Individual differences in nonverbal prediction and vocabulary size in infancy

Tracy Reuter, Lauren Emberson, Alexa Romberg, and Casey Lew-Williams

Contents

Nonverbal Prediction Task **2**
- AEMs in Both Blocks . . . 3
- AEMs in Block 1 . . . 5
- AEMs by Trial in Block 1 . . . 7
- AEMs in Block 2 . . . 8
- Updating AEMs in Block 2 . . . 10
- Updating AEMs by Trial in Block 2 . . . 11
- AEMs and Age . . . 12

Visual Processing (Gap-Overlap) Task **14**
- Reaction Time (RT) by Condition . . . 14
- Correlating Visual Processing Speed and AEMs . . . 15

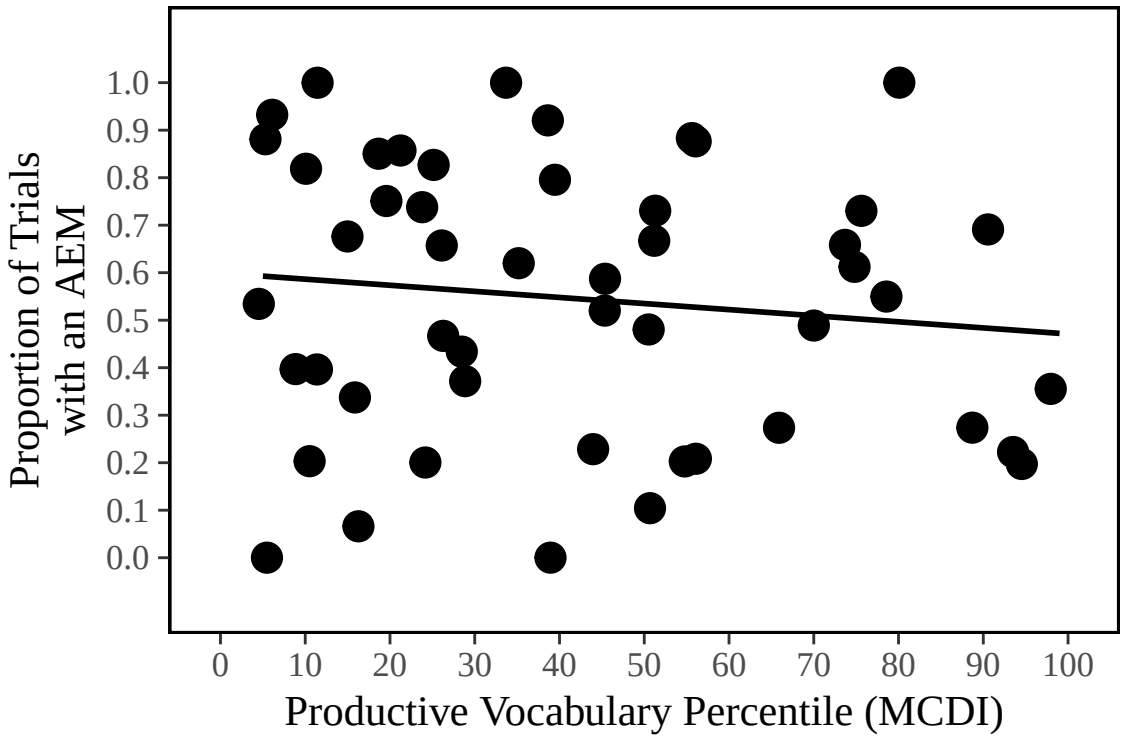
Additional Analyses (Response to Reviewers) **16**

Nonverbal Prediction Task

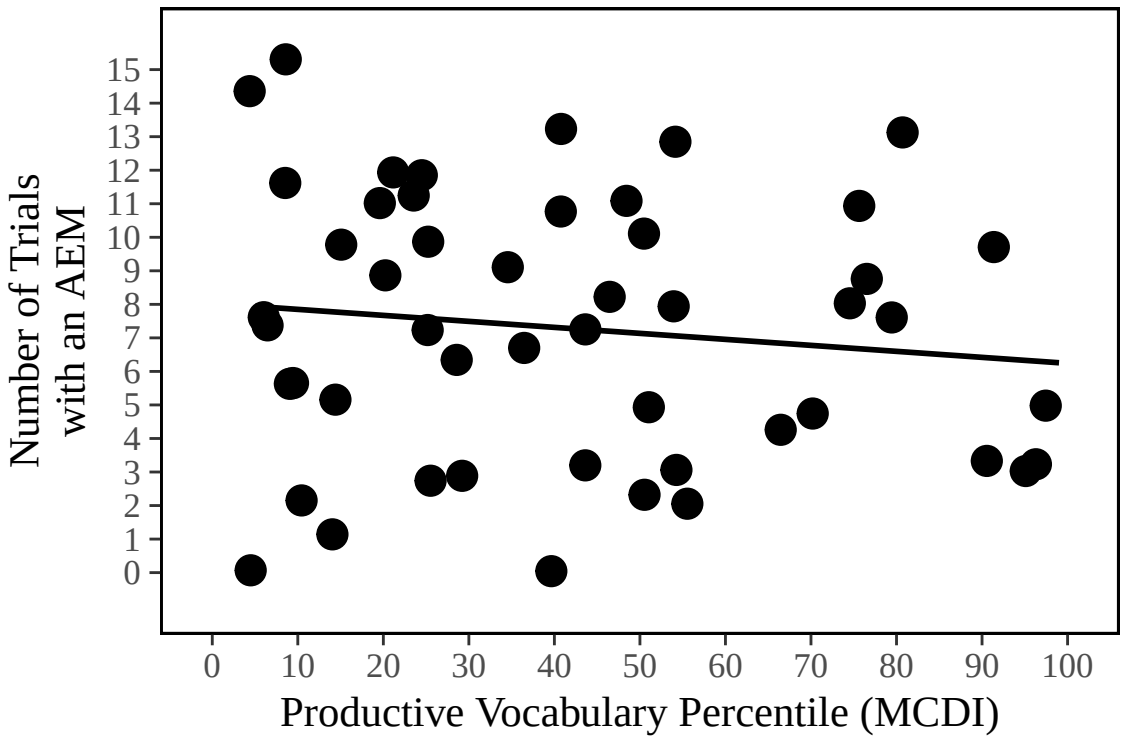
We hypothesized that infants with larger vocabularies would make more predictions and would update predictions more readily than infants with smaller vocabularies. To evaluate these hypotheses, we correlated infants' productive vocabulary percentile (MCDI) and anticipatory eye movements (AEMs) during the prediction task. We excluded Trial 1 from all analyses, as there was no basis for infants to generate predictions therein. We also compared high-vocabulary and low-vocabulary groups, based on a median split in MCDI scaled scores. Finally, we analyzed visual processing speed to evaluate alternative explanations for individual variability in infants' performance on the prediction task.

