Development and Psychopathology (2026), **38**, 1677–1691
doi:10.1017/S0954579426101527

CAMBRIDGE UNIVERSITY PRESS

Regular Article

Association between infant and toddler gut microbiota composition and later executive functioning

Venla Huovinen[1], Anna-Katariina Aatsinki[1,2], Eeva Eskola[1,3,4], Giulio Benedetti[5], Heidi Isokääntä[6], Elisabeth Nordenswan[1,3], Eveliina Munukka[7], Santosh Lamichhane[8], Akie Yada[3,4,9,10], Siobhain M. O'Mahony[11,12], Alex Dickens[13], Leo Lahti[5], Hasse Karlsson[1], Riikka Korja[1,3,4], Saara Nolvi[1,3,4] and Linnea Karlsson[1,2,14,15]

[1]FinnBrain Birth Cohort Study, Turku Brain and Mind Center, Department of Clinical Medicine, University of Turku, Finland; [2]Centre for Population Health Research, Turku University Hospital and University of Turku, Finland; [3]The Centre of Excellence for Learning Dynamics and Intervention Research (InterLearn), University of Turku, Turku, Finland; [4]Department of Psychology and Speech-Language Pathology, University of Turku, Finland; [5]Department of Computing, University of Turku, Finland; [6]Research Centre for Infections and Immunity, Institute of Biomedicine, University of Turku, Finland; [7]Turku Clinical Microbiome Biobank, Turku University Hospital and Institute of Biomedicine, University of Turku, Finland; [8]Turku Bioscience Centre, University of Turku and Abo Akademi University, Finland; [9]Department of Psychology, Faculty of Education and Psychology, University of Jyväskylä, Finland; [10]Department of Education, Faculty of Education and Psychology, University of Jyväskylä, Finland; [11]APC Microbiome Ireland, University College Cork, Ireland; [12]Department of Anatomy and Neuroscience, APC Microbiome Ireland, University College Cork, Ireland; [13]Department of Chemistry, University of Turku, Finland; [14]Department of Public Health, University of Turku and Turku University Hospital, Finland and [15]Department of Child Psychiatry, Turku University Hospital, Finland

Abstract

Background: Studies indicate that alterations in gut microbiota composition (GMC) during the first 1,000 days of life are associated with neurodevelopment and further behavioral development. However, research on the associations between GMC and executive functions (EFs) in childhood is scarce. This study aims to improve the understanding of the biological processes underlying behavioral development by exploring the associations between GMC and EFs early in life.

Methods: Study population ($n = 373$) is part of the longitudinal FinnBrain Birth Cohort Study. GMC was analyzed using infant and toddler stool sample 16S rRNA sequencing and targeted and untargeted metabolomic assays. EF was assessed using the Spin the Pots and Snack Delay tasks at 2.5 years and the Spin the Pots task, Delay of Gratification task, EF Touch battery and BRIEF-2 questionnaire at 5 years.

Conclusions: Alpha diversity in infancy was negatively associated with preschool EF. Additionally, EFs differed between microbial groups based on dominant genera. Bacterial genera abundances were related to some EFs, but no associations were found between microbial metabolites and EF. This study is among the first to investigate associations between GMC and EF in childhood, a crucial developmental stage characterized by significant changes in both the brain and microbiota.

Keywords: Child development; executive functions; gut-brain axis; gut microbiota; microbial metabolites

(Received 17 April 2025; revised 13 February 2026; accepted 20 March 2026; first published online 8 May 2026)

Introduction

The human gut microbiota (GM) comprises trillions of microorganisms that have been found to affect host health in a bidirectional manner through neuronal mechanisms (including the vagus nerve and enteric nervous system), immune responses, and the production of metabolites (Cryan et al., 2019). The first few years of life are crucial for GM establishment along with overall rapid growth of the brain and internal organs (Derrien et al., 2019). These simultaneous developmental processes further contribute to the development of metabolic and immune systems as well as cognitive and neurological development. Atypical early neurodevelopment has been found to predict various later psychiatric problems (Monk et al., 2019). Therefore, early childhood has recently been a subject of interest in gut-brain research. Furthermore, gut microbiota composition (GMC) in infancy differs substantially from that of toddler microbiota due to the transition from milk nutrition to solid foods (Stewart et al., 2018), and maturation continues through childhood and adolescence before reaching an adultlike composition (Derrien et al., 2019). Similarly, the human brain undergoes rapid development in childhood, doubling in size from birth to one year old and reaching 80% of adult brain size by 6 years of age (Gilmore et al., 2018). After 2 years, brain maturation continues at a slower pace.

As said, early life represents a period when both GMC and brain development are highly dynamic and particularly sensitive to

Corresponding author: Venla Huovinen; Email: vamhuo@utu.fi

Cite this article: Huovinen, V., Aatsinki, A-K, Eskola, E., Benedetti, G., Isokääntä, H., Nordenswan, E., Munukka, E., Lamichhane, S., Yada, A., O'Mahony, S. M., Dickens, A., Lahti, L., Karlsson, H., Korja, R., Nolvi, S., & Karlsson, L. (2026). Association between infant and toddler gut microbiota composition and later executive functioning. *Development and Psychopathology* **38**: 1677–1691, https://doi.org/10.1017/S0954579426101527

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike licence (https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the same Creative Commons licence is used to distribute the re-used or adapted article and the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use.

https://doi.org/10.1017/S0954579426101527 Published online by Cambridge University Press

environmental influences. Human studies have already reported associations between GMC and early neurocognitive development (e.g., Carlson et al., 2018; Sordillo et al., 2019). Although these findings do not demonstrate causality, they suggest that the parallel trajectories of brain and microbiota development may be functionally linked, which highlights the importance of studying the potential biological markers in these developmental trajectories. This is particularly relevant for executive functioning, which undergoes rapid maturation during early childhood and are critical for the development of later cognitive, behavioral, and emotional outcomes.

Executive functions (EFs) can be defined as higher-order cognitive functions that support effortful and goal-directed behaviors (Diamond, 2013). EFs are required in situations where conscious planning and concentration are required, and more automatic behaviors are typically not sufficient to lead to adaptive outcomes (Diamond, 2013). EFs are typically classified into inhibition, working memory and shifting (Miyake et al., 2000), which refer to the ability to inhibit the most prominent response to produce an alternative answer; to keep items in mind for a short period of time; and to shift the focus of attention between tasks, rule sets and mental states. The development of EFs in early childhood plays a critical role in shaping later cognitive, academic, and behavioral outcomes, highlighting their importance for long-term success in various aspects of life.

