Infants' curiosity impacts cognitive capacity in early childhood

Eline R. de Boer[1*] Francesco Poli,[1,2] Marlene Meyer,[1] Rogier B. Mars,[1,3] Sabine Hunnius[1]

[1]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen 6525 XZ, The Netherlands, [2]MRC Cognition and Brain Sciences Unit, University of Cambridge CB2 1TN, Cambridge, United Kingdom, [3]Wellcome Centre for Integrative Neuroimaging, Centre for Functional MRI of the Brain (FMRIB), Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, University of Oxford, Oxford OX1 2JD, United Kingdom.

[*]Corresponding author, eline.deboer@donders.ru.nl

Abstract

Recent research has shown that infants are curious and actively seek out situations from which they can learn. When presented with different stimulus sequences on a computer screen, babies tend to allocate their attention to stimuli that offer opportunities for information gain. Interestingly, however, the degree to which attention is guided by information gain varies among individual infants. Our study provides empirical evidence that these early interindividual differences in infants' curiosity have long-term consequences for their cognitive development. We found that the extent to which infants' attention was guided by information gain at 8 months predicted their IQ scores at 3½ years of age. These findings are the first to demonstrate the lasting consequences of early existing differences in curiosity-driven exploration for later childhood cognitive development.

Keywords: Cognitive development, Curiosity, Infants, Intelligence, Longitudinal

Infants are extraordinarily quick learners who swiftly acquire knowledge about the world. In the past decades, researchers have dedicated extensive efforts to understanding which information-processing strategies underlie this rapid learning (e.g., 1-3). Several recent studies have shown that infants are sensitive to whether stimuli around them are informative (4-8) and sample their surroundings to maximize the amount of information they can gain (9). This intrinsic sensitivity towards information gain has been identified as a key facet of curiosity (10-12) and has fueled recent theoretical and empirical investigations of its role for learning and cognitive development (13, 14). Yet, evidence on whether early curiosity has a positive impact on later developmental outcomes is still lacking.

Recent experimental and computational work has shown that the intrinsic sensitivity to information is not equal across infants. Some infants display a greater inclination to attend to new information than others (15), as evidenced by their prolonged looking towards information-rich compared to information-poor stimuli (Fig. 1v). This raises the question whether individual variations in infants' sensitivity to information are predictive of later cognitive functioning.

In this longitudinal study (preregistered at OSF: osf.io/u876h), we examined whether variations in infant curiosity – measured as sensitivity to information – predict cognitive outcomes in childhood. After measuring infants' curiosity at 8 months of age (Fig. 1i-v), we assessed their cognitive outcomes at 3.5 years using an intelligence test (WPPSI-IV-NL; 16; Fig. 1vi-vii). We suggest that infant curiosity is a key factor underlying positive cognitive outcomes and hypothesized that infants' early sensitivity to information predicts childhood IQ. The study outline is depicted in Figure 1.