AEMs in Both Blocks

To evaluate whether infants with larger vocabularies generated more predictions overall (regardless of accuracy), we first correlated infants' vocabulary size with their proportion of total trials with an AEM to either location. There was striking variation in both vocabulary size (MCDI percentile range=5-99, M=42, SD=28) and proportion of trials with an AEM (range=0-100, M=55, SD=29). Including all infants, regardless of how many AEMs they generated, we found no correlation between infants' vocabulary size and overall proportion of trials with an AEM [r(48)=-0.12, p=0.39]. Examining each Block separately, there was no correlation in Block 1 [r(48)=-0.03, p=0.82] nor in Block 2 [r(47)=-0.17, p=0.23]. Thus, contrary to our first hypothesis, infants with relatively larger vocabularies were not more likely to anticipate the target's appearance.

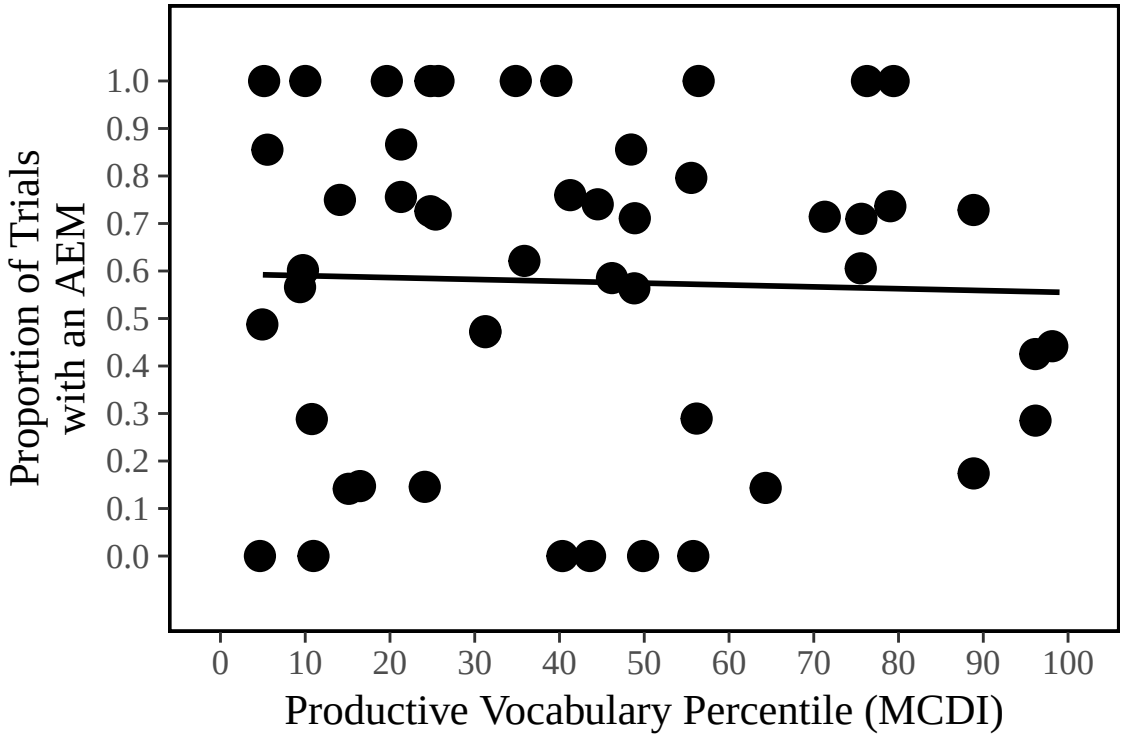



```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$MCDI and corr.data$Any.AEM
## t = -0.86879, df = 48, p-value = 0.3893
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.3892894  0.1594459
## sample estimates:
##       cor
## -0.124424
```

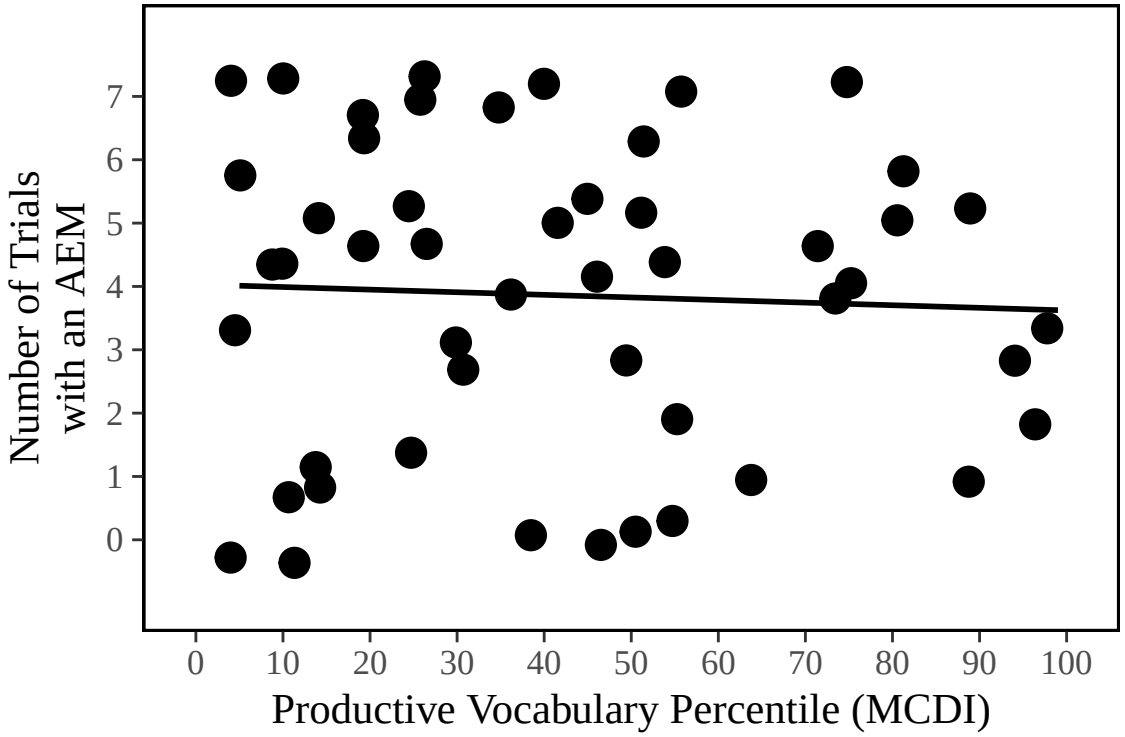



```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$MCDI and corr.data$AEM.count
## t = -0.87491, df = 48, p-value = 0.386
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.3900335  0.1585907
## sample estimates:
##        cor
## -0.1252876
```