EFs can be observed from the second half of the first year onward, and they develop rapidly during childhood (Best & Miller, 2010; Bridgett et al., 2013), with great enhancements occurring between 3 and 5 years of age (Broomell & Bell, 2022). Unlike general cognitive abilities, which demonstrate greater stability early in life, EF development in childhood is highly dynamic and sensitive to various environmental influences (e.g., caregiving, early experiences) and reaches moderate to high stability in adolescence (Bridgett et al., 2015). The brain networks underlying EFs (e.g. executive attention network) are considered highly sensitive to disruptions in development (e.g. environmental influences) (Posner et al., 2014). On the other hand, variation in EF performance in childhood can be due to contextual factors (Miller et al., 2023). This large variation and plasticity in EF skills in early childhood underscores the need to understand different factors associated with this variation, including co-occurring biological and environmental influences.

Importantly, EFs play an important role in daily life and performance with disruptions in normative development during childhood potentially leading to diverse health implications later in life (Diamond, 2013; Yang et al., 2022). Recent meta-analysis found that EFs in childhood were inversely associated with broad internalizing and externalizing problems later (Yang et al., 2022). Additionally, direct links between deficits in EF skills in preschool age to depression and anxiety symptoms in elementary school have been made (Nelson et al., 2018). Overall, EF skills have been found to be essential for both physical and mental health as well as school and social life and overall cognitive, emotional, and psychological development (Diamond, 2013), underlying the relevance of EF-GMC-connections as a study topic since understanding the intricate interplay between EFs and GMC in early life may provide valuable insights into well-being later in life.

Currently, the potential associations between early GMC and later EFs in young children remain largely unexplored, without a few notable exceptions. Fujihara et al. (2023) found that toddlers who had higher developmental risk for inhibitory control problems ($SD = 1.5$ higher than the mean score in BRIEF-P) had higher abundances of genera *Actinomyces* and *Sutterella*. Additionally, Willemsen et al. (2023) found that alpha diversity at two weeks of age was associated with better performance in an EF task at three years of age. Furthermore, the presence of *Parabacteroides* at two weeks was associated with enhanced EFs and less externalizing behaviors at three years, while the presence of *Streptococcus* at two weeks was related to poorer EFs and increased externalizing behaviors at the same age. In contrast, a study of older children, Eckermann et al. (2022) did not find any associations between child fecal microbiota in infancy and the ages of 6 and 10 years and EFs at ages 8 and 10 years. They speculate that the lack of associations might be due to a smaller, low-risk sample. Furthermore, although not a study on EFs as such, Tamana et al. (2021) studied early cognitive and language skills in general and found that higher abundance of *Bacteroides* in late infancy was associated with enhanced cognitive and language skills at 2 years of age. Additionally, it is noteworthy that previous work on general cognitive development, such as Carlson et al. (2018), has demonstrated associations between early microbiome composition and broader cognitive measures, highlighting the potential broader context of early gut–brain interactions. In general, these findings highlight the potential associations between early GMC and EFs, although inconsistencies across studies underscore the need for further investigation that accounts for timepoints and sample characteristics that may moderate these associations.

Additionally, it is important to note that previous research in gut-brain associations has revealed sex specificity. Both rodent (Cryan et al., 2019) and human studies (Aatsinki et al., 2019; Tamana et al., 2021) suggest that GMC and neurodevelopment differ between sexes. This highlights the importance of examining the role of biological sex in the associations between GMC and EFs.

Finally, many of the studies in the field to date have been correlational, mainly focusing on the associations between different indicators of GMC and EF outcomes. Despite evidence from animal research suggesting that gut microbiota metabolites (e.g., short-chain fatty acids, SCFAs) may influence neurobehavioral processes (Dalile et al., 2019; Gronier et al., 2018), corresponding studies in humans, particularly linking GM metabolites to EFs, are lacking. As mentioned, research on GMC and EFs in childhood remains scarce while investigating potential associations between early GMC and EFs in young children is important for understanding factors that may contribute to variation in EF development. Given that EFs are foundational to cognitive and behavioral development and their significance for later mental health and a range of other life outcomes, this research could provide valuable insights into early biological correlates of EF and inform future hypothesis-driven and even intervention studies.

In this study, our aim was to examine how GMC and fecal metabolites associate with EFs in early childhood in a general population sample of children. More specifically, we studied whether GMC at 2.5 months and 30 months of age is associated with EFs at 2.5 and 5 years of age. This design allowed us to examine whether early gut microbiota is related to later EFs. Importantly, the first years of life represent a sensitive period for both gut and brain development, making infancy and toddlerhood particularly meaningful ages for investigating links between microbial profiles and subsequent EF outcomes.

We hypothesized that GMC in infancy and toddlerhood would be associated with EF, both cross-sectionally and longitudinally during toddlerhood and preschool age, potentially in a sex-specific manner. Given that existing human research on GMC–EF associations is still scarce, prior work does not support a strong

https://doi.org/10.1017/S0954579426101527 Published online by Cambridge University Press

hypothesis concerning the direction of these linkages. For this reason, we emphasized an exploratory approach, aiming to capture potential associations without restricting the hypotheses too narrowly. However, by focusing on EF, we aimed to examine a developmental marker with known implications for later behavioral and emotional symptomatology, which underscores the potential relevance of early gut–brain associations, as early microbial variation may contribute to developmental pathways reflected in EF-related outcomes. To enhance reliability, we are combining both laboratory assessments across multiple time points with parent-reported measures. This approach allows us to capture both structured task-based performance and everyday EF behaviors, providing a more comprehensive picture of early EF development.

Method

Study sample

Our study sample is part of the FinnBrain Birth Cohort Study (www.finnbrain.fi) that was set up in 2010 to study the effects of early life stress on child development and health (Karlsson et al., 2018). The data was collected in the areas of Turku and Åland Islands in Finland, and the size of the entire cohort is $n = 3{,}837$. A subcohort of this study included children with at least one microbial sample and one EF measurement (overall $n = 373$), and from these children, 277 had fecal samples from 2.5 months and 204 from 30 months. Overall, 108 participants had samples from both time points. Parents provided a written consent on their own and on their child's behalf for participation in the study. The study protocol has been approved by the Ethics Committee of the Hospital District of Southwest Finland. This study followed the principles of the Helsinki Declaration.

Executive functions

EFs were assessed at age 2.5 and 5 years during a developmental research visit at the FinnBrain Child Development and Parental Functioning Lab. The visit included an EF assessment, along with other tasks and observational situations (including tests of temperament, general cognition, emotional attention, and parent–child interaction). The full visit lasted 1.5 hours at 2.5 years and 2 hours at 5 years. Study visits were conducted by clinical psychologists or trained advanced psychology students under the supervision of clinical psychologists. The parental reports of EF at the age of 5 years used in this study were gathered by asking the parent to fill out questionnaires in the waiting room during the child's research visit.