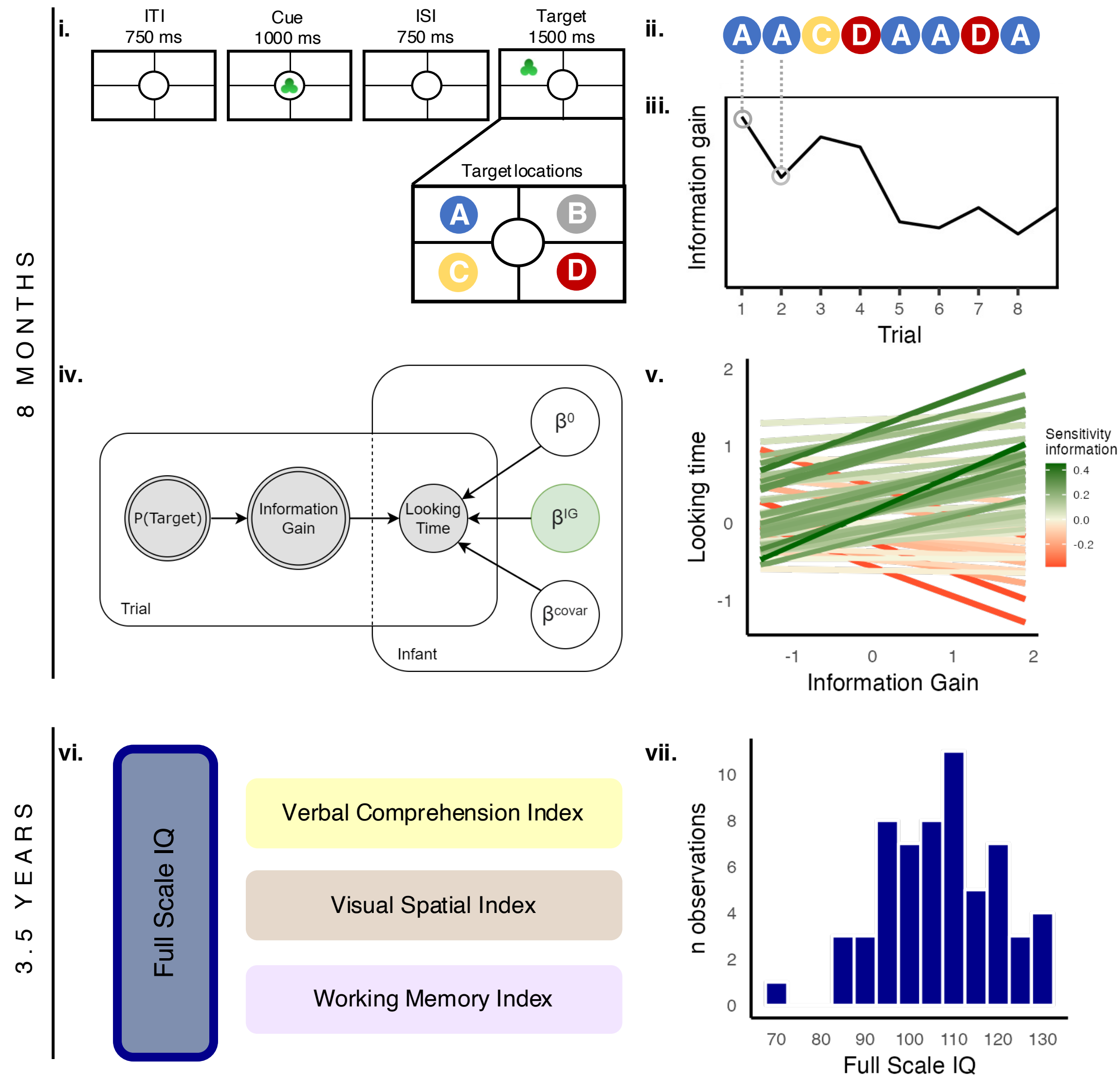

Fig. 1. Study outline. i. Infants were shown multiple sequences of cue-target trials, with each sequence featuring a different stimulus shape (in this example, a trefoil). The cue first appeared in the center of the screen and then was followed by an identical target stimulus in one of four possible locations (i.e. A-D). ii. The probabilistic structure of each sequence was statistically manipulated so that the target was more likely to appear in one location (in this example, location A) than in the others. iii. This resulted in each target stimulus providing a different amount of information about the sequence's probabilistic structure. This measure of information gain was quantified using KL divergence. iv. A hierarchical Bayesian model assessed infants' sensitivity to information by relating their looking time to the target stimulus with the information gain provided by the stimulus. v. Overall, infants' looking times were modulated by the information gain offered by the stimuli. However, individual infants displayed varying degrees of sensitivity to information (i.e., different slopes). A stronger positive relation indicates that infants tailored their looking time more to the available information gain, reflecting more curiosity-driven learning. vi. At 3.5 years of age, participants' cognitive abilities were assessed using the standardized WPPSI-IV-NL intelligence test which provides a general intelligence factor (Full Scale IQ) along with three subindices. vii. The Full Scale IQ that was used as our main outcome measure showed sufficient spread (n = 60).