AEMs in Block 1



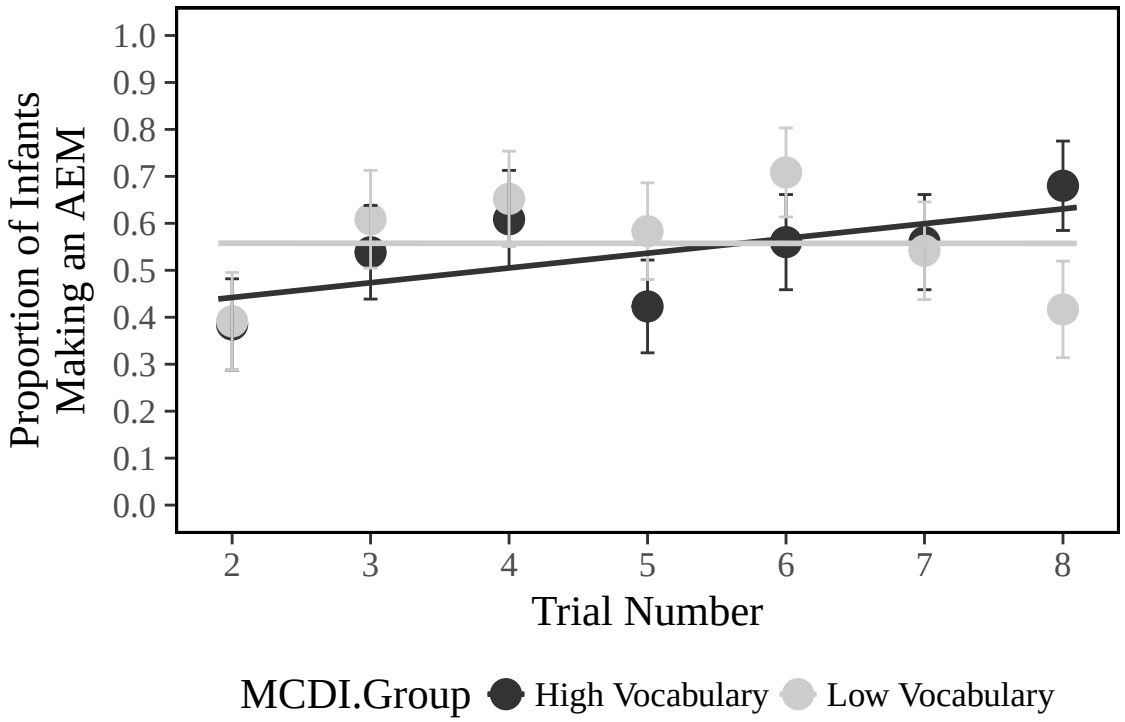

```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$MCDI and corr.data$Any.AEM
## t = -0.22355, df = 48, p-value = 0.8241
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.3078349  0.2483263
## sample estimates:
##         cor
## -0.03225059
```




```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$MCDI and corr.data$AEM.count
## t = -0.33755, df = 48, p-value = 0.7372
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.3226414  0.2328376
## sample estimates:
##         cor
## -0.04866398
```

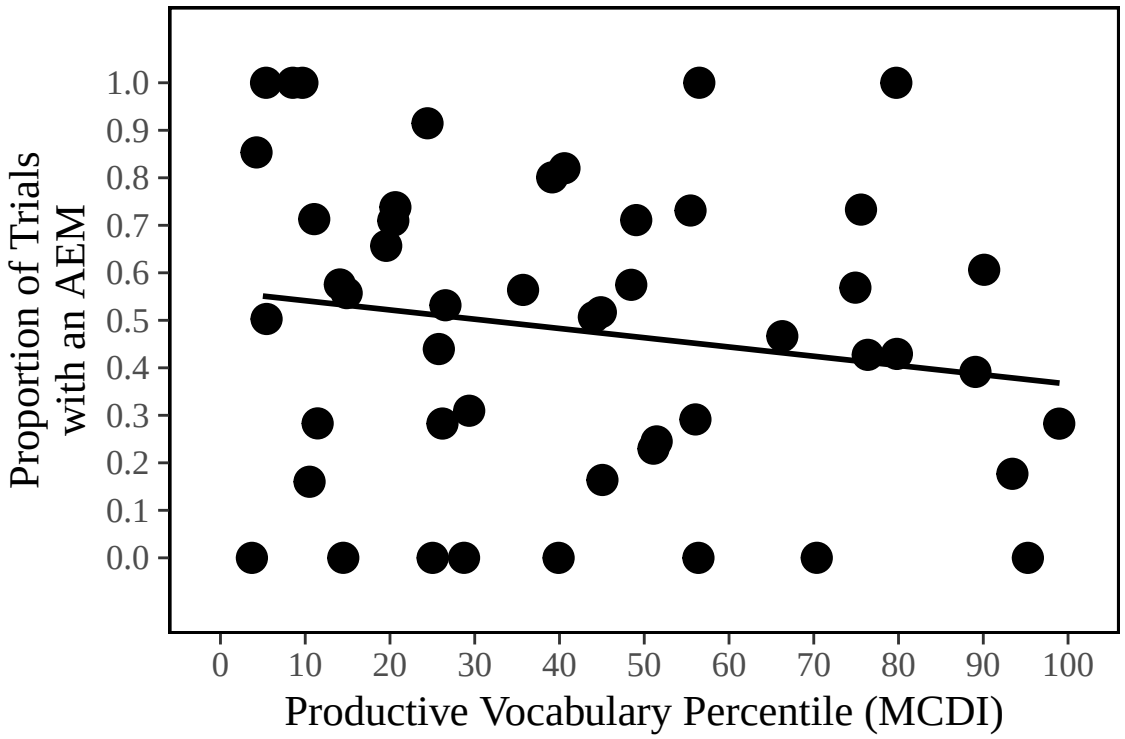
AEMs by Trial in Block 1

We also examined AEMs trial-by-trial in Block 1. We find that high-MCDI infants do not generate predictions faster in the learning process than low-MCDI infants. Again, this suggests both groups learned equally well in the first block. The rate of AEMs from trial to trial also suggested that this task may have been too easy for infants. On each trial, about 50% of infants were generating predictions, with very little, if any, increase in slope. (An increase in slope might indicate learning - that infants wait for more information before generating predictions.)

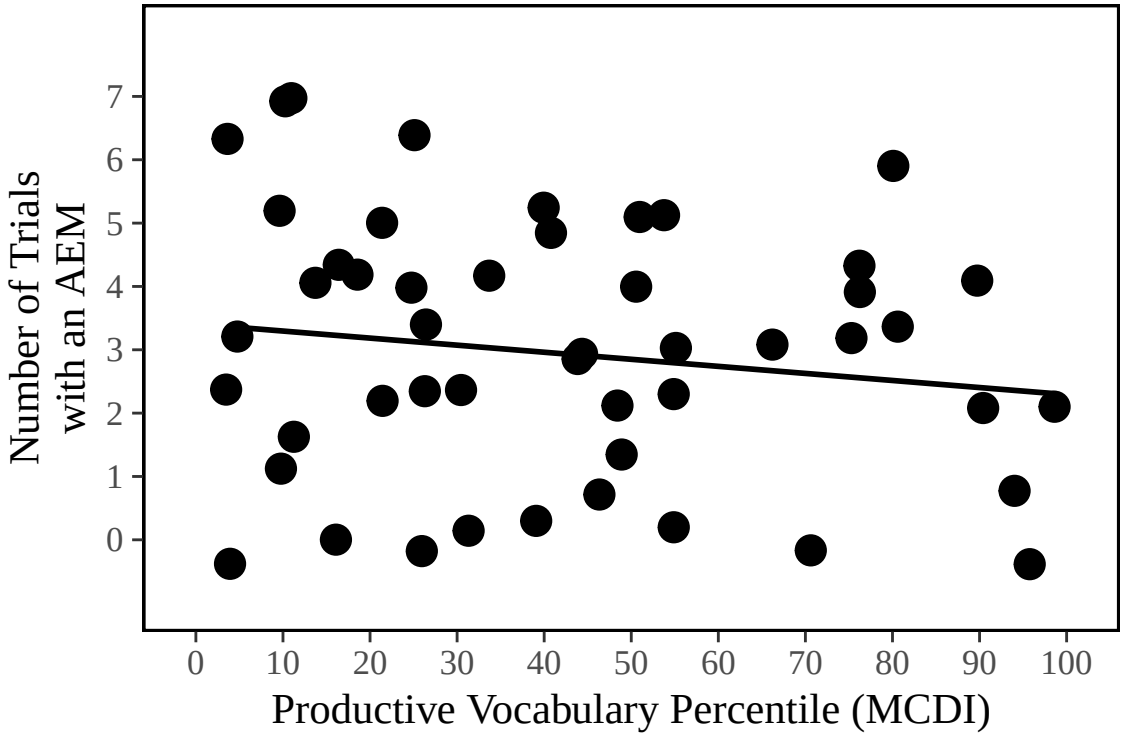


AEMs in Block 2

MCDI does not correlate with AEMs (to either target location - novel or familiar) in Block 2. (This analysis excludes Trial 9, as there is no basis to generate a prediction towards the novel target location yet.)