Spin the Pots task

Visuospatial working memory was assessed using the Spin the Pots task (Hughes & Ensor, 2005). The task at 2.5 years included a Lazy Susan Tray with eight visually distinct boxes glued to it. The experimenter hides six stickers in the boxes, covers the tray with an opaque scarf, and rotates it 180 degrees. Subsequently, the child is instructed to find a box with a sticker in it. Covering and rotating the tray is repeated between each trial. The task is discontinued after sixteen tries or when the child finds all six stickers. The final score is calculated by subtracting the number of unsuccessful attempts to find a sticker from the number of trials making the highest possible score in the task 16, with higher scores reflecting better working memory. The spin the Pots task at 5 years was nearly identical to the 30 months version; the difficulty level was increased by having 10 stickers in 12 boxes. Additionally, to increase the level of difficulty, the pots were designed to appear more similar to each other than in the 2.5 years version. The maximum number of trials was 20, making the maximum score in this task 20.

Snack delay task

Inhibitory control at 2.5 years was assessed using the modified Snack Delay task (Kochanska et al., 2000). In the task, the child is asked to keep their hands on the table while the experimenter puts the snack under a clear plastic cup. The child has to wait for the experimenter to ring a bell before taking the snack. The modified task in this study consists of six trials, with the duration of the task varying between 5 to 60 s. Trials were coded on a scale from 0 to 4 (0 = "Child eats the snack before the experimenter has touched the bell and rung it," 4 = "Child does not touch the bell or the cup before the experimenter has rung the bell"). Additionally, up to 2 extra points were given based on how the child was able to keep their hands on the mat (0 = "The child is unable to hold their hands on the mat," 1 = "The child is able to hold one hand on the mat," 2 = "The child is able to hold both hands on the mat") (Spinrad et al., 2007). The highest possible score in the task is 36, with higher scores reflecting better inhibitory control.

Delay of gratification task

Inhibitory control at 5 years of age was assessed using the Delay of Gratification task (Beck et al., 2011). In this task, the experimenter informs the child that they can either get one treat/sticker/penny now or two/four/six of these items later, at the end of the visit. After that, the child chooses the order: one now, four later (coin), one now, two later (sticker), one now, six later (coin), one now, four later (sticker), one now, two later (treat), one now, six later (sticker), one now, four later (treat), one now, two later (coin), one now, six later (treat). Each "later" answer gives one point, with the maximum delay score being 9 points.

EF Touch tasks: arrows, pigs and farmer

At 5 years, the child EF was assessed using three tasks from the computerized EF Touch Battery (Willoughby et al., 2012). Inhibitory control was assessed using the Spatial Conflict Arrows and Animal Go/No-Go (Pigs) tasks. In the Spatial Conflict Arrows task, the child needs to press the side of the screen where the arrow is pointing, despite the location of the arrow. The task consists of 36 items presented in three blocks (congruent, incongruent, and mixed conditions) with 12 items in each. Only the incongruent items were used in this study, since they specifically require inhibition of the prepotent response established during the congruent trials. Each item was displayed for 3,000 ms, and both the accuracy and reaction time of the responses were recorded. Responses faster than 400 ms were considered implausible (likely indicating a response from the previous screen) and were marked as missing (Willoughby et al., 2019). In the Animal Go/No-Go task, the child needs to press the screen when they see an animal, unless the animal presented is a pig, in which case the screen must not be pressed. Here again, each item was displayed for 3,000 ms and responses faster than 400 ms were considered implausible (likely indicating a response from the previous screen) and were marked as missing. Visuospatial working memory was assessed using the Farmer task, where the child needs to remember the order in which animals "walked" on four squared presented on the screen. In order to form factor scores for each EF Touch task, an IRT analysis was conducted to account for item level difficulty and

https://doi.org/10.1017/S0954579426101527 Published online by Cambridge University Press

discrimination properties. In the IRT analysis, the child's response for each item was rescaled in a binary manner (pass = 1, fail = 0). The IRT analysis is described in more detail by Yada et al. (2025).

BRIEF-2

Additionally, EFs at 5 years were assessed by asking the parent to rate the child's EF using Behavior Rating Inventory of Executive Functions, Second Edition (BRIEF-2) (Gioia et al., 2015). The BRIEF-2 consists of 63 child behavior rating items, for which the parents were asked to answer based on the frequency of certain behavior of their child using a Likert scale (0 = Never, 1 = Sometimes, 2 = Often). These items, in turn, form three indices: the Behavior Regulation Index (BRI; including factors Inhibit and Self-Monitor), the Emotional Recognition Index (ERI; including factors Shift and Emotional Control), and the Cognitive Regulation Index (CRI; including factors Initiate, Working Memory, Plan/Organize, Task-Monitor, and Organization of Materials) which further form the Global Executive Composite (GEC) factor used in this study. The GEC was used in this study to capture overall executive functioning in everyday life.

A Confirmatory Factor Analysis (CFA) was conducted to test the one-factor structure of the BRIEF-2. The initial model demonstrated an acceptable fit to the data: $\chi^2(1779) = 12{,}976.771$, $p < .05$; CFI = 0.82; TLI = 0.81; RMSEA = 0.047 (90% CI: 0.045–0.049); SRMR = 0.065. The BRIEF-2 factor structure was modified based on empirical fit indices and item characteristics. Items with very low frequency of endorsement (items 18, 36, 54) or low explained variance (items 55 and 61, $R^2 < 0.20$) were excluded. Additional items were removed due to poor factor loadings and inadequate model fit (items 47, 53, and 13). Modification indices were inspected to identify potential sources of model misfit, and adjustments were made accordingly (e.g., residual covariances with high modification indices were addressed; see Supplementary Table 1 for full details). After applying modifications, the model was refined and showed an improved fit: $\chi^2(1485) = 12{,}070.355$, $p < .05$; CFI = 0.87; TLI = 0.87; RMSEA = 0.042 (90% CI: 0.039–0.044); SRMR = 0.06. Due to the complexity of the model, we used the two-index strategy by Hu & Bentler (1999), where a combination of RMSEA $\leq$ 0.06 and SRMR $\leq$ 0.09 are considered an acceptable fit for the data.

Gut microbiota

At 2.5 months and at 30 months of age, stool samples were collected from children by their parents, either from diapers or a potty, depending on the child's age. Parents were instructed to store the samples in their home refrigerator until the study visit. The samples were transported to the clinic within 24 hours in a cooler bag with cold packs ("refrigerated plastic pouch"). At the clinic, samples were kept refrigerated until processed by the laboratory nurse. In the lab, samples were aliquoted into four cryotubes and stored at −75°C until further analysis. The time from collection to freezing was monitored for each sample, although the precise temperature during transport was not recorded beyond the use of cold packs. To account for potential effects of sample handling, sensitivity analyses were conducted for the 30-month samples by adjusting for the delay between sample collection and freezing, as this information was available at the 30-month time point (see Supplementary Materials).

