glmer_tab$CIH[2]
PHASE_WALD_CHISQ <- Anova(cn_supp_glmer)['phase','Chisq']
PHASE_WALD_P <- reportP(Anova(cn_supp_glmer)['phase','Pr(>Chisq)'])

anova(cn_supp_glmer_null, cn_supp_glmer)
```

```
321  ## Data: cn_supp_stacked_r2
322  ## Models:
323  ## cn_supp_glmer_null: cbind(target_bins, nontarget_bins) ~ 1 + (1 | subject_id) + (1 | nou
324  ## cn_supp_glmer: cbind(target_bins, nontarget_bins) ~ phase + (1 | subject_id) + (1 | noun_
325  ##                    npar    AIC    BIC logLik deviance Chisq Df Pr(>Chisq)
326  ## cn_supp_glmer_null    3 115594 115609 -57794    115588
327  ## cn_supp_glmer         4 114953 114973 -57473    114945   643  1      <2e-16 ***
328  ## ---
329  ## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
Anova(cn_supp_glmer)
```

```
330  ## Analysis of Deviance Table (Type II Wald chisquare tests)
331  ##
332  ## Response: cbind(target_bins, nontarget_bins)
333  ##       Chisq Df Pr(>Chisq)
334  ## phase   641  1     <2e-16 ***
335  ## ---
336  ## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

337   Model summary shows that model is fit on 1056 observations from 21 subjects (number
338 of trials = 1056/2 = 528, as model is fit to data in 'long' form, where two rows for each
339 trial: pre-naming and post-naming).

```
cn_supp_stacked_r2$ratio <- cn_supp_stacked_r2$target_bins/
  cn_supp_stacked_r2$nontarget_bins

cn_supp_age_glmer <- glmer(cbind(target_bins, nontarget_bins) ~
                            phase + age_centered +
                              (1|subject_id) + (1|noun_pair),
                           family=binomial, cn_supp_stacked_r2)

cn_pp_age_glmer_tab <- as.data.frame(
  cbind('OR'=op(exp(fixef(cn_supp_age_glmer))),
        'CIL'=op(exp(confint(cn_supp_age_glmer))[3:5,1]),
        'CIH'=op(exp(confint(cn_supp_age_glmer))[3:5,2])))

anova(cn_supp_glmer, cn_supp_age_glmer)
```

340 ## Data: cn_supp_stacked_r2
341 ## Models:
342 ## cn_supp_glmer: cbind(target_bins, nontarget_bins) ~ phase + (1 | subject_id) + (1 | noun
343 ## cn_supp_age_glmer: cbind(target_bins, nontarget_bins) ~ phase + age_centered + (1 | subj
344 ##   npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
345 ## cn_supp_glmer   4 114953 114973 -57473   114945
346 ## cn_supp_age_glmer   5 114955 114980 -57473   114945   0.1 1   0.75

```
Anova(cn_supp_age_glmer)
```

347 ## Analysis of Deviance Table (Type II Wald chisquare tests)
348 ##
349 ## Response: cbind(target_bins, nontarget_bins)
350 ##   Chisq Df Pr(>Chisq)
351 ## phase   641.1 1   <2e-16 ***
352 ## age_centered   0.1 1   0.75
353 ## ---
354 ## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```
cn_supp_ageint_glmer <- glmer(cbind(target_bins, nontarget_bins) ~
                            phase*age_centered +
                              (1|subject_id) + (1|noun_pair),
                           family=binomial, cn_supp_stacked_r2)
```

```r
cn_supp_ageint_glmer_tab <- as.data.frame(
  cbind('OR'=op(exp(fixef(cn_supp_ageint_glmer))),
        'CIL'=op(exp(confint(cn_supp_ageint_glmer))[3:6,1]),
        'CIH'=op(exp(confint(cn_supp_ageint_glmer))[3:6,2])))

## reporting effect of trial phase
POST_AGE_INT_PHASE_OR <- cn_supp_ageint_glmer_tab$OR[2]
POST_AGE_INT_PHASE_CIL <- cn_supp_ageint_glmer_tab$CIL[2]
POST_AGE_INT_PHASE_CIH <- cn_supp_ageint_glmer_tab$CIH[2]
POST_AGE_INT_PHASE_WALD_CHISQ <-
  Anova(cn_supp_ageint_glmer)['phase','Chisq']
POST_AGE_INT_PHASE_WALD_P <-
  reportP(Anova(cn_supp_ageint_glmer)['phase','Pr(>Chisq)'])

## reporting interaction between phase and age
POST_AGE_INT_OR <- cn_supp_ageint_glmer_tab$OR[4]
POST_AGE_INT_CIL <- cn_supp_ageint_glmer_tab$CIL[4]
POST_AGE_INT_CIH <- cn_supp_ageint_glmer_tab$CIH[4]
POST_AGE_INT_WALD_CHISQ <-
  Anova(cn_supp_ageint_glmer)['phase:age_centered','Chisq']
POST_AGE_INT_WALD_P <-
  reportP(Anova(cn_supp_ageint_glmer)['phase:age_centered','Pr(>Chisq)'])

## comparing interaction model to model with phase alone
cn_supp_ageint_anova <- anova(cn_supp_glmer, cn_supp_ageint_glmer)