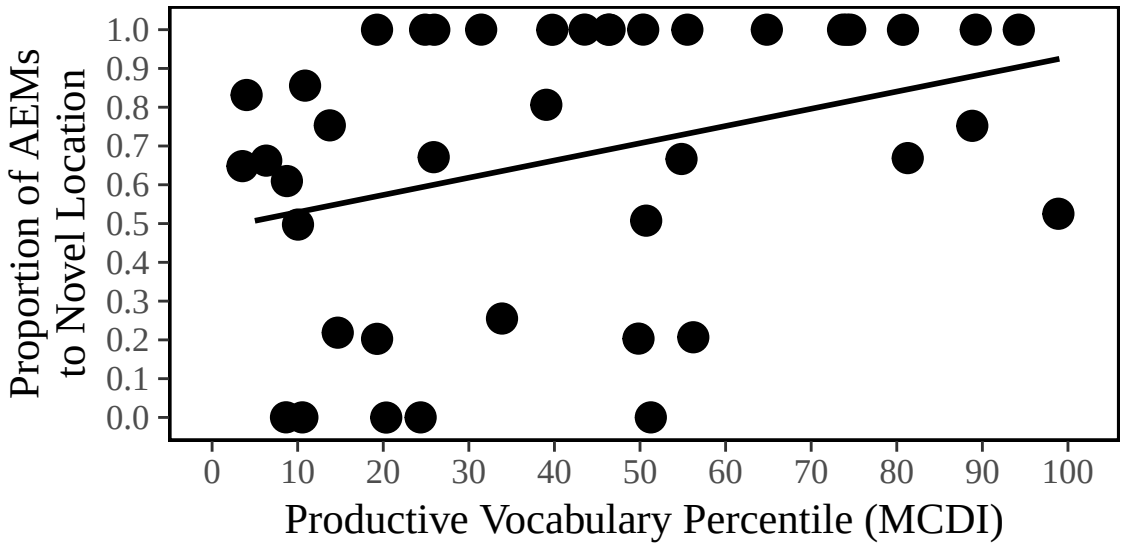

```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$MCDI and corr.data$Any.AEM
## t = -1.2061, df = 47, p-value = 0.2338
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.4333538  0.1134527
## sample estimates:
##       cor
## -0.1732717
```




```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$MCDI and corr.data$AEM.count
## t = -1.0972, df = 47, p-value = 0.2781
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.4205454  0.1288866
## sample estimates:
##       cor
## -0.1580377
```

Updating AEMs in Block 2

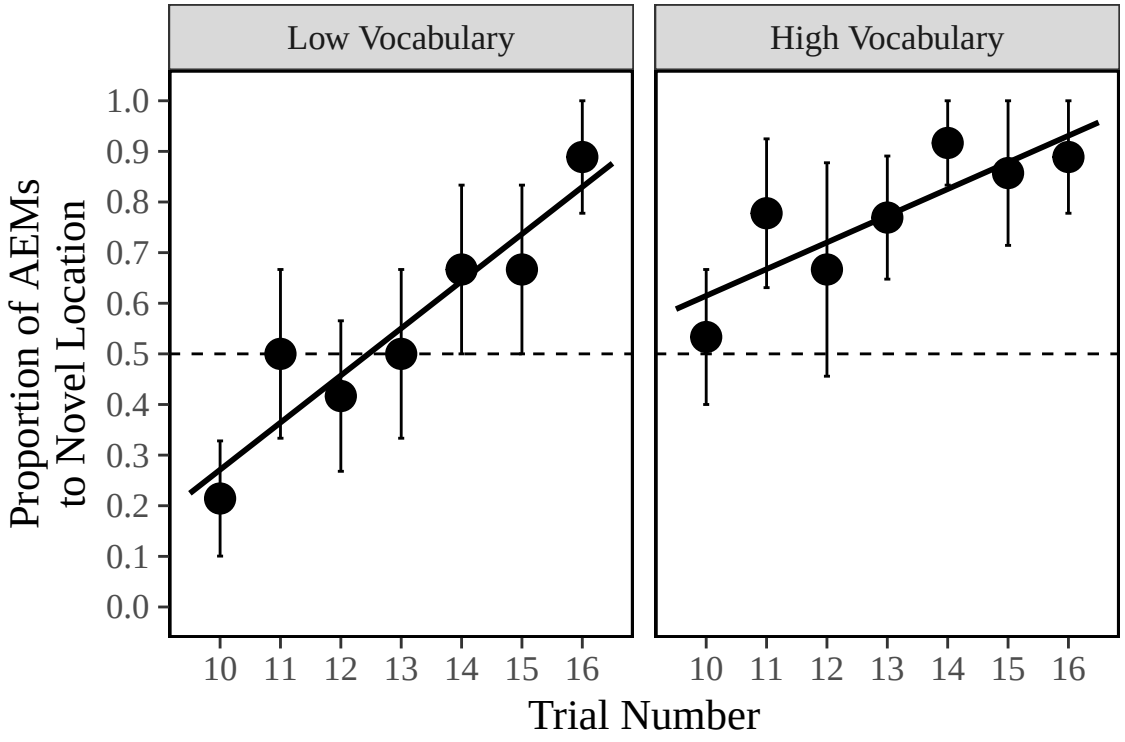
To test our second hypothesis – that infants with larger vocabularies would update predictions more readily – we classified AEMs in Block 2 as directed toward the novel location (i.e., the Block 2 target location) or the familiar location (i.e., the Block 1 target location). Trial 9 was excluded from this analysis, because this was the first trial in which the target appeared in the novel location. Thus, infants had no basis for directing an AEM to the novel target location during Trial 9. Infants who did not make any AEMs (n=8) or who did not have any code-able trials in Block 2 (n=1) were not included in this analysis. For each infant, we calculated the proportion of their AEMs that were to the novel location (i.e., novel AEMs divided by total AEMs). We found a significant correlation between infants' vocabulary size and the proportion of their AEMs to the novel location [r(39)=0.34, p=0.028]. Consistent with our second hypothesis, when making an AEM in Block 2, infants with larger vocabularies were more likely to make an AEM to the novel target location than infants with smaller vocabularies.




```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$MCDI and corr.data$AEM.ratio
## t = 2.2835, df = 39, p-value = 0.02793
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.03998657 0.58885024
## sample estimates:
##       cor
## 0.3434124
```

Updating AEMs by Trial in Block 2

To examine the time-course of prediction updating, we analyzed AEMs trial by trial in Block 2. We divided infants into high-vocabulary and low-vocabulary groups based on a median split in MCDI percentile scores (high vocabulary: n=22, M=64, SD=19; low vocabulary: n=19, M=17, SD=9). Then, on each trial we calculated the proportion of AEMs to the novel location as the number of infants who made an AEM to the novel location divided by the number of infants who made AEMS to either location. A binomial test was used to test this proportion against a chance value of 0.5. The low-vocabulary group only performed above chance on trial 16 (p=0.039), making significantly more AEMs to the novel than the familiar location. The high-vocabulary group was marginally above chance in trial 13 (p=0.092), above chance in trial 14 (p=0.006), and above chance in trial 16 (p=0.039).