The GXT Stool Extraction Kit VER 2.0 (Hain Lifescience GmbH) was used to extract bacterial DNA, the concentration of which was then measured with the Qubit dsDNA HS Assay Kit and the Qubit 2.0 fluorometer (Thermo Fisher Scientific). DNA sequencing was performed at the Medical microbiology of University of Turku. The V4 region of 16S rRNA genes was amplified with custom primers, and the resulting gene library was sequenced with the Illumina MiSeq platform, as previously described (Rintala et al., 2018). Sequencing reads were assembled into amplicon sequence variants (ASVs) with DADA2 (Callahan et al., 2016) and taxonomic annotation was performed on the SILVA rRNA database v138 (Quast et al., 2013) as previously described by Keskitalo et al. (2021). Before further analysis, samples were rarefied to the minimum count across all samples to reduce sequencing amplification bias (Weiss et al., 2017). Additionally, ASVs with a prevalence lower than 10% and a relative abundance greater than 0.1% were removed, which has been observed to improve the overall robustness of methods, specifically in differential abundance analysis (Nearing et al., 2022).

Metabolomic assays

The Bile Acids, SCFAs, and untargeted metabolome were previously described and adapted for this publication by (Aatsinki et al., 2025). Briefly, the fecal samples were randomized before extraction. Two aliquots (50 mg each) of each fecal sample were weighed, with one aliquot freeze-dried to determine dry weight and the other homogenized with homogenizer beads and 20 µL of water per mg of dry weight. The samples were frozen to at least −70°C and homogenized for five min using a bead beater. The bile acids (BAs) analyzed included 33 compounds listed previously (Aatsinki et al., 2025). BAs were extracted by adding 40 µL of fecal homogenate to 400 µL of methanol containing internal standards, followed by filtration using a Supelco protein precipitation filter plate. The dried extracts were resuspended in 20 µL of methanol: water (40:60) containing an injection standard. Quality control (QC) samples were prepared by pooling an aliquot from each sample, vortexing, and treating them identically to the other samples. Blank samples and calibration curves were prepared similarly, with standard dilutions ranging from 0.0025 to 600 ppb. The LC separation was conducted on a Sciex Exion AD 30 system with a Waters Aquity UPLC HSS T3 column. Eluent A was 0.1% formic acid in water, and eluent B was 0.1% formic acid in methanol. The gradient increased from 15% to 100% B over 18 min, with a flow rate of 0.5 mL/min and a 5 µL injection volume. A Sciex 5500 QTrap mass spectrometer operating in scheduled multiple reaction monitoring mode was used for detection.

Short-chain fatty acids (SCFAs) were analyzed by modifying a targeted analysis protocol (Trimigno et al., 2017). Fecal homogenates were prepared with water (10 µL per mg dry weight) and homogenized with a bead beater. SCFAs were extracted by crashing 50 µL of fecal homogenate with 500 µL of methanol containing internal standards. The filtered extracts were analyzed on an Agilent 5890B GC system equipped with a Zebron ZB-WAXplus column. Gas chromatography-mass spectrometry (GC-MS) was performed on an Agilent 5977A MSD in SIM mode, with standard curves (0.1 to 100 ppm) prepared for acetic, propionic, butyric, valeric, hexanoic, isobutyric, and iso-valeric acids.

Polar metabolites were extracted using methanol containing internal standards and filtered using protein precipitation plates. The extracts were dried, stored at −80°C, and derivatized with methoxamine and *N*-Methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA). Gas chromatography was conducted on an Agilent 7890B GC system equipped with an Agilent DB-5 MS column. Mass spectrometry was carried out on a LECO Pegasus BT system with

https://doi.org/10.1017/S0954579426101527 Published online by Cambridge University Press

ChromaTOF software for data acquisition. The analysis was performed in nine batches, each including calibration curves, blanks, QCs, and standard samples. Instrument maintenance, including the replacement of septa and liners, was performed between batches.

Covariates

Covariates used in this study include the child's biological sex assigned at birth, duration of pregnancy (gestational age), sampling age, delivery mode, breastfeeding status at the time of sampling, and maternal age, weight and education. Mothers' education and breastfeeding status at 2.5 months and 30 months of child age were acquired by questionnaires. Information about child sex, duration of pregnancy, mothers' age and pre-pregnancy body mass index (BMI), as well as mode of delivery (vaginal or caesarean) was drawn from the Wellbeing County of Southwest Finland registries. Duration of the pregnancy was defined as the number of weeks, while sampling age was defined as the number of days from the participant's birth to the sampling. Delivery mode was categorized as either vaginal birth or cesarean section. Breastfeeding status at the time of sampling is based on maternal reports which were divided into four categories: no breastfeeding, cessation before timepoint, partial breastfeeding or full (exclusive) breastfeeding. Mothers' weight was derived from their body mass index (BMI) and categorized into normal (BMI < 25), overweight ($25 <$ BMI < 35) and obese (BMI > 35). Mother's education was divided into three categories: 1 = up to upper secondary school, 2 = up to tertiary level, university, and 3 = higher than tertiary level.

Statistical analyses

The analytical pipeline was implemented in the R statistical language (R Core Team, 2024) using TreeSummarizedExperiment (Huang et al., 2021) and MultiAssayExperiment (Ramos et al., 2017) data containers, along with the package mia v1.8.0 (Borman et al., 2025). Within-sample diversity, or alpha diversity was calculated in terms of Shannon index and observed richness, whereas between-sample diversity or beta diversity was visualized using Bray-Curtis dissimilarity and Principal Coordinate Analysis (PCoA). For the remaining analyses, samples were agglomerated by genus, and missing taxonomic information at the genus level was replaced with the corresponding family labels.

Associations between alpha diversity and EFs were preliminarily studied using a correlation analysis with the non-parametric Kendall τ coefficient, examining the relationships between each EF score, the Shannon index, and observed richness. After analyzing correlations for the whole sample, we conducted separate analyses for boys and girls to explore potential sex-based differences. Statistically significant correlations were further examined through linear regression analysis, both unadjusted and adjusted for covariates.