Results

To link individual differences in infants' curiosity to their later cognitive capabilities, we used a generalized additive model relating measures of sensitivity to information with childhood IQ while controlling for differences in socioeconomic status (SES) between the infants' families. Infants' sensitivity to information was assessed through the degree to which infants' looking time to the stimulus correlated with the information gain that was provided by that stimulus (Fig. 1v). This value was standardized, with higher values indicating higher sensitivity to information. Childhood IQ was measured as the score on the Full Scale Intelligence Quotient (Full Scale IQ) of the WPPSI-IV-NL (Fig. 1vii, $M = 106$, $SD = 12.6$). SES was determined by the average education level of the caregiver(s). In 93% of the cases, at least one of both caregivers completed a form of higher education (higher vocational or university education) and in 63% of the cases, both caregivers completed a form of higher education. Results are depicted in Figure 2.

Individual differences in infants' sensitivity to information significantly predicted childhood IQ ($F = 3.87$, edf (effective degrees of freedom) $= 3.584$, $p = .006$) with the model explaining 22% of the variance in childhood intelligence. The number of effective degrees of freedom was significantly different from 1, indicating that the observed effect of infants' sensitivity to information on their later IQ was non-linear: While greater values of infants' sensitivity to information gain were associated with higher childhood intelligence, lower values of infants' sensitivity to information gain were not necessarily related to lower childhood intelligence scores (Fig. 2). SES did not have a significant effect on IQ ($t = 0.98$, $p = .33$). We also explored the relationship between infants' sensitivity to information and each intelligence subindex separately. There was a significant non-linear relationship between infants' sensitivity to information and their later Verbal Comprehension Index (VCI; $M = 106$, $SD = 15.6$, $F = 2.96$, edf $= 3.23$, $p = .028$). The relationships between infants' sensitivity to information and the Visual Spatial Index (VSI; $M = 103$, $SD = 11.3$) and Working Memory Index (WMI; $M = 109$, $SD = 9.5$) followed a similar non-linear pattern but were not significant ($F = 2.05$, edf $= 2.88$, $p = .108$ and $F = 1.2$, edf $= 2.97$, $p = .286$).