AEMs and Age

To assess whether nonverbal prediction varied as a function of age, we repeated the above correlations with infants' age in months as a factor. We found no correlation between infants' age and overall proportion of trials with an AEM [r(48)=-0.25, p=0.08]. There was no correlation in Block 1 [r(48)=-0.23, p=0.11] nor in Block 2 [r(47)=-0.22, p=0.13]. Finally, there was no correlation between infants' age and the proportion of their AEMs to the novel location in Block 2 [r(39)=0.10, p=0.52].

```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$Age and corr.data$Any.AEM
## t = -1.7436, df = 48, p-value = 0.08763
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.48917489  0.03678855
## sample estimates:
##       cor
## -0.2440583
```

```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$Age and corr.data$Any.AEM
## t = -1.6018, df = 48, p-value = 0.1158
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.47388959  0.05664502
## sample estimates:
##       cor
## -0.2252543
```

```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$Age and corr.data$Any.AEM
## t = -1.5066, df = 47, p-value = 0.1386
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.46761246  0.07083278
## sample estimates:
##       cor
## -0.214639
```

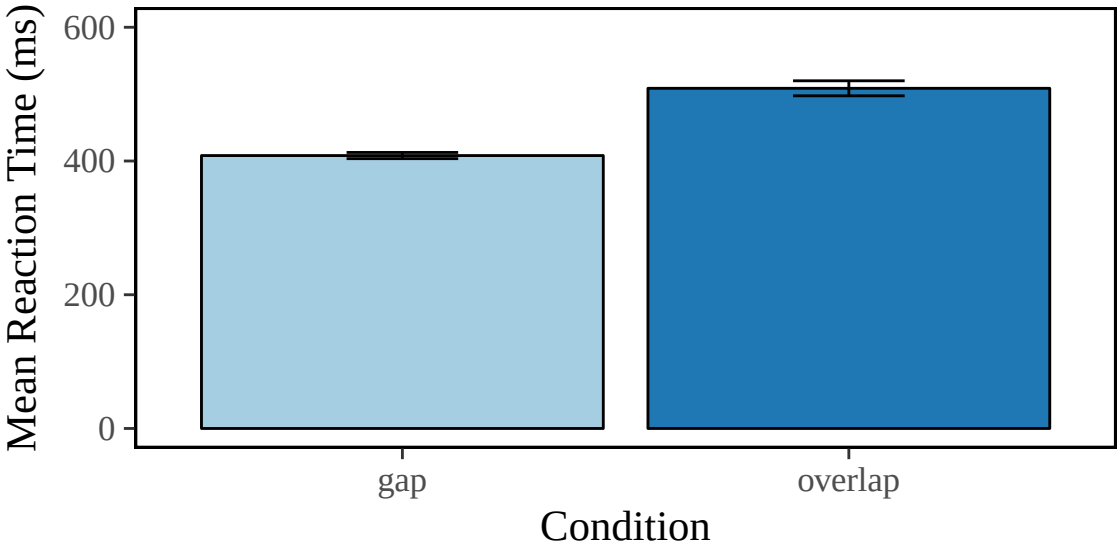
```
##
```

```
##  Pearson's product-moment correlation
##
## data:  corr.data$Age and corr.data$AEM.ratio
## t = 0.6541, df = 39, p-value = 0.5169
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.2102179  0.3990320
## sample estimates:
##       cor
## 0.1041696
```

Visual Processing (Gap-Overlap) Task

The gap-overlap task allowed us to examine whether individual differences in AEMs were due to differences in visual processing speed. First, we confirmed that infants were slower to shift their eyes to the target on overlap trials when there was a competing stimulus than on gap trials when there was no competing stimulus [gap trials: M=408 ms, SD=31; overlap trials: M=509 ms, SD=74; two-tailed t(42)=-10.4, p<.001].

Reaction Time (RT) by Condition




```
##
##  Paired t-test
##
## data:  RT.data$Gap.RT and RT.data$Overlap.RT
## t = -10.384, df = 42, p-value = 3.598e-13
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -120.12572  -81.03204
## sample estimates:
## mean of the differences
##               -100.5789
```

Correlating Visual Processing Speed and AEMs

Next, we computed RT difference scores for each infant (i.e., overlap trials minus gap trials) and used this difference score as the measure of visual processing speed independent of prediction. We found no correlation between infants' RT difference scores and their proportion of AEMs overall [r(40)=-0.09, p=0.56] and no correlation between infants' RT difference scores and their proportion of AEMs to the novel target location in Block 2 of the prediction task [r(33)=0.18, p=0.29]. Together, these findings suggest that infants' performance on the prediction task was not related to variation in orienting to the dynamics of visual events.

```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$RT.Score.mean and corr.data$AEM.Any.mean
## t = -0.58449, df = 40, p-value = 0.5622
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.3851821  0.2180048
## sample estimates:
##         cor
## -0.09202392
```