POST_AGE_INT_DF <- cn_supp_ageint_anova$Df[2]
POST_AGE_INT_CHISQ <- cn_supp_ageint_anova$Chisq[2]
POST_AGE_INT_P <- reportP(cn_supp_ageint_anova$`Pr(>Chisq)`[2])

## print summary, model comparison, type 2 tests
summary(cn_supp_ageint_glmer)
```

```
355  ## Generalized linear mixed model fit by maximum likelihood (Laplace
356  ##    Approximation) [glmerMod]
357  ##  Family: binomial  ( logit )
358  ## Formula: cbind(target_bins, nontarget_bins) ~ phase * age_centered + (1 |
359  ##     subject_id) + (1 | noun_pair)
360  ##    Data: cn_supp_stacked_r2
361  ##
362  ##      AIC      BIC   logLik deviance df.resid
363  ##   114556   114585   -57272   114544     1050
364  ##
365  ## Scaled residuals:
```

```
366  ##     Min      1Q  Median      3Q     Max
367  ## -19.818  -7.472   0.582   7.865  19.878
368  ##
369  ## Random effects:
370  ##  Groups     Name         Variance Std.Dev.
371  ##  subject_id (Intercept) 0.0320   0.179
372  ##  noun_pair  (Intercept) 0.0235   0.153
373  ## Number of obs: 1056, groups:  subject_id, 21; noun_pair, 8
374  ##
375  ## Fixed effects:
376  ##                                Estimate Std. Error z value Pr(>|z|)
377  ## (Intercept)                     0.02609    0.06713    0.39    0.698
378  ## phasepost-naming                0.25195    0.00985   25.58   <2e-16 ***
379  ## age_centered                   -0.02692    0.01499   -1.80    0.072 .
380  ## phasepost-naming:age_centered  0.07640    0.00381   20.03   <2e-16 ***
381  ## ---
382  ## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
383  ##
384  ## Correlation of Fixed Effects:
385  ##            (Intr) phsps- ag_cnt
386  ## phspst-nmng -0.059
387  ## age_centerd  0.045 -0.001
388  ## phspst-nm:_ -0.001  0.018 -0.106
```

```
anova(cn_supp_glmer, cn_supp_ageint_glmer)
```

```
389  ## Data: cn_supp_stacked_r2
390  ## Models:
391  ## cn_supp_glmer: cbind(target_bins, nontarget_bins) ~ phase + (1 | subject_id) + (1 | noun
392  ## cn_supp_ageint_glmer: cbind(target_bins, nontarget_bins) ~ phase * age_centered + (1 | s
393  ##                       npar    AIC     BIC logLik deviance Chisq Df Pr(>Chisq)
394  ## cn_supp_glmer           4 114953 114973 -57473    114945
395  ## cn_supp_ageint_glmer    6 114556 114585 -57272    114544   402  2     <2e-16 ***
396  ## ---
397  ## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Anova(cn_supp_ageint_glmer)
```

```
398  ## Analysis of Deviance Table (Type II Wald chisquare tests)
399  ##
400  ## Response: cbind(target_bins, nontarget_bins)
401  ##               Chisq Df Pr(>Chisq)
402  ## phase         635.8  1     <2e-16 ***
403  ## age_centered    0.1  1       0.75
```

```
404  ## phase:age_centered 401.0  1      <2e-16 ***
405  ## ---
406  ## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
write(texreg(list(cn_supp_glmer, cn_supp_age_glmer)),
      here('supplement/tables/exp_1', 'cn_prepost_glmers_w_wo_age_tab.tex')
      )


write(texreg(list(cn_supp_glmer, cn_supp_age_glmer, cn_supp_ageint_glmer)),
      here('supplement/tables/exp_1', 'cn_prepost_3glmers_tab.tex')
      )
```

407   The odds ratio for trial phase (POST-NAMING *OR*= 1.28, *95% CI:* [1.26, 1.31]) in-
408 dicates that infants dedicated a significantly greater share of their visual attention to the
409 target image *after* hearing it labeled than *before* hearing it labeled, controlling for subject–
410 and item–level variability (*Wald* $\chi^2(1) = 641.07$, $p < .001$, $d = 1.05$; see Supplement).

411   A model which additionally included infant age and its interaction with trial phase
412 resulted in a significantly better fit ($\chi^2(2) = 401.95$, $p < .001$), showing a reliable effect of
413 trial phase (POST-NAMING *OR*= 1.29, *95% CI:* [1.26, 1.31], *Wald* $\chi^2(1) = 635.84$, $p < .001$,
414 $d = 1.07$) and interaction with age, such that older children showed a greater increase in
415 the ratio of targe to non-target looking after hearing the target word (*OR*= 1.08, *95% CI:*
416 [1.07, 1.09], *Wald* $\chi^2(1) = 401.02$, $p < .001$, $d = 0.32$).

417   ***GLMERs by Age Group.***   *From B&S 2012:*

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