The samples were classified based on the most dominant genus. The analysis was performed on genus-level relative abundance data and the most dominant genus was determined for each sample in terms of relative abundance. Genera not among the top 5 genera with the highest mean abundance were classified as "Other". This stratification by the most dominant genus provides a simple and interpretable way to characterize microbial communities, and it has been shown to align with broad community types (Salosensaari et al., 2021). While more nuanced data-driven clustering approaches such as Dirichlet Multinomial Models (DMM) are possible, their interpretation is more complex and, in our data, DMM did not yield additional or clearer groupings in this sample. This approach was chosen to identify patterns in microbial community composition, enabling comparisons between groups that may reveal ecological or functional differences associated with the dominant taxa. Next, differences in EFs between these dominant genus groups for the whole sample and separately for boys and girls were determined with the non-parametric Kruskal–Wallis test by ranks followed by the post-hoc Dunn's test.

To determine whether certain members of the microbial community associate with EFs, we used LinDA (Zhou et al., 2022) to test if any genera were differentially abundant. Each EF score was analyzed with a separate model, once for each time point (2.5 months and 30 months), resulting in a total of 20 tests. Every model was adjusted for covariates and adjusted for multiple comparisons using the Benjamini–Hochberg procedure. Next, the same procedure was conducted for boys and girls separately to examine potential sex differences in these associations.

Lastly, log-transformed metabolite concentrations were correlated with EF scores using Spearman correlation and p-values were adjusted for multiple comparisons with the Benjamini–Hochberg method. Additionally, as a sensitivity analysis, we also adjusted for infant and toddler antibiotic exposure to assess the robustness of the observed associations.

Results

Participant characteristics

Participant characteristics are presented in Table 1. Overall composition of the data did not change remarkably between time points, excluding the breastfeeding status. Almost all participants stopped breastfeeding by 30 months of child age with only one case of ongoing partial breastfeeding. All participants with successfully collected and analyzed stool samples available did not necessarily have successful EF assessment. Since EF tasks were administered separately and at different time points, the participant pool throughout the tasks greatly overlaps, but is not identical. Some children contributed data at more than one time point (e.g., stool samples at 2.5 and 30 months, or EF assessments at 30 months and 5 years). To address this, we conducted separate analyses for each microbiota–EF combination: stool at 2.5 months with EF at 30 months, stool at 30 months with EF at 30 months, stool at 2.5 months with EF at 5 years, and stool at 30 months with EF at 5 years. Each model therefore included only one microbiota–EF pairing per participant. This approach allowed us to maximize the available sample size while avoiding non-independence within a single analysis and enabled us to examine associations across different developmental periods. Figure 1 shows the degree of overlap among participants with GMC and EF data across time points.

EF scores had only weak correlations between each other ($\tau = 0.01$–0.19) (see Supplemental Figure 1). This is consistent with previous studies, which usually show weak to moderate correlations between tasks intended to measure the same performance (Willoughby et al., 2014). In addition, correlations in EF tasks over time potentially reflect the rapid and heterogeneous development of these skills in early life (Miller et al., 2023).

Alpha diversity and executive functions

The Shannon index did not correlate at any of the time points with EF scores. However, the observed richness index at 2.5 months of age correlated negatively with the EF Touch Pigs task at 5 years ($\tau = -0.16$, adj.$p = 0.03$). When testing separately for boys and

https://doi.org/10.1017/S0954579426101527 Published online by Cambridge University Press

Table 1. Participant characteristics

Mean/count (SD/%)		All (373)	Girls (164)	Boys (209)
Microbial sample at 2.5 months		277	124 (44.8%)	153 (55.2%)
Microbial sample at 30 months		204	86 (42.2%)	118 (57.8%)
Sampling age at 2.5 months		65.2 (14.1)	64.6 (13.8)	65.8 (14.3)
	Missing information	1	0	1
Sampling age at 30 months		914 (62.8)	907 (94.1)	918 (16.1)
	Missing information	2	1	1
Breastfeeding status at 2.5 months	Full	189	88	101
	Partial	39	13	26
	Post	6	3	3
	None	3	2	1
	Missing information	40	18	22
Breastfeeding status at 30 months	Full	0	0	0
	partial	1	0	1
	Post	170	73	97
	None	2	1	1
	Missing information	31	12	19
Gestational age (weeks)		39.98 (1.38)	40.1 (1.15)	39.9. (1.53)
	Missing information	3	1	2
Delivery mode	Vaginal	301	131	170
	Caesarean	68	32	36
	Missing information	4	1	3
Maternal BMI		24.51 (4.87)	24.9 (5.29)	24.2 (4.49)
Maternal obesity	Normal	241	102	139
	Overweight	35	18	17
	Obese	93	43	50
	Missing information	4	1	3
Maternal age at birth (years)		31.16 (4.48)	31.0 (4.64)	31.2 (4.36)
	Missing information	0	0	0
Maternal education	1	89	37	52
	2	110	52	58
	3	161	71	90
	Missing information	13	4	9
Antibiotics before 2.5 month stool sample	Yes	5	NA	NA
	No	368	163	205
Antibiotics 6 months prior to 30 month stool sample	Yes	110	39	71
	No	263	125	138
Snack Delay 2.5 years		27.22 (8.65)	28.61 (8.47)	26.14 (8.66)
	Missing information	96	43	53
Spin the Pots 2.5 years		11.82 (3.52)	12.13 (3.37)	11.59 (3.62)
	Missing information	77	35	42
Delay of Gratification 5 years		6.07 (3.07)	5.88 (3.22)	6.23 (2.94)
	Missing information	100	41	59
EF Touch – Arrows 5 years		0.71 (0.28)	0.71 (0.28)	0.71 (0.28)
	Missing information	107	40	67

(*Continued*)

https://doi.org/10.1017/S0954579426101527 Published online by Cambridge University Press

Table 1. (*Continued*)

Mean/count (SD/%)		All (373)	Girls (164)	Boys (209)
EF Touch – Farmer 5 years		0.49 (0.17)	0.49 (0.17)	0.50 (0.17)
	Missing information	142	51	91
EF Touch – Pigs 5 years		0.90 (0.16)	0.49 (0.17)	0.89 (0.18)
	Missing information	114	42	72
Spin the Pots 5 years		10.91 (3.02)	11.46 (3.20)	10.44 (2.77)
	Missing information	110	43	67
BRIEF – GEC		87.42 (18.20)	85.13 (16.64)	89.34 (19.23)
	Missing information	109	44	65

Note. EF task metrics: Spin the Pots: number of successful finds (max = 16 at 2.5 years; max = 20 at 5 years), higher scores indicate better working memory. Snack Delay: sum score across six trials (0–36), higher scores indicate better inhibitory control. Delay of Gratification: number of "later" choices (0–9), higher scores indicate better inhibitory control. Arrows, Pigs, and Farmer tasks: proportion correct (0–1), where 0 indicates no correct items and 1 indicates all items correct. Overlap in microbial samples: 108.