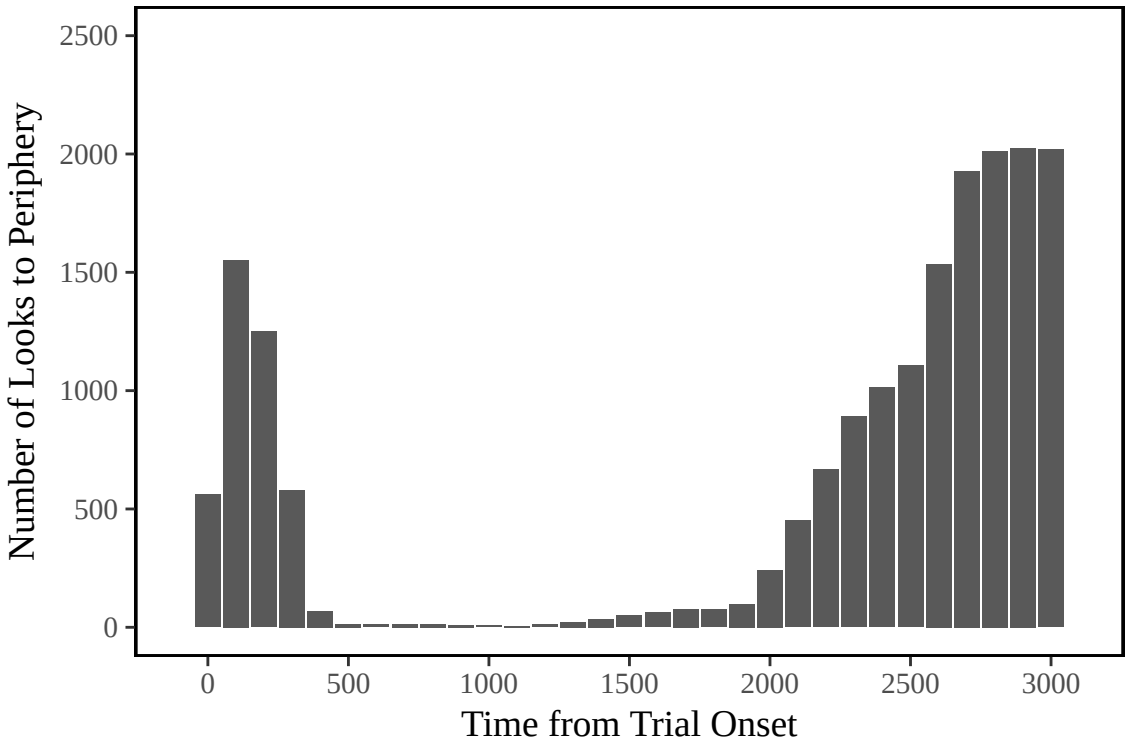
```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$RT.Score.mean and corr.data$AEM.Tn.mean
## t = 1.0702, df = 33, p-value = 0.2923
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.1598531  0.4866917
## sample estimates:
##     cor
## 0.18315
```

Additional Analyses (Response to Reviewers)

We completed a number of additional analyses, which were included in our response to reviewers.

Why start the AEM measurement window 200ms BEFORE center fixation image offset? In principle, it seems like you could start the window AT fixation image offset or 200ms AFTER fixation image offset, but that any other selection is going to be arbitrary.

Behaviorally, infants may not generate anticipatory eye movements (AEMs) with precise timing. Therefore, we suspected that some AEMs might occur prior to the offset of the central fixation. To ensure that we included such eye movements in analyses, we used a 200-ms buffer at both the onset and offset of our AEM window. But your comment prompted us to dig deeper. In the below figure, we plotted infants' looks to the periphery over time. However, very few looks to the periphery (1.2%) occurred in the 200-ms window prior to fixation offset (i.e., between 1300 and 1500 ms). Thus, the decision to include this 200-ms buffer is unlikely to have impacted AEM results. (Also note that the high number of looks to the periphery at the onset of the trial is due to residual looking behavior from the prior trial.) In light of this, we decided not to alter our original plan for window size.



Why no auditory stimuli or moving visual stimuli in the visual processing task? Did this make the task less engaging?

We decided to use silent, static images, in order to make our design highly comparable to prior studies. Specifically, we felt it appropriate to hold the parameters from Hood and Atkinson (1993) constant. However, there is no reason to believe that this specific design decision was crucial to evaluating visual processing speed. To assess whether this design made the task less engaging, we calculated the proportion of trials included in analysis per subject in each task, and compared the proportions with a Welch t-test. We find that the proportion of trials included in analyses did not vary by task ($t(83) = 0.50$, $p = 0.62$), suggesting that the two tasks were equally engaging for infants.

```
##
##  Welch Two Sample t-test
##
## data:  task1$prop and task2$prop
## t = 0.50418, df = 83.234, p-value = 0.6155
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.04308719  0.07235076
## sample estimates:
## mean of x mean of y
## 0.8625000 0.8478682
```

I wonder if having a "standardized difference score" for this task (parallel to the standardized MCDI score) might be a better predictor of prediction probability/prediction updating.

Thank you for this analysis suggestion. To assess this possibility, we transformed each of the RT difference scores into a z-score (i.e., subtracted the mean from every value and divided the difference by the standard deviation), and completed correlations among RTs and AEMs. We find the same pattern of results as with the raw RT difference scores: We find no correlation between infants' RT difference scores and their proportion of AEMs overall [r(40)=-0.09, p=0.56] and no correlation between infants' RT difference scores and their proportion of AEMs to the novel target location in Block 2 of the prediction task [r(33)=0.18, p=0.29]. We are not including these statistics in the manuscript, but would be willing to if you think it is important.

```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$RT.Score.z.mean and corr.data$AEM.Any.mean
## t = -0.58449, df = 40, p-value = 0.5622
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.3851821  0.2180048
## sample estimates:
##         cor
## -0.09202392
```

```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$RT.Score.z.mean and corr.data$AEM.Tn.mean
## t = 1.0702, df = 33, p-value = 0.2923
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.1598531  0.4866917
## sample estimates:
##     cor
## 0.18315
```

Why the difference scores? Would RT on the gap trials or on the overlap trials alone have been a better predictor?

For the sake of brevity in the manuscript, we used an RT difference score (overlap RT minus gap RT) to evaluate whether infants' performance on the visual processing task might explain their performance on the nonverbal prediction task. However - and again not included in the manuscript for the sake of brevity - we also analyzed RTs for gap trials and overlap trials independently. We found that infants' RTs do not correlate with their proportion of AEMs overall [Gap trials: r(40) = -0.21, p = 0.18; Overlap trials: r(40) = -0.17, p = 0.29], nor do infants' RTs correlate with their proportion of AEMs correct in the second block of the nonverbal prediction task [Gap trials: r(33) = -0.10, p = 0.56; Overlap trials: r(33) = 0.13, p = 0.47].

```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$Gap.RT.mean and corr.data$AEM.Any.mean
## t = -1.3693, df = 40, p-value = 0.1785
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.48438116  0.09867597
## sample estimates:
##        cor
## -0.2116015
```

```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$Overlap.RT.mean and corr.data$AEM.Any.mean
## t = -1.0708, df = 40, p-value = 0.2907
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.4481307  0.1443174
## sample estimates:
##        cor
## -0.1669362
```

```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$Gap.RT.mean and corr.data$AEM.Tn.mean
## t = -0.59024, df = 33, p-value = 0.5591
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.4211123  0.2391841
## sample estimates:
##        cor
```

```
## -0.1022093
```

```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$Overlap.RT.mean and corr.data$AEM.Tn.mean
## t = 0.72927, df = 33, p-value = 0.471
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.2163899  0.4406896
## sample estimates:
##       cor
## 0.1259382
```

Any hint that the kids who were excluded for fussiness came from a particular portion of your sample? Low-vocab, young? I always wonder once there are more than a couple who fuss out, especially with a relatively large age range like this.

We excluded 8 infants (12.5%) from all analyses due to fussiness (i.e., fewer than 50% of their trials were code-able). In response to your comment, we compared age and vocabulary measures for excluded and included infants, and found no differences in age [t(8.58) = 0.06, p = 0.95], MCDI comprehensive vocabulary size [t(11.26) = -0.63, p = 0.54], or MCDI productive vocabulary size [t(11.71) = -0.40, p = 0.695]. We included this information as a footnote on p. 5 of the manuscript.