```
cn_phase_6_9_glmer <- glmer(cbind(target_bins, nontarget_bins) ~
                     phase + (1|subject_id) + (1|noun_pair),
                   family=binomial, cn_supp_stacked_r2[
                    cn_supp_stacked_r2$bebe_meses<10,])


cn_phase_10_13_glmer <- glmer(cbind(target_bins, nontarget_bins) ~
                     phase + (1|subject_id) + (1|noun_pair),
                   family=binomial, cn_supp_stacked_r2[
                     cn_supp_stacked_r2$bebe_meses>=10 &
                       cn_supp_stacked_r2$bebe_meses<14,])


cn_phase_14_16_glmer <- glmer(cbind(target_bins, nontarget_bins) ~
```

```
                            phase + (1|subject_id) + (1|noun_pair),
                        family=binomial, cn_supp_stacked_r2[
                          cn_supp_stacked_r2$bebe_meses>=14 &
                            cn_supp_stacked_r2$bebe_meses<17,])
```

```
write(texreg(list(
  cn_phase_6_9_glmer, cn_phase_10_13_glmer, cn_phase_14_16_glmer),
  custom.model.names=c('5-9 months', '10-13 months', '14-16 months')),
  here('supplement/tables/exp_1/age_binned', 'cn_prepost_agegroup_glmers.tex')
  ) # 3 columns for age

# individual tables for each age group
phase_glmer_6_9mos_tab  <-
  cbind('Estimate'=summary(cn_phase_6_9_glmer)$coefficients[,'Estimate'],
      'Standard Error'= summary(cn_phase_6_9_glmer)$coefficients[,'Std. Error'],
      'P value'=summary(cn_phase_6_9_glmer)$coefficients[,'Pr(>|z|)'])

phase_glmer_10_13mos_tab  <-
  cbind('Estimate'=summary(cn_phase_10_13_glmer)$coefficients[,'Estimate'],
      'Standard Error'= summary(cn_phase_10_13_glmer)$coefficients[,'Std. Error'],
      'P value'=summary(cn_phase_10_13_glmer)$coefficients[,'Pr(>|z|)'])

phase_glmer_14_16mos_tab  <-
  cbind('Estimate'=summary(cn_phase_14_16_glmer)$coefficients[,'Estimate'],
      'Standard Error'= summary(cn_phase_14_16_glmer)$coefficients[,'Std. Error'],
      'P value'=summary(cn_phase_14_16_glmer)$coefficients[,'Pr(>|z|)'])

# write to tex
write(apa_table(phase_glmer_6_9mos_tab,
                caption='5--9 month-olds'),
      here('supplement/tables/exp_1/age_binned', 'cn_phase_6mos_tab.tex')
      )

write(apa_table(phase_glmer_10_13mos_tab,
                caption='10--13 month-olds'),
      here('supplement/tables/exp_1/age_binned', 'cn_phase_10_13mos_tab.tex')
      )

write(apa_table(phase_glmer_14_16mos_tab,
                caption='14--16 month-olds'),
      here('supplement/tables/exp_1/age_binned', 'cn_phase_14_16mos_tab.tex')
      )

# replicate final table in B&S 2012 SI:
```

```r
write(apa_table(
  rbind(
    phase_glmer_6_9mos_tab,
    phase_glmer_10_13mos_tab,
    phase_glmer_14_16mos_tab)),
      here('supplement/tables/exp_1/age_binned', 'cn_phase_agegroups_bsrep_tab.tex')
  )
# will add midrules in tex for ease
```

```r
cn_fin %>%
  group_by(age_group) %>%
  summarize(pre_prop = na.mean(pre_target_prop),
            post_prop = na.mean(post_target_prop))
```

```
424  ## # A tibble: 3 x 3
425  ##   age_group    pre_prop post_prop
426  ##   <fct>           <dbl>     <dbl>
427  ## 1 5-9 months      0.526     0.482
428  ## 2 10-13 months    0.498     0.541
429  ## 3 14-16 months    0.500     0.560
```

430 ## Trial Counts Across Analyses

431   How many more trials were we able to analyze in the pre-/post-looking analysis,
432 relative to the paired difference score calculation, where so many trials with otherwise
433 useable data had to be dropped?

```r
stim_diffs_summary_tab <- cn_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)) %>%
  group_by(subject_id) %>%
  summarize(n=n()) %>%
  ungroup() %>%
  summarize(min_stimdiffs=min(n),
            max_stimdiffs=max(n),
```

```
            mean_stimdiffs=mean(n),
            med_stimdiffs=median(n),
            mode_stimdiffs=DescTools::Mode(n))

#stim_diffs_summary_tab$min_stimdiffs
#stim_diffs_summary_tab$max_stimdiffs
#stim_diffs_summary_tab$mean_stimdiffs
#stim_diffs_summary_tab$mode_stimdiffs
#stim_diffs_summary_tab$med_stimdiffs
```

```
TOTAL_TRIALS_N <- nrow(cn_fin)

MEAN_DIFF_TRIALS_N <- cn_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

TRIALS_DROPPED_FOR_DIFF_CALC <- TOTAL_TRIALS_N - MEAN_DIFF_TRIALS_N
TRIALS_DROPPED_FOR_DIFF_CALC/TOTAL_TRIALS_N
```

434   `## [1] 0.18`

435   94 trials lost (out of 528), or 17.80% of non-excluded trials.