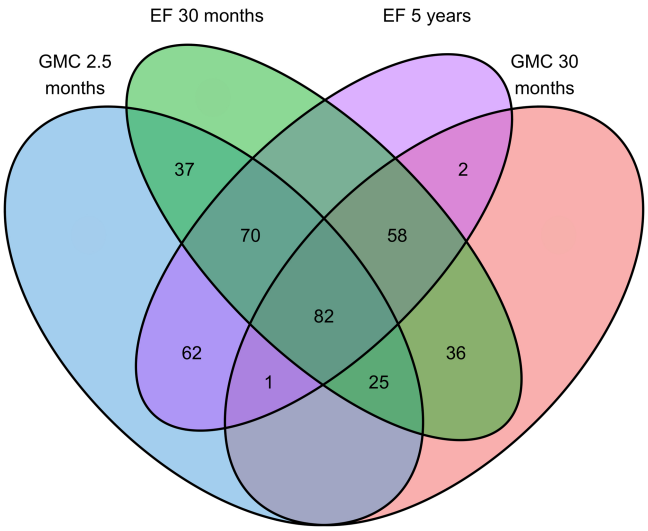


Figure 1. Venn diagram showing availability of stool samples and EF measurements across timepoints. Circles indicate children with stool samples at 2.5 months (blue) and 30 months (red), and EF measures at 30 months (green) and 5 years (purple). Overlaps show individuals with data at multiple timepoints, illustrating longitudinal versus cross-sectional coverage.

girls, the correlations were similar to the whole sample with the observed richness index at 2.5 months of age correlating negatively with the EF Touch Pigs task at 5 years for girls ($\tau = -0.23$, adj.$p = 0.03$) but for boys there were no statistically significant associations. The correlation coefficient was higher when tested for in girls, suggesting that the correlation in the whole sample is driven by girls. Linear regression models were built first for the whole sample and then only for girls based on the correlations. For the whole sample, a higher observed richness index at 2.5 months was associated with decreased performance in the EF Touch Pigs task at 5 years both in the unadjusted analysis (β= -0.03, $R^2 = 0.03$, $p < .01$) and when adjusted for covariates ($b = -0.40$, $R^2 = 0.08$, $p < .01$) (Figure 2). Results were similar for girls with higher observed richness index at 2.5 months associating with decreased performance in the EF Touch Pigs task at 5 years both in the unadjusted analysis ($\beta = -0.05$, $R^2 = 0.14$, $p < .001$) and when adjusted for covariates ($b = -.60$, $R^2 = 0.34$, $p < .001$) (Figure 2). For alpha diversity, including infant antibiotic exposure did not change the statistical significance of the results.

Community composition and executive functions

We used simple stratification by the five most dominant bacterial genera, which corresponded to *Bacteroides* ($n = 104$), *Bifidobacterium* ($n = 68$), *Clostridium* ($n = 77$), *Escherichia* ($n = 91$) and *Veillonella* ($n = 76$). The remaining samples ($n = 65$) were labeled as "Other" (Supplemental Figure 2). As for changes over time, the proportion of infants in the *Bacteroides* and *Bifidobacterium*-dominant groups decreased from 41% at 2.5 months to 28% at 30 months, while the *Clostridium*-dominant group increased from 10% to 23% over the same period. We observed differences in EF performance between the dominant genera groups (Table 2).

Regarding EF scores, when analyzed for the whole sample, at 2.5 months of age the *Bacteroides*-dominant group differed from the *Escherichia*-dominant group with the *Escherichia*-dominant group at 2.5 months of age performing better in the EF Touch Pigs task at 5 years of age (adj.$p = .03$) (Table 2 & Figure 3(a)). When analyzed separately for boys and girls we found that for girls, the *Clostridium*-dominant group at 2.5 months performed better in the EF Touch Pigs task at 5 years of age compared to the *Bacteroides*-dominant (adj.$p = .02$) and the *Bifidobacterium*-dominant (adj.$p = .02$) groups (Table 2 & Figure 3(b)). For boys, the *Bifidobacterium*-dominant group at age 30 months performed better at the Arrows task at 5 years of age compared to the *Escherichia*-dominant (adj.$p = .03$) and the *Veillonella*-dominant (adj.$p = .03$) groups (Table 2 & Figure 3(d)). Sensitivity analyses including infant antibiotic exposure showed largely consistent results. In the difference-in-differences analyses, two previously significant associations were slightly attenuated, and one new association reached statistical significance.

Differential abundances and executive functions

The associations between bacterial abundances at genus level and each EF variable are shown in Table 3 and illustrated in Figures 4 and 5. For the whole 2.5 month GMC sample, children's higher performance in the Spin the Pots task at 2.5 years was negatively associated with the genus *Escherichia,* and better performance in the EF Touch Pigs task at 5 years showed negative associations with

https://doi.org/10.1017/S0954579426101527 Published online by Cambridge University Press

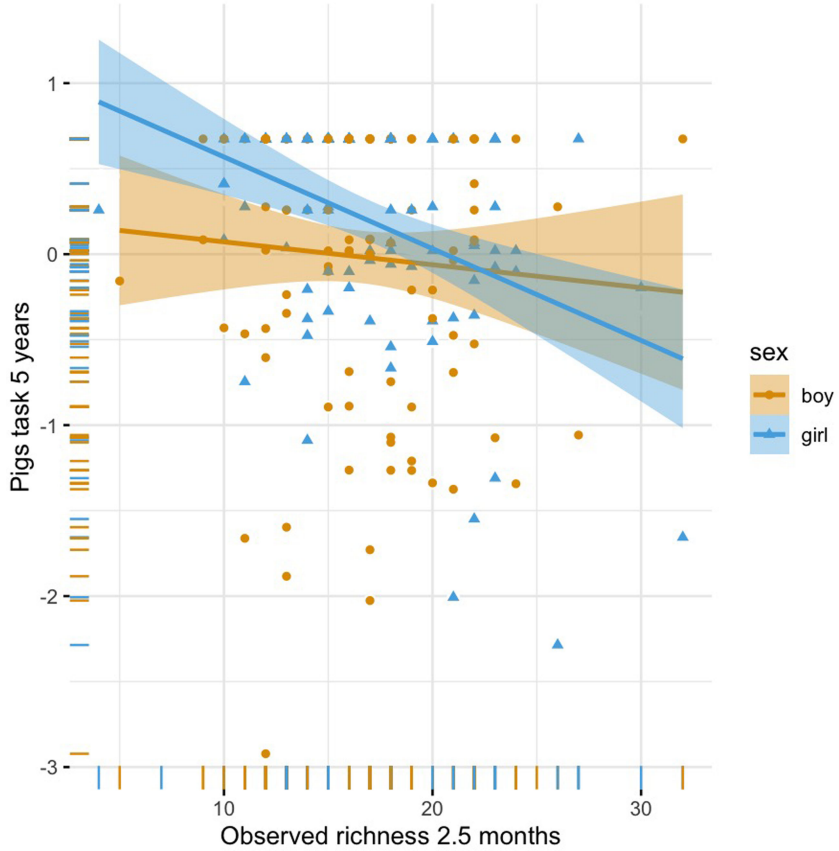


Figure 2. Alpha diversity (Observed richness) at 2.5 months and EF Touch Pigs task at 5 years of age. Regression lines displayed by sex.