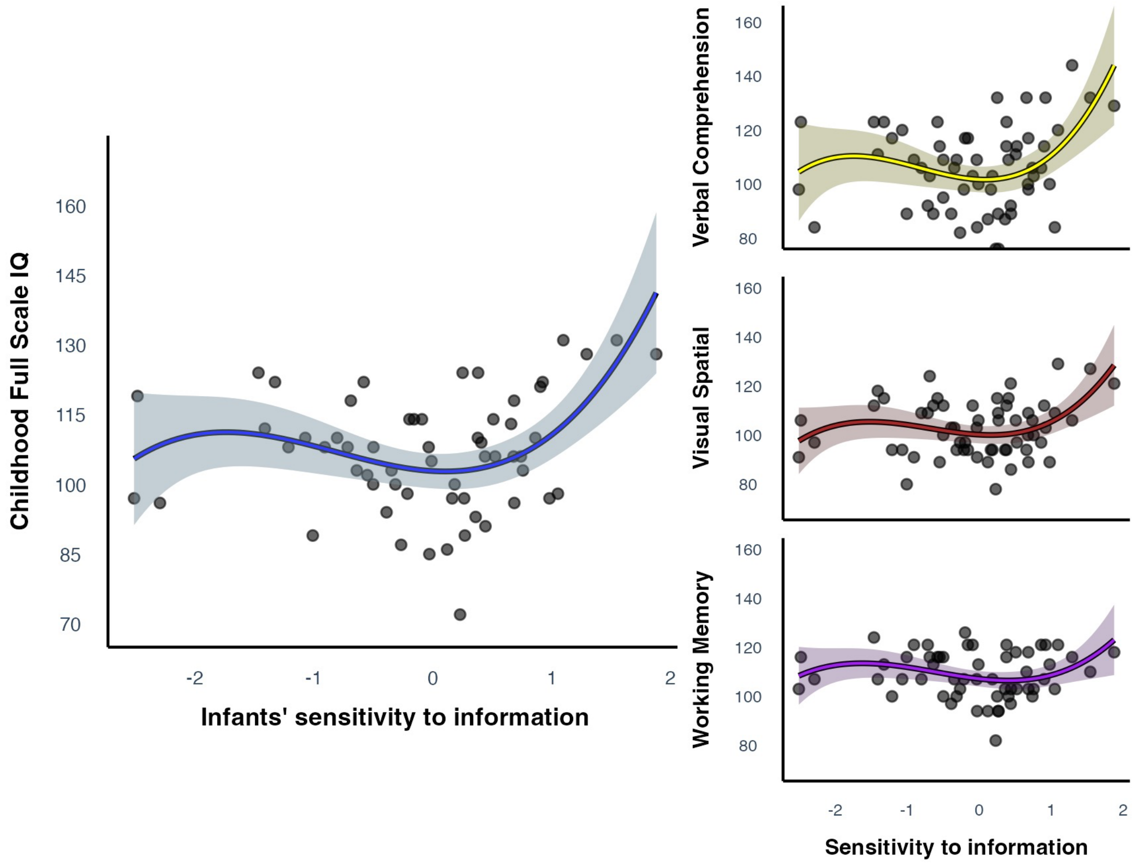

Fig. 2. Results of the generalized additive model relating infants' sensitivity to information with childhood Full Scale IQ and intelligence sub-indices. Infants' sensitivity to information was significantly related to the Full Scale IQ scores in childhood (p=.006, n=59). The relationship between infants' sensitivity to information and the IQ subindices followed a similar non-linear relationship. Of those, only the relationship between infants' sensitivity to information and the Verbal Comprehension Index was significant.

Note. The shaded areas surrounding the plots represent confidence intervals.

Discussion

In this study, we examined the longitudinal implications of individual differences in early manifestations of curiosity for subsequent cognitive development. We tested 8-month-old infants on a visual learning task and extracted interindividual differences in their sensitivity to information using a hierarchical Bayesian model. Three years later, we tested the same children on their intelligence. We show that infants' sensitivity to information predicts childhood cognitive abilities over a period of almost three years, underscoring the pivotal role that curiosity might play in early learning.

Our findings suggest that children who displayed greater curiosity as infants tend to have a more favorable cognitive development. This implies a potentially cascading effect, where infants who pay more attention to stimuli they can learn from early in development, encounter more learning opportunities, thereby enhancing their cognitive development over time (17, 18). Consistent with this interpretation, we found that this effect was mostly driven by the relationship between infants' sensitivity to information and their later verbal comprehension, as opposed to their visual-spatial reasoning and working memory. Whereas the latter two represent more fundamental cognitive abilities and in the WPPSI-IV-NL are assessed through solving a puzzle, recreating geometric patterns, or recalling a series of visual stimuli, the assessment of verbal comprehension mainly relies on world knowledge and vocabulary (16). This supports our notion of a cascading effect: Children who display more curiosity towards their surroundings at a young age are likely exposed to more relevant language input (e.g. through the contingent verbal responses of their caregivers, 19), which in turn aids their understanding and acquisition of words and facts (20, 21).