```
cn_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))
```

436   `## # A tibble: 3 x 5`

```
437  ##       min   max  mean   med  mode
438  ##     <dbl> <dbl> <dbl> <dbl> <dbl>
439  ## 1     16    32  25.1    26    16
440  ## 2     16    32  25.1    26    25
441  ## 3     16    32  25.1    26    31
```

442   **Trial Durations**

```
CN_MED_TRIAL_DUR_S <- median(cn_fin$trialtolookingoffset_dur_s)
CN_MIN_TRIAL_DUR_S <- min(cn_fin[cn_fin$trial_dur_s<30,]$trialtolookingoffset_dur_s)
CN_MAX_TRIAL_DUR_S <- max(cn_fin[cn_fin$trial_dur_s<30,]$trialtolookingoffset_dur_s)
CN_MEAN_TRIAL_DUR_S <- mean(
  cn_fin[cn_fin$trial_dur_s<30,]$trialtolookingoffset_dur_s)
CN_CIL_TRIAL_DUR_S <- CN_MEAN_TRIAL_DUR_S -
  ci.low(cn_fin[cn_fin$trial_dur_s<30,]$trialtolookingoffset_dur_s)
CN_CIH_TRIAL_DUR_S <- CN_MEAN_TRIAL_DUR_S +
  ci.high(cn_fin[cn_fin$trial_dur_s<30,]$trialtolookingoffset_dur_s)

ggplot(cn_fin, aes(x=trialtolookingoffset_dur_s)) +
  geom_histogram(fill='#bae4bc') +
  sb.density.theme +
  geom_vline(xintercept=CN_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 = 15s',
    x = CN_MED_TRIAL_DUR_S+2, y = 20, size = 4, colour = 'red')
```
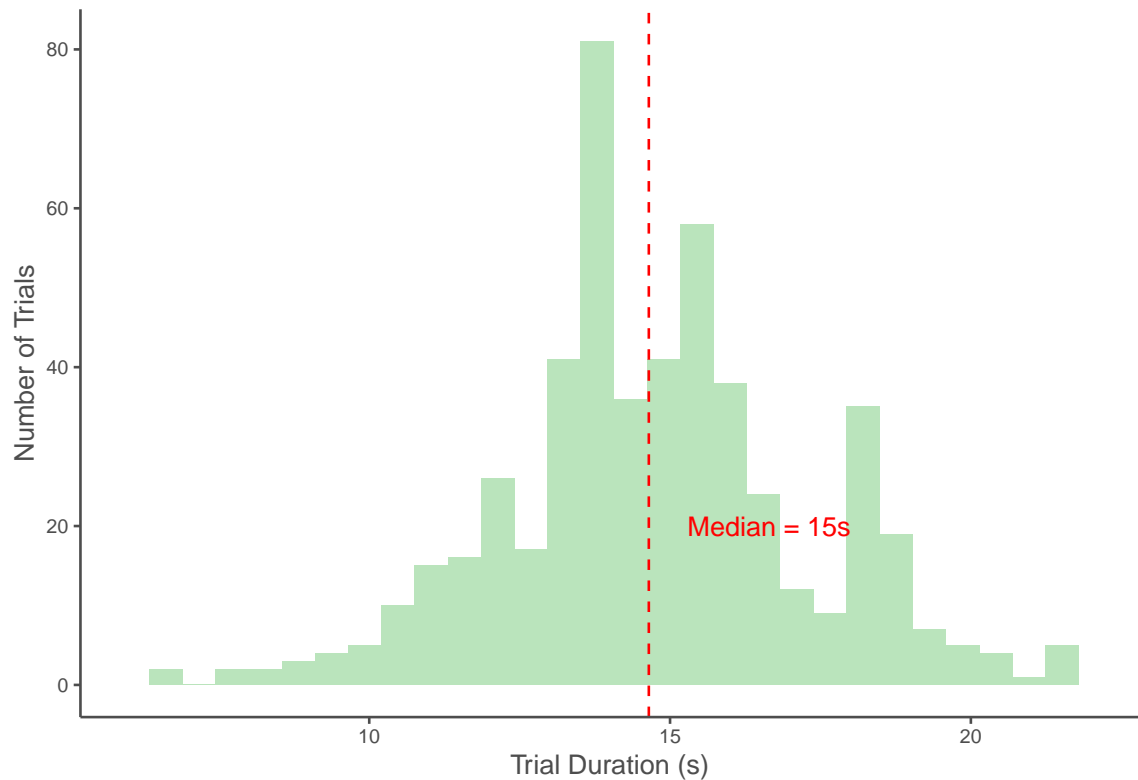
```
ggsave(here('supplement/plots/exp_1/pdfs', 'cn_trialdur_histogram.pdf'),
       device='pdf', width=2.5, height=1.5, units='in', scale=2.5)
ggsave(here('supplement/plots/exp_1/pngs', 'cn_trialdur_histogram.png'),
       device='png', width=2.5, height=1.5, units='in', scale=2.5)
```