the genus *Lactobacillus*. The Delay of Gratification task (5 years) showed a positive association with unidentified genus belonging to the family Enterococcaceae. When stratified for sex, for girls, the Delay of Gratification task (5 years) was positively associated with unidentified genus belonging to the family Enterococcaceae at 2.5 months. For boys, the Snack Delay task at 2.5 years was negatively associated with the genus *Barnesiella* ($p < 0.05$ for all).

When focusing on the 30-month GMC, the Spin the Pots task at 2.5 years showed positive associations with the genera *Escherichia* and *Ruminococcus*. Additionally, at 30 months, an unidentified genus belonging to the family Lachnospiraceae was negatively associated with the EF Touch Farmer task at 5 years. With regard to 30-month GMC, there were no associations when stratified by sex. When accounting for the effects of sample handling, the results remained highly similar to the primary analyses, with comparable effect directions, and all reported associations remaining statistically significant (see Supplementary Table 2).

Metabolites and executive functions

We found that microbial metabolites, including polar metabolites, bile acids and SCFAs, were associated with EF traits. Specifically, Snack Delay and Spin the Pots tasks at 2.5 years were negatively associated with the microbe-modified secondary bile acids 7-oxo-DCA and GCA at 2.5 months based on a nominal p-value < .05. Likewise, the Spin the Pots at 2.5 years was negatively correlated with 7-oxo-DCA at 30 months.

The Delay of Gratification task at 5 years was negatively associated with valeric acid at 2.5 months of age and aMCA at 30 months of age. The Spin the Pots task at 5 years of age was negatively associated with propionic and isovaleric acids at 2.5 months. The BRIEF-2 GEC negatively associated with propionic acid at 2.5 months. Similarly, the EF Touch Pigs task at 5 years showed a negative association with CA and aMCA, whereas the EF Touch Arrows task at 5 years exhibited a positive association with 7-oxo-DCA and CA.

When stratified by sex, these associations showed some differences. For instance, for boys only, 7-oxo-DCA at 30 months was positively correlated with the EF Touch Arrows task at 5 years and butyric acid at 30 months positively associated with the Spin the Pots task at 5 years. For girls, GUDCA at 2.5 months was positively associated with the Spin the Pots task at 5 years.

These associations between stool metabolites and each EF variable are detailed in Supplementary Tables 3, 4, 5 and 6 and Supplementary Figures 3 and 4. None of these metabolites remained significant after correcting for multiple comparisons using a false discovery rate (FDR) threshold of 0.05.

Discussion

In this study, our aim was to examine how GMC and fecal metabolites in infancy and toddlerhood associate with EFs in toddlerhood and preschool age in a general population-based sample of children. We found some associations between alpha diversity and EFs and differences in EFs between microbial groups based on the most dominant genus as well as differential abundances of bacterial genera related to EFs. However, our findings differed somewhat from the limited and methodologically varied previous research, while comparable study populations and designs have been scarce.

First, regarding alpha diversity, we found that higher observed richness at 2.5 months of age was associated with poorer

https://doi.org/10.1017/S0954579426101527 Published online by Cambridge University Press

Table 2. Results of Dunn's post-hoc test showing differences in EF scores across groups defined by dominant bacterial genera

Timepoint		EF task	Most dominant genus group	Most dominant genus group	*Z*-statistic	adj.*p*
2.5	all	EF Touch Pigs task 5 years	Bacteroides ($n = 48$)	Escherichia ($n = 41$)	3.08	.03*
	girls	EF Touch Pigs task 5 years	Bacteroides ($n = 23$)	Clostridium ($n = 8$)	3.24	.02*
			Bifidobacterium ($n = 14$)	Clostridium ($n = 8$)	2.94	.02*
30	boys	EF touch arrows task 5 years	Bifidobacterium ($n = 9$)	Escherichia ($n = 13$)	−2.96	.03*
			Bifidobacterium ($n = 9$)	Veillonella ($n = 13$)	−2.87	.03*

P-values have been corrected for multiple testing. * $p < .05$.