While we find that greater levels of curiosity promote positive cognitive outcomes, we also find that infants displaying lower levels of curiosity are not necessarily at a disadvantage. Curiosity might thus act as a boost factor – enriching further development – but not as a risk factor, that is, a lack of curiosity does not necessarily lead to lower cognitive capacity. However, it is important to consider that our sample was biased towards children with a higher-than-average IQ: only 10% of our participants had intelligence scores that were below average, and those were only minimally below average (the clinical cutoff for below-average IQ scores on the WPPSI-IV-NL is typically set at 90; 16, 22). Moreover, our sample consisted predominantly of children growing up in high SES households. Although our findings do not show evidence for a relation between SES and IQ – potentially due to the low variability in SES within our sample –, previous studies did indicate that children who grow up in higher SES households tend to have higher IQ scores (23, 24). Future empirical studies examining more diverse samples of children will help in refining our understanding of the relationship between infant's curiosity and later IQ.

Until recently, individual differences in fundamental cognitive mechanisms early in life were difficult to quantify (25) and variability was therefore often discarded. Poli and colleagues (9, 15) introduced a hierarchical Bayesian model to harness these variations in early behavioral measures to extract meaningful differences in cognitive processes that are more precise than what could be perceived directly from the behavioral data alone. Here, we adopted this novel model-based approach to obtain individual differences in infants' sensitivity to information – as inferred from their looking behaviors in a visual learning task – to examine their relationship with later cognitive outcomes. We show that this approach can be applied to predict real-world developmental outcomes from individual differences in processing abilities in infancy using a longitudinal design.

In a nutshell, how do our findings further our understanding of cognitive development and the role of infants' processing abilities therein? Prior research demonstrated that infants as young as 8 months selectively allocate their attention towards stimuli from which they can gain information and that infants display individual differences in this mechanism (9, 15). Building on this, we show that early-existing interindividual differences in curiosity-driven learning play a key role in cognitive development and allow predicting differences in cognitive capacity over a time span of almost three years. Benefitting from this discovery, finding ways to stimulate curiosity might be a promising avenue for boosting exploratory behavior and supporting learning in early childhood.

Materials and Methods

This longitudinal study was a follow-up of the experiment by Poli and colleagues (9). Infants' sensitivity to information was measured at 8-months of age. Intelligence scores were obtained from the same and several additional participants at 3.5 years of age. The study was approved by the local ethics review board (Ethical approval number: ECSW-2020-096) and preregistered at OSF (osf.io/u876h). We deviated from the preregistered analyses in three ways: We used the continuous estimate looking time instead of the dichotomous estimate look-away to determine infants' sensitivity to information gain, applied an additive instead of a linear model in our main analysis, and we added an outlier removal procedure. The rationale behind these deviations is outlined in more detail below.

Participants

Eye-tracking data were gathered from a sample of ninety infants (27). Of the ninety 8-month-olds, twenty did not provide usable data due to poor calibration or fussiness. At 3.5 years of age, intelligence scores were obtained from the same participants. Eight children dropped out for practical reasons such as relocation or time constraints, and two were excluded from the analyses due to incomplete administration of the intelligence test. The final sample consisted of sixty children. The mean age (in months) of participants at the initial test was M = 7.9 (SD = 0.39) and M = 43.0 (SD = 0.36) at the follow-up. The distribution of gender was equal (50% boys, 50% girls).

Measures

Sensitivity to information at 8 months. Infants' sensitivity to information was measured with a visual learning task during which looking behaviors were recorded using eye-tracking (Fig. 1i). The task was presented on a video screen that was divided into four same-sized quadrants (target locations) with in the middle a circle (cue location). The task consisted of 16 sequences, each featuring one of eight unique stimuli (each stimulus was presented twice). Each sequence was composed of 15 trials, with each trial consisting of 4 phases: a cue phase (1000 ms) during which a stimulus was displayed in the cue location; an interstimulus phase (750 ms) during which the stimulus was not visible; a target phase (1500 ms) during which the stimulus was displayed in one of four target locations; and an intertrial interval (750 ms) during which the stimulus was not visible before it was presented again during the cue phase. All 15 trials were shown unless the infant looked away from the screen for longer than one second, triggering the sequence to stop. Once the infant looked back at the screen, the following sequence would start.