```
CN_MED_TRIAL_LOOKING_S <- median(
  cn_fin$totaltrialtime_looking_sum_s)
CN_MIN_TRIAL_LOOKING_S <-
  min(cn_fin$totaltrialtime_looking_sum_s)
CN_MAX_TRIAL_LOOKING_S <-
  max(cn_fin$totaltrialtime_looking_sum_s)
CN_MEAN_TRIAL_LOOKING_S <-
  mean(cn_fin$totaltrialtime_looking_sum_s)
CN_CIL_TRIAL_LOOKING_S <- CN_MEAN_TRIAL_LOOKING_S -
  ci.low(cn_fin$totaltrialtime_looking_sum_s)
CN_CIH_TRIAL_LOOKING_S <- CN_MEAN_TRIAL_LOOKING_S +
  ci.high(cn_fin$totaltrialtime_looking_sum_s)

ggplot(cn_fin, aes(x=totaltrialtime_looking_sum_s)) +
  geom_histogram(fill='#7bccc4') +
  sb.density.theme +
  geom_vline(xintercept=CN_MED_TRIAL_LOOKING_S,
```

```
                color='red', lty=2) +
  xlim(2,14) +
  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 = 6.8s',
    x = CN_MED_TRIAL_LOOKING_S+1.5, y = 69, size = 4, colour = 'red')
```
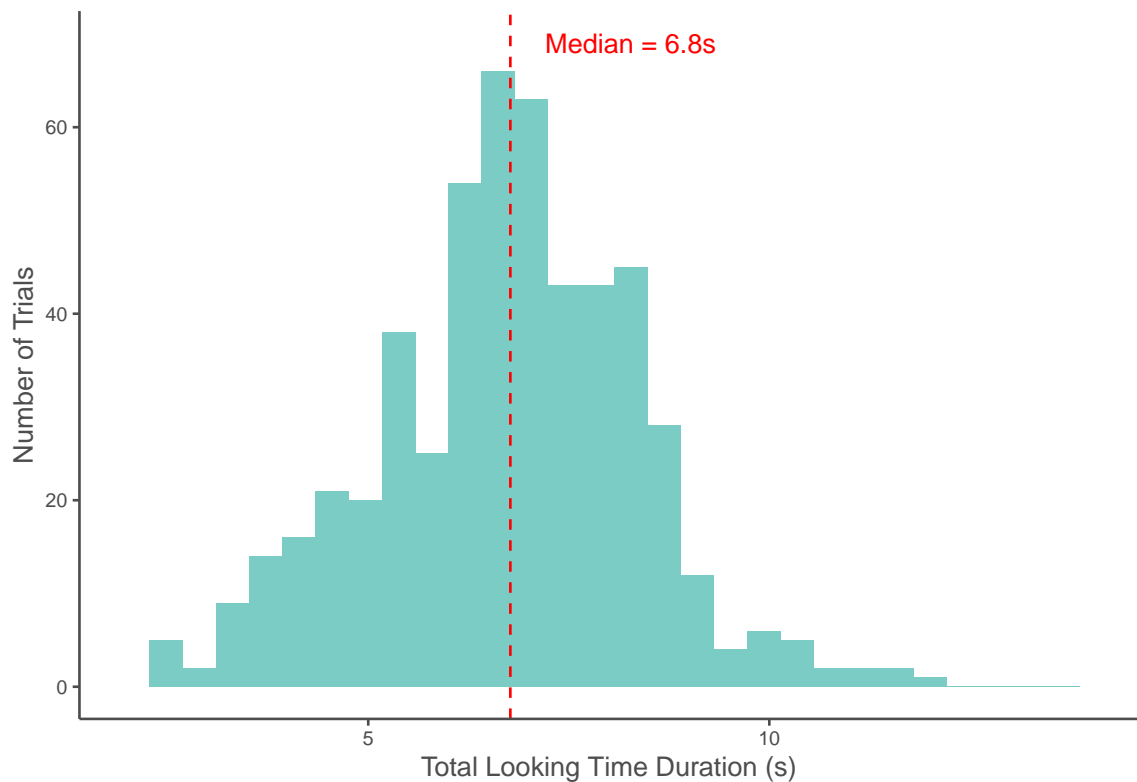

444

```
ggsave(here('supplement/plots/exp_1/pdfs', 'cn_lookingtime_histogram.pdf'),
       device='pdf', width=2.5, height=1.25, units='in', scale=2.5)
ggsave(here('supplement/plots/exp_1/pngs', 'cn_lookingtime_histogram.png'),
       device='png', width=2.5, height=1.25, units='in', scale=2.5)
```