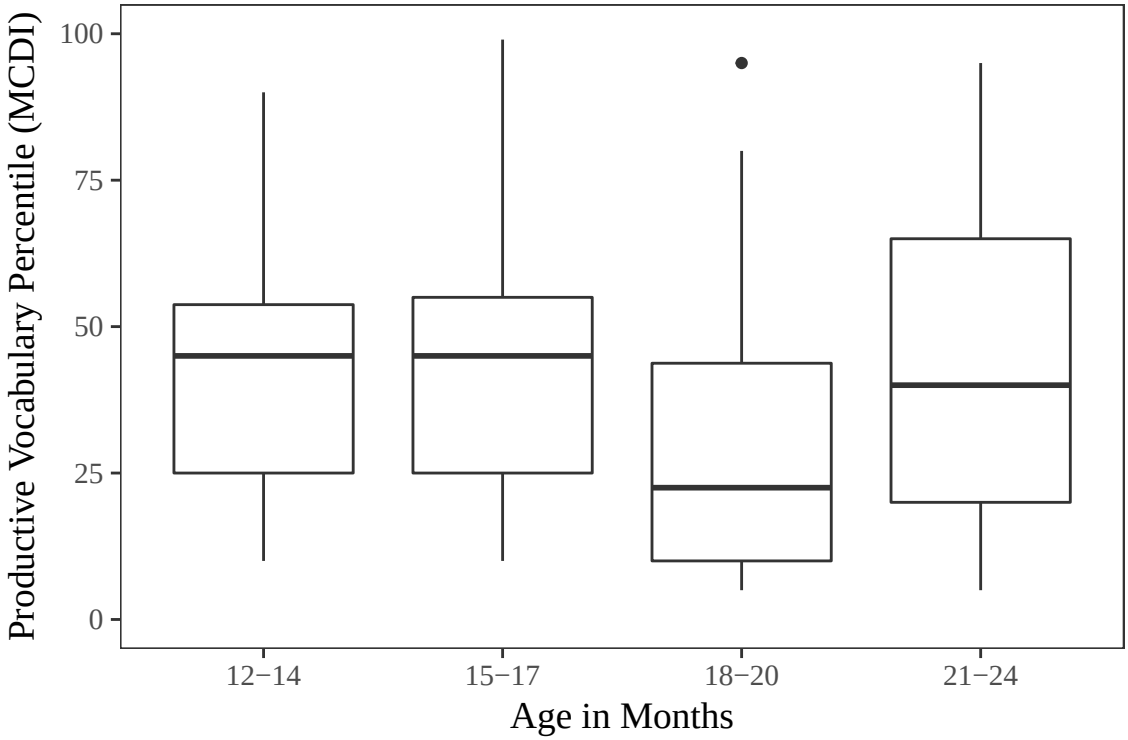
```
##
##  Welch Two Sample t-test
##
## data:  excluded.kids$Age and included.kids$Age
## t = -0.063606, df = 8.5816, p-value = 0.9507
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.830427  3.622427
## sample estimates:
## mean of x mean of y
##    18.500    18.604
```

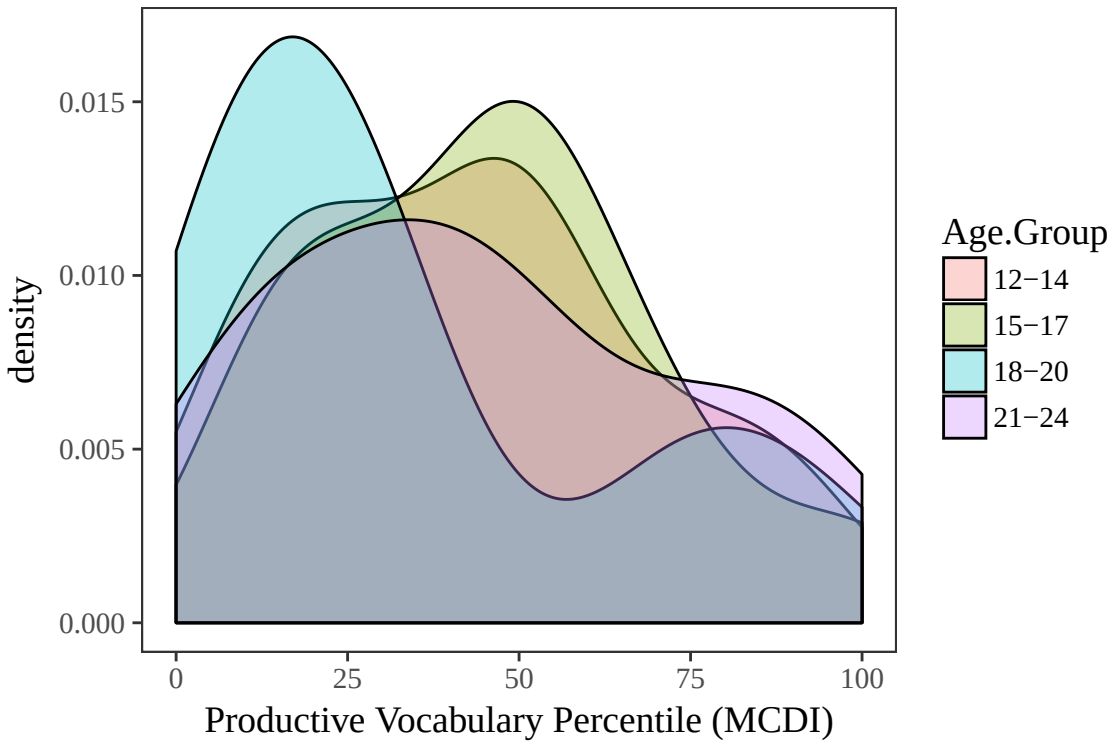
```
##
##  Welch Two Sample t-test
##
## data:  excluded.kids$MCDI.Total.Words.Understands and included.kids$MCDI.Total.Words.
## t = -0.62915, df = 11.256, p-value = 0.5418
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -132.77545   73.61545
## sample estimates:
## mean of x mean of y
##    123.00    152.58
```

```
##
##  Welch Two Sample t-test
##
## data:  excluded.kids$MCDI.Total.Words.Says and included.kids$MCDI.Total.Words.Says
## t = -0.40194, df = 11.705, p-value = 0.695
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -123.21481   84.92481
## sample estimates:
## mean of x mean of y
##   103.875   123.020
```

The combination of analysis using correlation, the use of standardized vocabulary scores, and the large age range make me wonder about the distribution of standard scores across ages. Do you have a good distribution of standard vocab scores in, say, each quartile of your age range?