The location where the stimulus would reappear (see Fig. 1ii for an example) during the target phase was statistically manipulated following three scenarios. In each scenario, the target location was predictable – it was thus possible for the infants to learn the most likely target location for each sequence – but predictability levels (100%, 80% or 60%) varied across scenarios. In 4 out of 16 sequences, the stimulus always appeared at the same target location (100% predictable). In 6 out of 16 sequences, the stimulus appeared at the same target location 80% of the trials while appearing randomly at one of the three other locations for the remainder 20%. In the remaining 6 out of 16 sequences, the stimulus appeared at the same target location 60% of the trials while appearing randomly at one of the three other locations for the remainder 40%.

For each trial, we quantified the information gain it offered (Fig. 1iii) using Kullback-Leibler Divergence (KL divergence):

$$D_{KL}\left(p^{j} \parallel p^{j-1}\right) = \sum_{k=1}^{K} p\left(x^{j} = k \middle| X^{j}, \alpha\right) log_{2} \frac{\left(x^{j} = k \middle| X^{j}, \alpha\right)}{\left(x^{j} = k \middle| X^{j-1}, \alpha\right)}$$

For each trial, KL divergence computes how much the new event (the stimulus appearing in the target location in the current trial) changes prior probabilities about the target location. For example, before the first trial, the stimulus is equally likely to appear in any four locations (probability is 25%). When in the first trial, the stimulus appears in the top left corner (location A in Fig. 1i), the probability of this location increases (e.g. to 40%) while the probability of the other three locations decreases (e.g. to 20%). The more the new event changes these prior probabilities, the more information gain it offers. More information about these computations can be found in Poli and colleagues (9).

Finally, to determine infants' sensitivity to the information of each trial, we measured their looking time to the target stimulus on each trial. Then, by using a hierarchical Bayesian model (Fig. 1iv), we correlated the looking time towards each trial with the information gain of that specific trial. Infants displayed individual differences in the extent to which their looking time to the target stimuli was correlated to the stimuli's information gain (Fig. 1v).

As described in our preregistration, we initially planned to use the estimate of look-away (a dichotomous measure indicating whether the infant kept looking until the end of the trial or looked away before the trial ended) in relation to information gain as our index of infants' sensitivity to information. However, while look-away is a good measure to estimate group effects (26) looking time turned out to be a more refined measure for studying individual differences (15). Thus, based on incremental insights from the study by Poli and colleagues (15), we decided to use looking time during target presentation instead. We made this decision prior to the start of our analyses.

IQ at 3.5 years. We assessed children's IQ with the Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition - Dutch version (WPPSI-IV-NL; 15) for children aged between 2.5 and 4 years. This standardized intelligence test comprises seven distinct tasks that provide insights into participants' Full Scale Intelligence Quotient (Full Scale IQ) and three intelligence subindices, namely Verbal Comprehension Index (VCI), Visual Spatial Index (VSI) and Working Memory Index (WMI) (Fig. 1vi). In our primary analysis, we used the FSIQ, considered to be a representative indicator of global intellectual functioning, as our main outcome measure (15).

Socioeconomic status. For both caregivers (if applicable, N=1 child was raised by a single caregiver), we collected information on their highest level of education, ranging from 1 (secondary education), 2 (intermediate vocational education), 3 (higher vocational education) to 4 (university education).

Procedure

Families were invited to the Baby and Child Research Center in Nijmegen, The Netherlands, twice – once when their child was 8 months old and once when their child was 3.5-years old. At the first visit, participants' eye movements were measured during the visual learning task through which infants' sensitivity to information was assessed. During the task, infants sat in a baby seat that was positioned on the parent's lap in front of the eye-tracker (Tobii X300) monitor that displayed the stimuli. Caregivers were instructed to refrain from interacting with their child and, when their infant sought their attention, not to attempt to redirect their attention back to the screen. The experiment ended when the infant had watched all 16 sequences or when they became fussy.