```r
CN_MED_LOOKING_PROP <- median(
  cn_fin$trialtolookingoffset_prop)
CN_MIN_LOOKING_PROP <- min(
  cn_fin$trialtolookingoffset_prop)
CN_MAX_LOOKING_PROP <- max(
  cn_fin$trialtolookingoffset_prop)
CN_MEAN_LOOKING_PROP <- mean(
  cn_fin$trialtolookingoffset_prop)
CN_CIL_LOOKING_PROP <- CN_MEAN_LOOKING_PROP - ci.low(
  cn_fin$trialtolookingoffset_prop)
CN_CIH_LOOKING_PROP <- CN_MEAN_LOOKING_PROP +ci.high(
  cn_fin$trialtolookingoffset_prop)

ggplot(cn_fin, aes(x=trialtolookingoffset_prop)) +
  geom_histogram(fill='#7bccc4') +
  sb.density.theme +
  geom_vline(xintercept=CN_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 = 0.47',
    x = CN_MED_LOOKING_PROP-.12, y = 68, size = 4, colour = 'red')
```
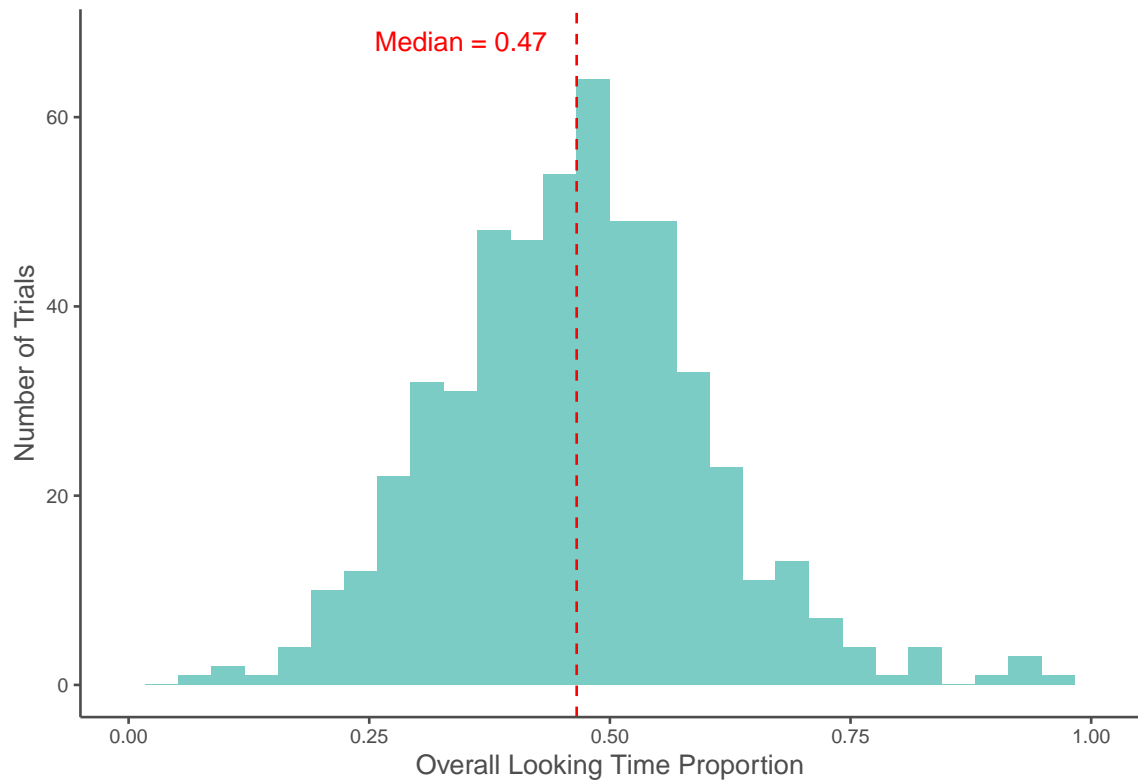
445

```
ggsave(here('supplement/plots/exp_1/pdfs', 'cn_lookingprop_histogram.pdf'),
       device='pdf', width=2.5, height=1.25, units='in', scale=2.5)
ggsave(here('supplement/plots/exp_1/pngs', 'cn_lookingprop_histogram.png'),
       device='png', width=2.5, height=1.25, units='in', scale=2.5)
```

```
CN_MED_PRE_DUR_S <- median(cn_fin$pre_dur_ms/1000)
CN_MIN_PRE_DUR_S <- min(cn_fin$pre_dur_ms/1000)
CN_MAX_PRE_DUR_S <- max(cn_fin[cn_fin$pre_dur_ms<10000,]$pre_dur_ms/1000)
CN_MEAN_PRE_DUR_S <- mean(cn_fin[cn_fin$pre_dur_ms<10000,]$pre_dur_ms/1000)

