

Manybabies1 Test-Retest Supplementary Information

1

2

3

4 S1: Relationship between the number of trials infants contribute in each session

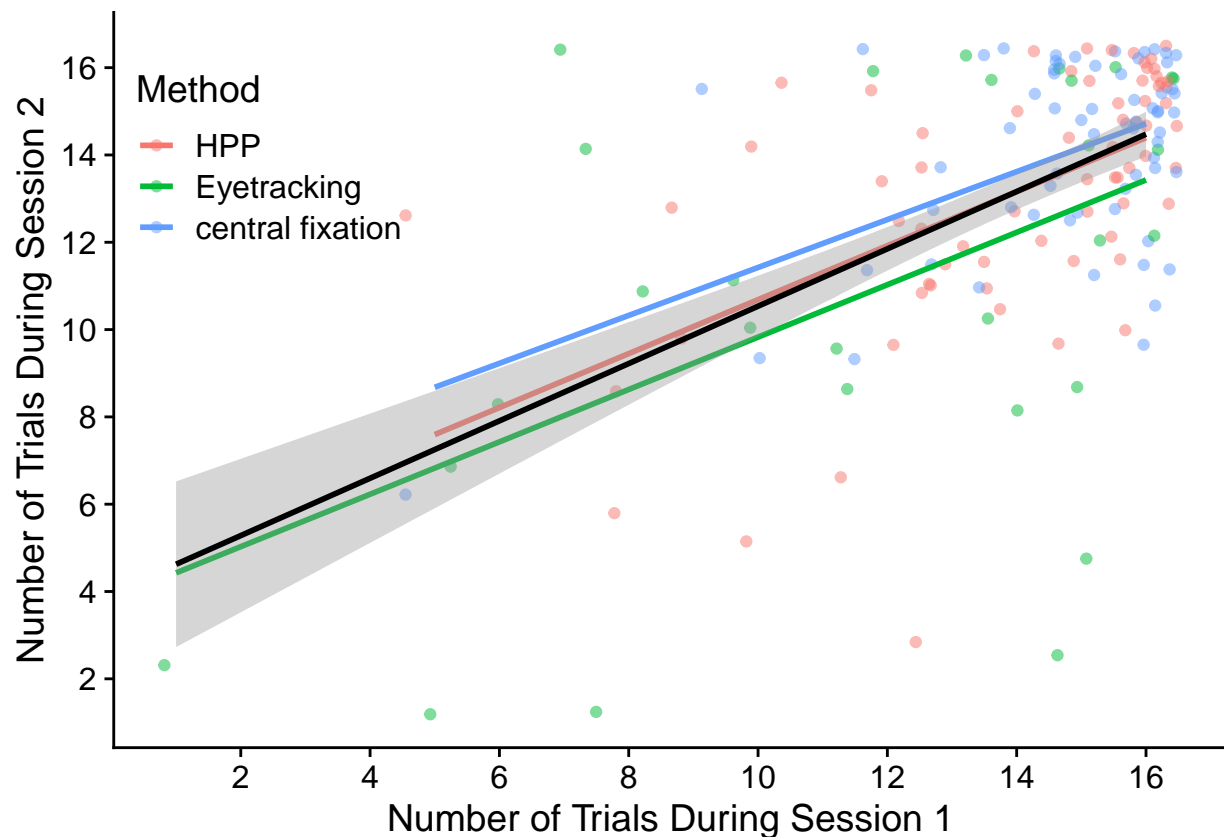


Figure 1. Correlation between the number of trials contributed in session 1 and session 2. Each data point represents one infant. Colored lines represent linear fits for each method.

Are there stable individual differences in how likely an infant is to contribute a high number of trials? To answer this question, we conducted an exploratory analysis investigating whether there is a relationship between the number of trials an infant contributed in session 1 and session 2. Do infants who contribute a higher number of trials during their first testing session also tend to contribute more trials during their second testing session? A positive correlation between trial numbers during the first and second session would indicate that there is some stability in a given infant's likelihood of remaining attentive throughout the experiment. On the other hand, the absence of a correlation would indicate that the number of trials a given infant contributes is not predictive of how many trials they might contribute during their next session.

We found a strong positive correlation between number of trials contributed during the first and the second session $r = .58$, 95% CI $[.47, .68]$, $t(159) = 9.05$, $p < .001$ (see Figure 1). This result suggests that if infants contribute a higher number of trials in one session, compared to other infants, they are likely to contribute a higher number of trials in their next session. This finding is consistent with the hypothesis that how attentive infants are throughout an experiment (and hence how many trials they contribute) is a stable individual difference, at least for some infant looking time tasks. Researchers should therefore be mindful of the fact that decisions about including or excluding infants based on trials contributed may selectively sample a specific sub-set of the infant population they are studying (Byers-Heinlein, Bergmann, & Savalei, 2021; DeBolt, Rhemtulla, & Oakes, 2020).

S2: Preference reversal

48.48 percent of the infants had a preference reversal from test to retest session.

S3: Deviations from the preregistration

Below, we document all deviations from the preregistered methods and analyses <https://osf.io/v5f8t>.

- All infants with usable data for both test and retest session were included in the analyses, regardless of the number of total of infants a lab was able to contribute after exclusion. This decision is consistent with past decisions in ManyBabies projects to be as inclusive about data inclusion as possible (ManyBabies Consortium, 2020).
- A small number of infants with a time between sessions above 31 days were also included in the analyses ($n=2$).
- Consistent with analytic decisions in ManyBabies 1 (ManyBabies Consortium, 2020), total looking times were truncated at 18 seconds (the maximum trial time) in the

38 small number of cases where recorded looking times were slightly greater than 18s
39 (presumably due to small measurement error in recording infant looking times).

References

- Byers-Heinlein, K., Bergmann, C., & Savalei, V. (2021). Six solutions for more reliable infant research. *Infant and Child Development*, e2296.
- DeBolt, M. C., Rhemtulla, M., & Oakes, L. M. (2020). Robust data and power in infant research: A case study of the effect of number of infants and number of trials in visual preference procedures. *Infancy*, 25(4), 393–419.
- ManyBabies Consortium. (2020). Quantifying sources of variability in infancy research using the infant-directed-speech preference. *Advances in Methods and Practices in Psychological Science*, 3(1), 24–52.