At the second visit, the WPPSI-IV-NL was administered to assess children's intelligence. During the test, the child sat behind a table across from an experimenter. The caregiver was positioned at a table behind the child ensuring proximity while minimizing their influence on the child's behavior. While the researcher administered the intelligence test with the child, the caregiver completed a set of questionnaires on a laptop. The intelligence test lasted approximately 45 to 60 minutes. Approximately halfway through the session, typically after completing task five or earlier if the child displayed signs of fatigue, a 15-minute break was allocated. Additional information was collected, including a computer-based game that was played by the child after the intelligence test and a questionnaire that was filled in by the caregiver while the intelligence test was administered. The results obtained from these measures were part of a different study and fall outside the scope of the present paper.

At both visits, there was some time reserved prior to the experiment to familiarize the child with the experimenter. After the experiment, the caregiver was debriefed and the child received a gift. At the first visit, the parent could choose between a monetary reward of 10 euros or a children's book. At the second visit, the gift consisted of a 'diploma' and either 20 euros or two children's books.

Statistical Analysis

We examined the relation between infants' sensitivity to information with childhood intelligence while controlling for SES using an additive model. In this model, participants' Full Scale IQ was included as dependent variable, infants' sensitivity to information as independent variable, and SES (indexed by the average educational level of caregiver(s)) was added as covariate. Prior to the analyses, we performed an outlier removal procedure for individuals who scored below or over 3 standard deviations from the mean on the sensitivity to information measure. This outlier removal procedure was set before the analysis but was not specified in the preregistration. We decided on the threshold of 3 standard deviations because it carries no negative consequence for type I error (28). It led to the exclusion of one infant. The analysis was performed in R 4.3.1. The data and analysis scripts are available on the Radboud Data Repository at https://data.ru.nl/login/reviewer-2842140730/GFZUZQCAO35XYP7GVA2OQNZWWB7NHWXULIVB55A

Although we preregistered the use of a linear model, based on additional statistical insights gained during previous studies (15) and prior to data analysis, we decided to use an additive model instead. The additive model has the advantage of being more flexible while keeping the desired properties of a linear model. It provides a clearer and more accurate representation of the underlying trends and patterns in the data, thus allowing for the detection of linear effects as well as non-linear ones. Consistently, the additive model shows a better fit (AIC = 459.73) than the linear model (AIC = 471.97) that we preregistered and that yielded no evidence for a linear relation between infants' sensitivity to information and childhood intelligence ($t = 1.12$, $\beta = 2.01$, $p = .27$).

For exploratory purposes, we examined the relation between infants' sensitivity to information and the intelligence subindices VCI, VSI, WMI. For this, we followed the same analytical procedure that we reported for the Full Scale IQ measure.