ggplot(cn_fin, aes(x=pre_dur_ms/1000))+
  geom_histogram(fill='#bae4bc') +
  sb.density.theme +
  geom_vline(xintercept=CN_MED_PRE_DUR_S, color='red', lty=2) +
  #geom_vline(xintercept = 3.133, color='red') +
  xlim(3,6.5)+
  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 = 5s',
    x = CN_MED_PRE_DUR_S+.4, y = 38, size = 4, colour = 'red')
```
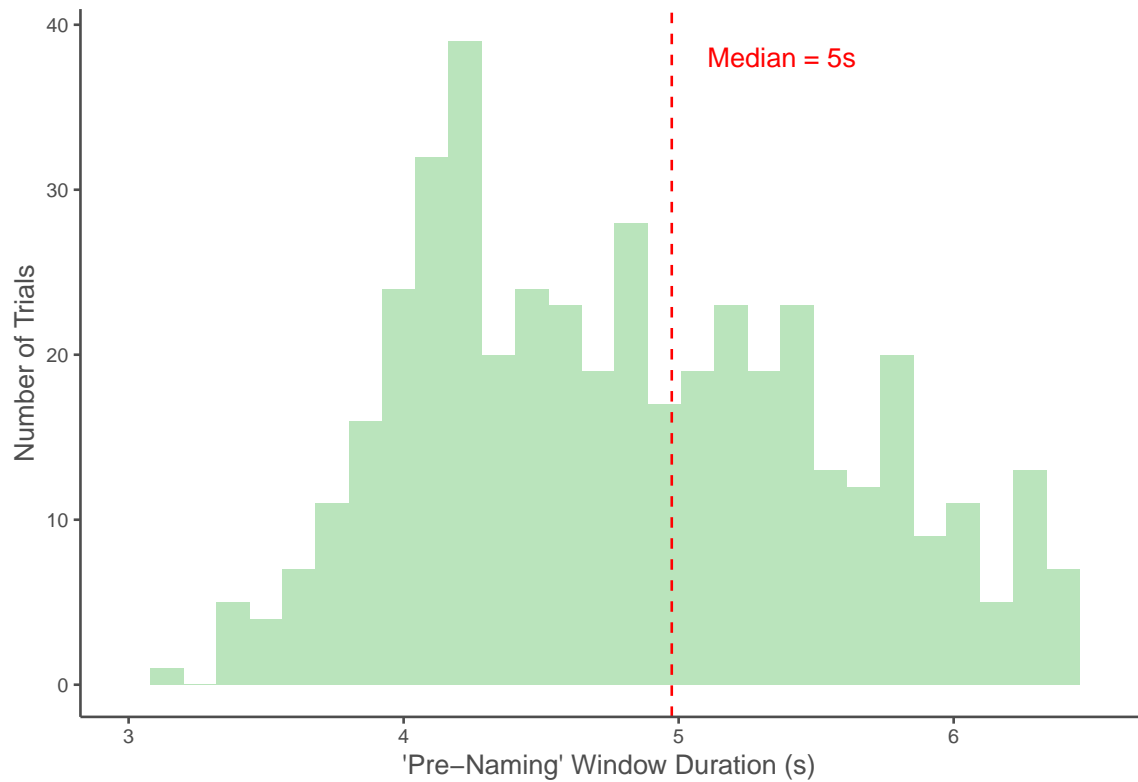

446

```
prelookingprop_df <- cn_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')

CN_PRE_MIN_PROP <- min(prelookingprop_df$looking_prop)
CN_PRE_MAX_PROP <- max(prelookingprop_df$looking_prop)
CN_PRE_MEAN_PROP <- mean(prelookingprop_df$looking_prop)
CN_PRE_MEDIAN_PROP <- median(prelookingprop_df$looking_prop)

postlookingprop_df <- cn_fin %>%
```

```r
  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')

CN_POST_MIN_PROP <- min(postlookingprop_df$looking_prop)
CN_POST_MAX_PROP <- max(postlookingprop_df$looking_prop)
CN_POST_MEAN_PROP <- mean(postlookingprop_df$looking_prop)
CN_POST_MEDIAN_PROP <- median(postlookingprop_df$looking_prop)

prepost_lookingprop_df <- rbind(prelookingprop_df, postlookingprop_df)
prepost_lookingprop_df$window <-
  factor(prepost_lookingprop_df$window, levels=c(
    "'Pre-Naming' Window","'Post-Naming'/Analysis Window"), ordered=T)

prepost_lookingprop_label_df <- prepost_lookingprop_df %>%
  group_by(window) %>%
  summarize(median=median(median),
            label=paste('Median =', round(median, 2), sep=' '))

ggplot(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(0,100) +
  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=prepost_lookingprop_label_df,
            aes(x=median-.24, label=label), y=21,
            color='red', size=4)
```
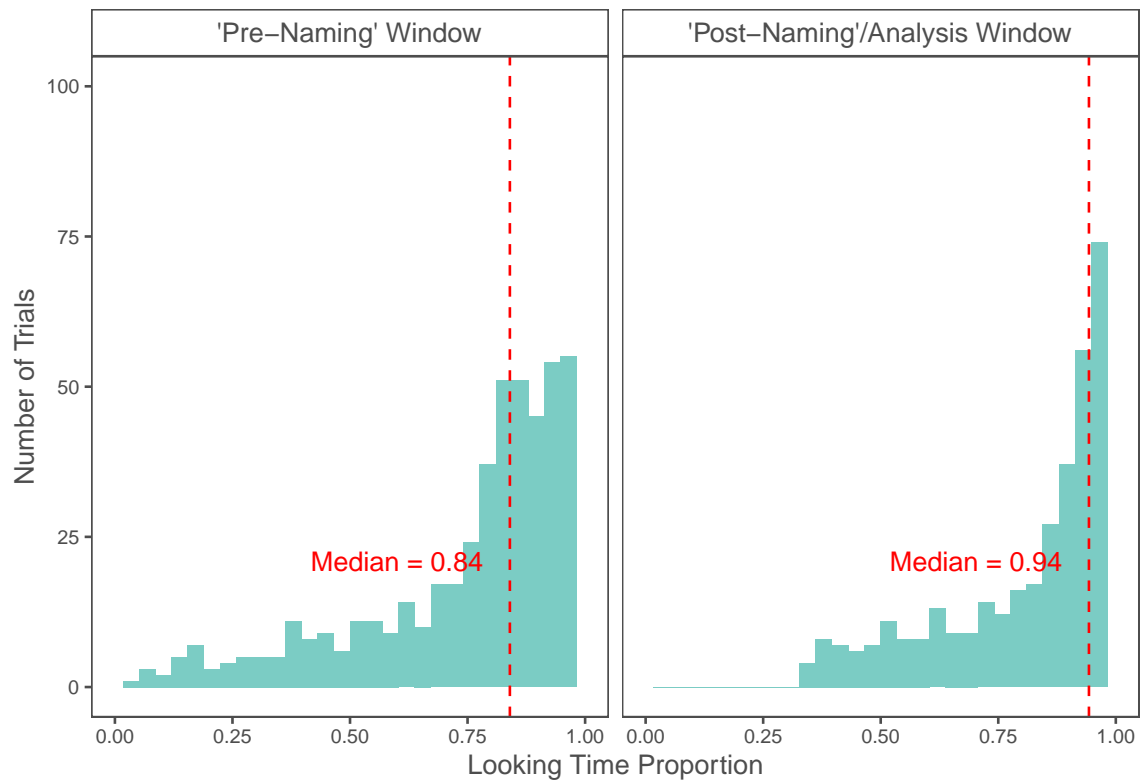
447