To address this question, we plotted the distribution of MCDI scores in each quartile of our age range. While the 18-20-month-olds show slightly lower standard vocabulary scores, as you can see below, the overall range is comparable across quartiles.




density
0.015
0.010
0.005
0.000
0
25
50
75
100
Productive Vocabulary Percentile (MCDI)
Age.Group
12−14
15−17
18−20
21−24

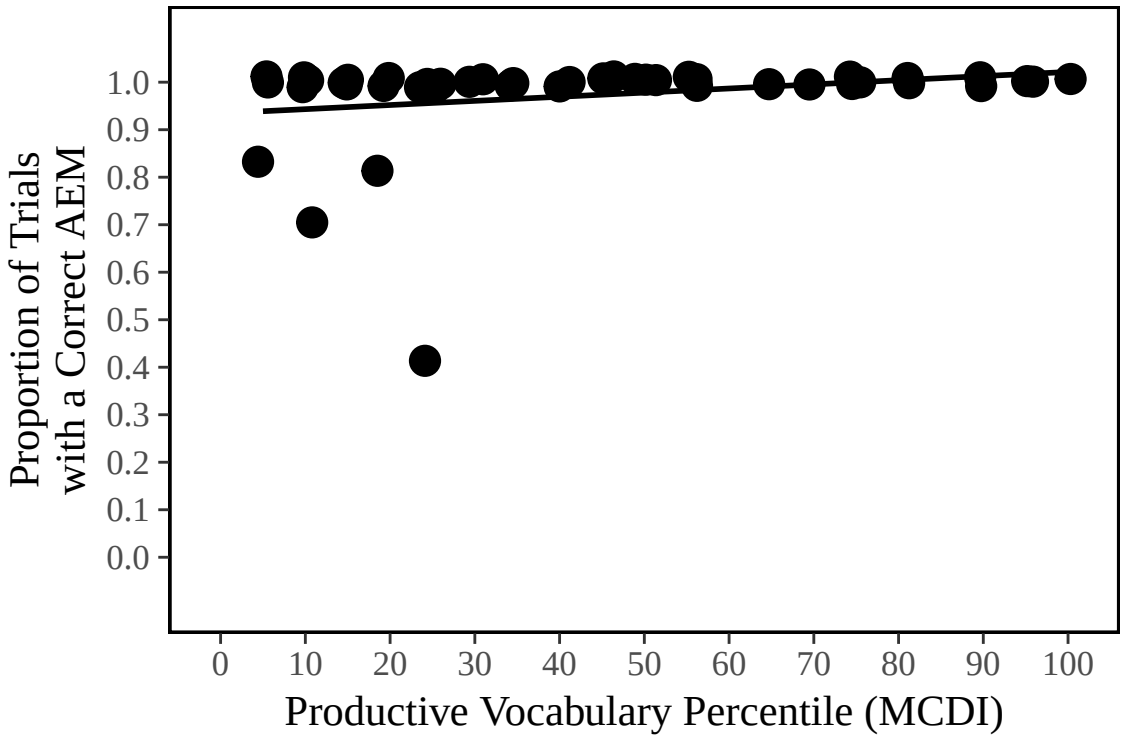
The negative r's for age and likelihood of an AEM (though non-significant) are interesting. Were older kids less engaged by the task?

To answer this question, we correlated each infant's proportion of trials included in analyses with their age in months. We found no correlation between age and number of trials included [r(48) = -0.06, p = 0.65], suggesting that older infants were not less engaged in the task. We appreciate that you generated many careful ideas for further data analysis.

```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$Age and corr.data$prop
## t = -0.46123, df = 48, p-value = 0.6467
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.3385150  0.2159136
## sample estimates:
##         cor
## -0.06642628
```

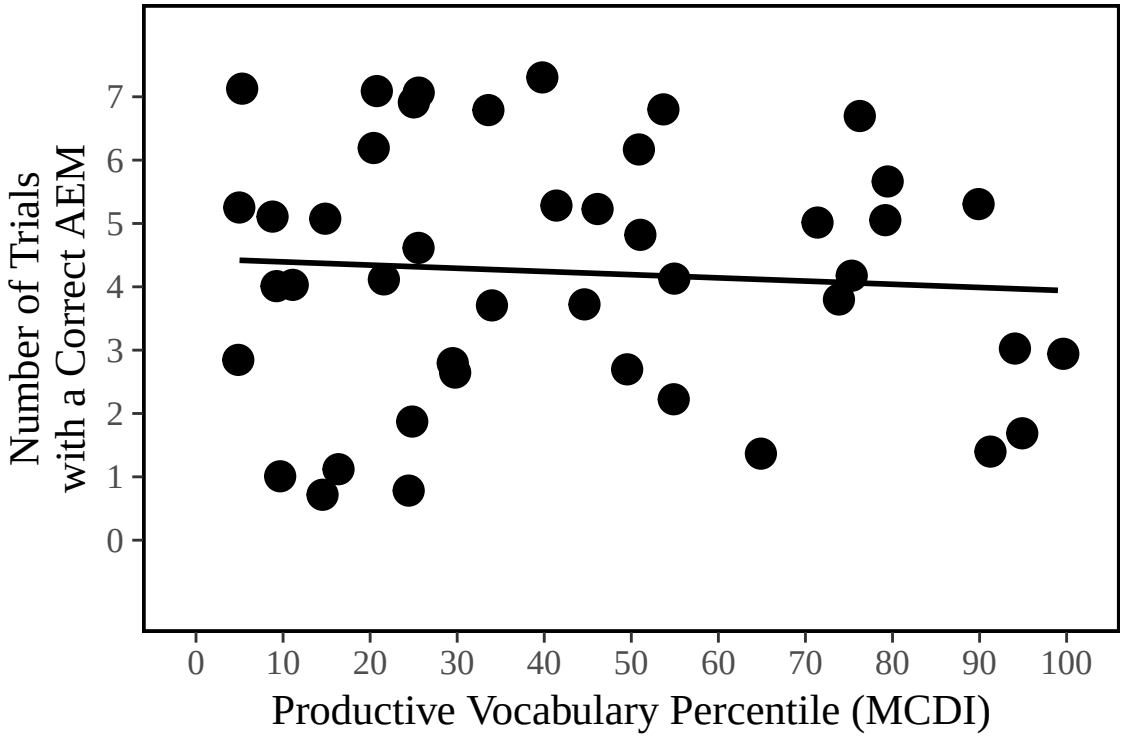
I am not clear why the authors did not measure children's proportion of AEMs to the correct location for trials 2-8. It is clear why looking at proportion of AEMs to the novel location on trials 10-16 provides us with a measure of children's updated anticipations, but it seems it would be helpful to have a baseline for low vs. high vocabulary children on how well they learn to make the primary anticipations that one would assume they are making in the first block of trials.

For the sake of brevity in the manuscript, we reported minimal results from Block 1. Specifically, on p. 9, we reported the null correlation between infants' proportion of trials with an AEM in Block 1 and vocabulary size [r(48)=-0.03, p=0.82]. However, addressing your question, we analyzed the proportion of infants' AEMs to the correct location in Block 1 (trials 2 through 8, excluding infants who did not make any AEMs) and again found a null correlation between infants' Block 1 AEMs and vocabulary size [r(42)=0.24, p=0.12]. We also find that the vast majority of infants' AEMs in Block 1 (189 of 193 total AEMs, 98%) were to the correct peripheral target location, indicating a ceiling effect. Lastly, we also analyzed AEMs trial-by-trial in Block 1, and found no differences in performance among high-vocabulary and low-vocabulary infants. These analyses are all included in supplementary materials, and we also added a sentence on p. 9-10 of the manuscript to address this comment.




```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$MCDI and corr.data$correct.AEM
## t = 1.6112, df = 42, p-value = 0.1146
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
```

```
##  -0.05990331  0.50217830
## sample estimates:
##       cor
## 0.2412678
```




```
##
##  Pearson's product-moment correlation
##
## data:  corr.data$MCDI and corr.data$AEM.count
## t = -0.48201, df = 42, p-value = 0.6323
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.3630564  0.2277244
## sample estimates:
##         cor
## -0.07417039
```