References

1. R. L. Fantz, Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science* **146**, 668-670 (1964).
2. M. Kavšek, Predicting later IQ from infant visual habituation and dishabituation: A meta-analysis. *J. Appl. Dev. Psychol*. **25**, 369-393 (2004).
3. S. A. Rose, J. F. Feldman, Memory and speed: Their role in the relation of infant Information processing to later IQ. *Child Dev.* **68**, 630-641 (1997).
4. C. Addyman, D. Mareschal, Local redundancy governs infants' spontaneous orienting to visual-temporal sequences. *Child Dev.* **84**, 1137-1144 (2013).
5. M. Bazhydai, G, Westermann, E. Parise, "I don't know but I know who to ask": 12-month-olds actively seek information from knowledgeable adults. *Dev. Sci.* **23**, e12938 (2020).
6. C. Kidd, S. T. Piantadosi, R. N. Aslin, The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLOS ONE* **7**, e36399 (2012).
7. A. Ruggeri, N. Swaboda, Z. L. Sim, A. Gopnik, Shake it baby, but only when needed: Preschoolers adapt their exploratory strategies to the information structure of the task. *Cognition* **193**, 104013 (2019).
8. K. E. Twomey, G. Westermann, Curiosity-based learning in infants: A neurocomputational approach. *Dev. Sci.* **21**, e12629 (2018).
9. F. Poli, G. Serino, R. B. Mars, S. Hunnius, Infants tailor their attention to maximize learning. *Sci. Adv.* **6**, eabb5053 (2020).
10. C. Baer, C. Kidd, Learning with certainty in childhood. *Trends Cogn. Sci.* **26**, 887-896 (2022).
11. C. Kidd, B. Y. Hayden, The psychology and neuroscience of curiosity. *Neuron* **88**, 449-460 (2015).
12. A. Ten, P. Kaushik, P.-Y Oudeyer, J. Gottlieb, Humans monitor learning progress in curiosity-driven exploration. *Nat. Commun.* **12**, 5972 (2021).
13. M. Bahzydai, K. Twomey, G. Westermann, "Curiosity and exploration" in Encyclopedia of Infant and Early Childhood Development (Elsevier BV, 2019), pp. 370-378.
14. F. Poli, J. X. O'Reilly, R. B. Mars, S. Hunnius, Curiosity and the dynamics of optimal exploration. *Trends Cogn. Sci.* S1364661324000287 (2024).
15. F. Poli et al., Individual differences in processing speed and curiosity explain infant habituation and dishabituation performance. *Dev. Sci.* **27**, e13460 (2024).
16. D. Wechsler, *Technical and interpretative manual: WPPSI-IV* (Pearson, New York, 2012).

17. A. S. Masten, D. Cicchetti, Developmental cascades. *Dev. Psychopathol.* **22**, 491-495 (2010).

18. J. M. Iverson, Developing language in a developing body, revisited: The cascading effects of motor development on the acquisition of language. *Wiley Interdiscip. Rev. Cogn. Sci.* **13**, e1626 (2022).

19. C. S. Tamis-LeMonda, Y. Kuchirko, L. Tafuro, From action to interaction: Infant object exploration and mothers' contingent responsiveness. *IEEE Trans. Auton. Ment. Dev.* **5**, 202-209 (2013).

20. N. Kartushina et al., COVID-19 first lockdown as a window into language acquisition: Associations between caregiver-child activities and vocabulary gains. *Lang. Dev. Res.* **2** (2021).

21. L. R. Masek et al., Where language meets attention: How contingent interactions promote learning. *Dev. Rev.* **60**, 100961 (2021).

22. S. A. J. Ruiter, P. P. M. Hurks, M. E. Timmerman, IQ-score is dringend aan modernisering toe. *Kind & Adolescent Praktijk* **16**, 16-23 (2017).

23. S. Von Stumm, R. Plomin, Socioeconomic status and the growth of intelligence from infancy through adolescence. *Intelligence* **48**, 30-36 (2015).

24. R. H. Bradley, R. F. Corwyn, Socioeconomic status and child development. *Annu. Rev. Psychol.* **53**, 371-399 (2002).

25. S. T. Piantadosi, C. Kidd, R. Aslin, Rich analysis and rational models: Inferring individual behavior from infant looking data. *Dev. Sci.* **17**, 321-337 (2014).

26. F. Poli, "Developing models for learning and exploration (Publication No. 302609)", doctoral dissertation, Radboud University, Nijmegen, The Netherlands (2024).

27. F. Poli, T. Ghilardi, R. B. Mars, M. Hinne, S. Hunnius, Eight-month-old infants meta-learn by downweighting irrelevant evidence. *Open Mind* **7**, 141-155 (2023).

28. M. Bakker, J. M. Wicherts, Outlier removal, sum scores, and the inflation of the type I error rate in independent samples t tests: The power of alternatives and recommendations. *Psychol. Methods* **19**, 409-427 (2014).