```r
ggsave(here('supplement/plots/exp_1/pdfs', 'cn_lookingprops_prepost.pdf'),
       device='pdf', width=2.75, height=1.5, units='in', scale=2.5)
ggsave(here('supplement/plots/exp_1/pngs', 'cn_lookingprops_prepost.png'),
       device='png', width=2.75, height=1.5, units='in', scale=2.5)
```

```r
prelookingdur_df <- cn_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')

CN_PRE_MIN_DUR <- min(prelookingdur_df$pre_looking_sum_ms)
CN_PRE_MAX_DUR <- max(prelookingdur_df$pre_looking_sum_ms)
CN_PRE_MEAN_DUR <- mean(prelookingdur_df$pre_looking_sum_ms)
CN_PRE_MEDIAN_DUR <- median(prelookingdur_df$pre_looking_sum_ms)

postlookingdur_df <- cn_fin %>%
  mutate(window = "'Post-Naming'/Analysis Window",
         looking_dur = post1_looking_sum_ms/1000,
         median = median(looking_dur),
         mean = mean(looking_dur)) %>%
```

```r
  dplyr::select('subject_id', 'window', 'looking_dur', 'median', 'mean')

CN_POST_MIN_DUR <- min(postlookingdur_df$looking_dur)
CN_POST_MAX_DUR <- max(postlookingdur_df$looking_dur)
CN_POST_MEAN_DUR <- mean(postlookingdur_df$looking_dur)
CN_POST_MEDIAN_DUR <- median(postlookingdur_df$looking_dur)

prepost_lookingdur_df <- rbind(prelookingdur_df, postlookingdur_df)
prepost_lookingdur_df$window <-
  factor(prepost_lookingdur_df$window, levels=c(
    "'Pre-Naming' Window","'Post-Naming'/Analysis Window"), ordered=T)

prepost_lookingdur_label_df <- prepost_lookingdur_df %>%
  group_by(window) %>%
  summarize(median=median(median),
            label=paste('Median =', round(median, 2), sep=' '))

cn_pp_durs <- ggplot(prepost_lookingdur_df, aes(x=looking_dur)) +
  geom_histogram(fill='#7bccc4') +
  sb.density.theme +
  geom_vline(aes(xintercept=median), color='red', lty=2) +
  xlim(0,9) +
  #ylim(0,100) +
  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=prepost_lookingdur_label_df,
            aes(x=median+2, label=label), y=250,
            color='red', size=4)

ggsave(here('supplement/plots/exp_1/pdfs', 'cn_lookingdurs_prepost.pdf'),
       device='pdf', width=2.75, height=1.5, units='in', scale=2.5)
ggsave(here('supplement/plots/exp_1/pngs', 'cn_lookingdurs_prepost.png'),
       device='png', width=2.75, height=1.5, units='in', scale=2.5)
```

448   Trials in Experiment 1 were $14.86s$ [14.60, 15.13] long on average (*range*: $5.42-24.82s$,
449 $M_{\mathrm{dur}} = 14.86s$, *Med*$= 14.65s$), and infants spent an average of $6.70s$ [6.55, 6.84] total looking
450 at the displays (*range*: $1.97 - 12.17s$, *Med*$= 6.77s$, or between 0.07 and 0.99 of the total
451 trial duration; $M = 0.46$ [0.45, 0.48]).

452   The pre-naming window was between 3.12 and $9.44s$ ($M = 5.22s$, $Med= 4.97s$).

453   The children in our final sample spent similar proportions of time looking at the
454   displays during the pre- and post-naming periods (PRE-NAMING *range*: $0.03 - 1$, $M = 0.77$,
455   *Med*= 0.84; POST-NAMING *range*: $0.34 - 1$, $M = 0.87$, *Med*= 0.94).