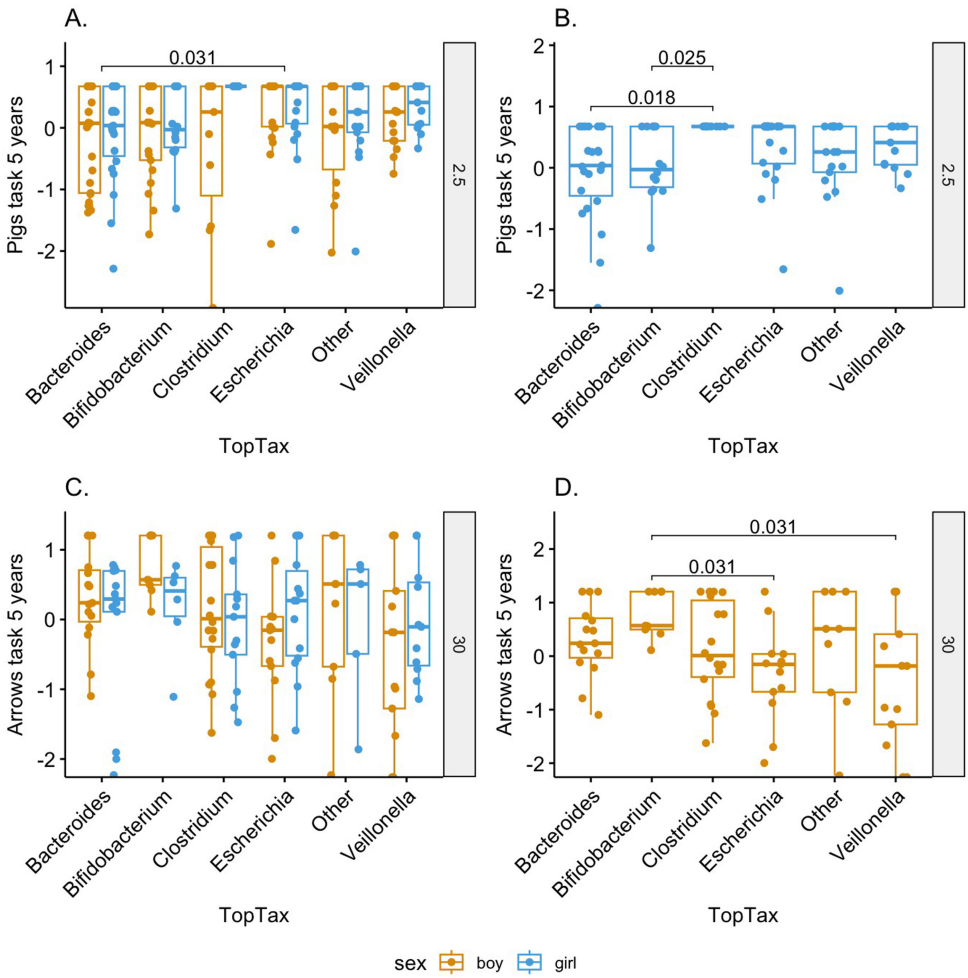


Figure 3. Dunn's test for community composition (grouping by most dominant bacterial genera) and EF. EF Touch inhibition task pigs at 5 years in the whole sample (a), EF Touch inhibition task pigs at 5 years for girls (b), EF touch inhibition task arrows at 5 years for the whole sample (c), EF touch inhibition task arrows at 5 years for boys (d). Statistically significant differences displayed by lines.

performance in the inhibition measuring EF Touch Pigs task at 5 years of age. Although high diversity is generally considered health-beneficial in adults, lower alpha diversity in infancy has been shown to be age appropriate (Bäckhed et al., 2015; Stewart et al., 2018). Breastfeeding affects infant GMC in a way that it is dominated by *Bifidobacterium* while richness and diversity of bacterial species increase over time with the introduction of solid foods (Stewart et al., 2018). This suggests that lower microbial diversity can be seen as typical for early infancy, while higher diversity at this age point might reflect an atypical pattern of GMC development. Contrary to our finding, Willemsen et al. (2023) found that higher alpha diversity (observed richness) measured at 2 weeks of age was linked to better performance in the inhibitory control task and parental reports of EF at three years. A key difference between these studies is the timing of the assessments: Willemsen et al., measured diversity very early, at 2 weeks of age using Chao1, whereas in our study diversity was assessed at 2.5 months using observed richness. Moreover, gut microbiota diversity is naturally low at birth and increases rapidly over the first months of life (Bäckhed et al., 2015), meaning that 2-week and 2.5-month measurements reflect different stages of microbial maturation. At 2.5 months, higher diversity may indicate

https://doi.org/10.1017/S0954579426101527 Published online by Cambridge University Press

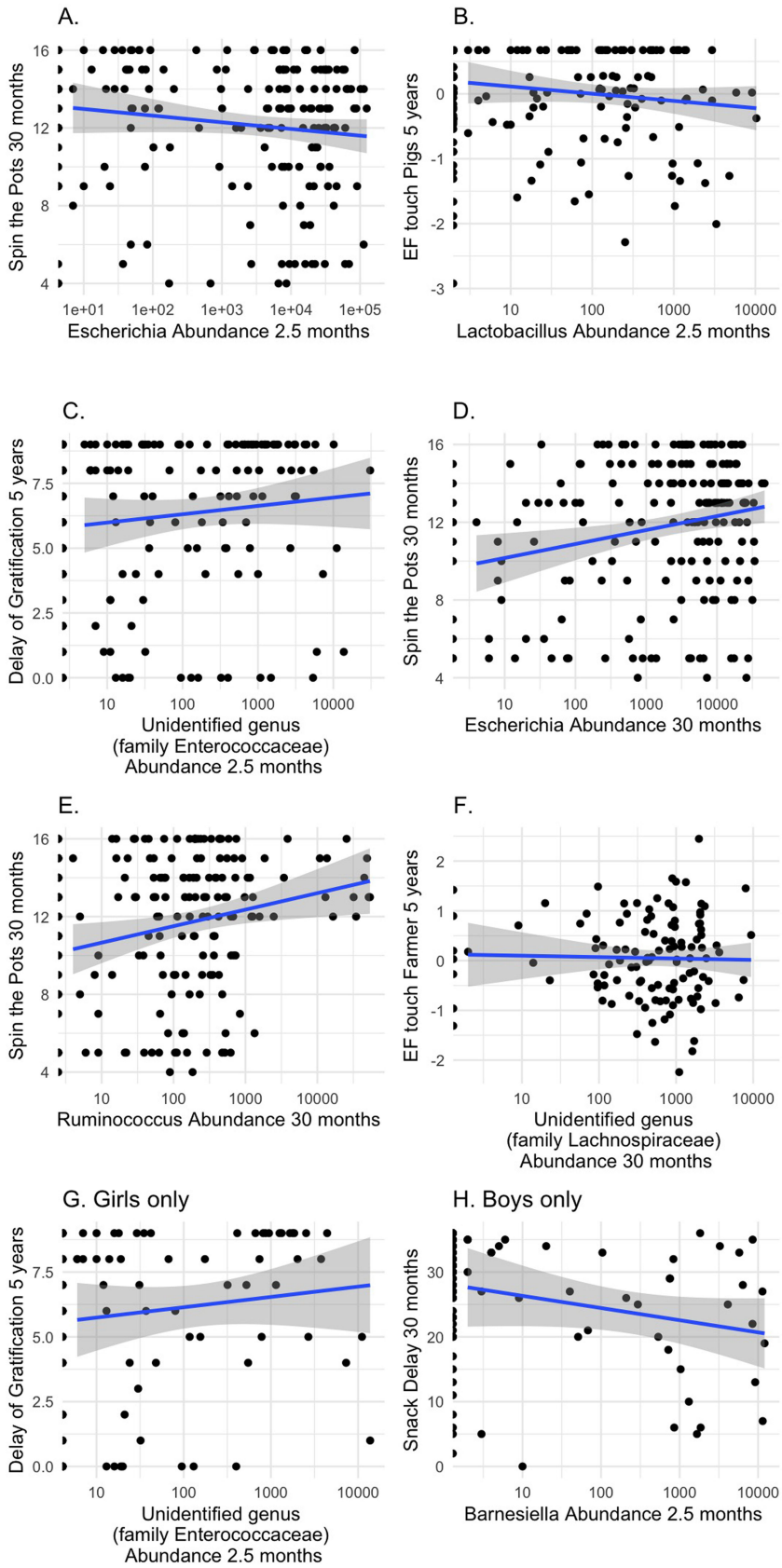


Figure 4. Scatter plots showing statistically significant associations between log-transformed abundances of genera and EF.

https://doi.org/10.1017/S0954579426101527 Published online by Cambridge University Press

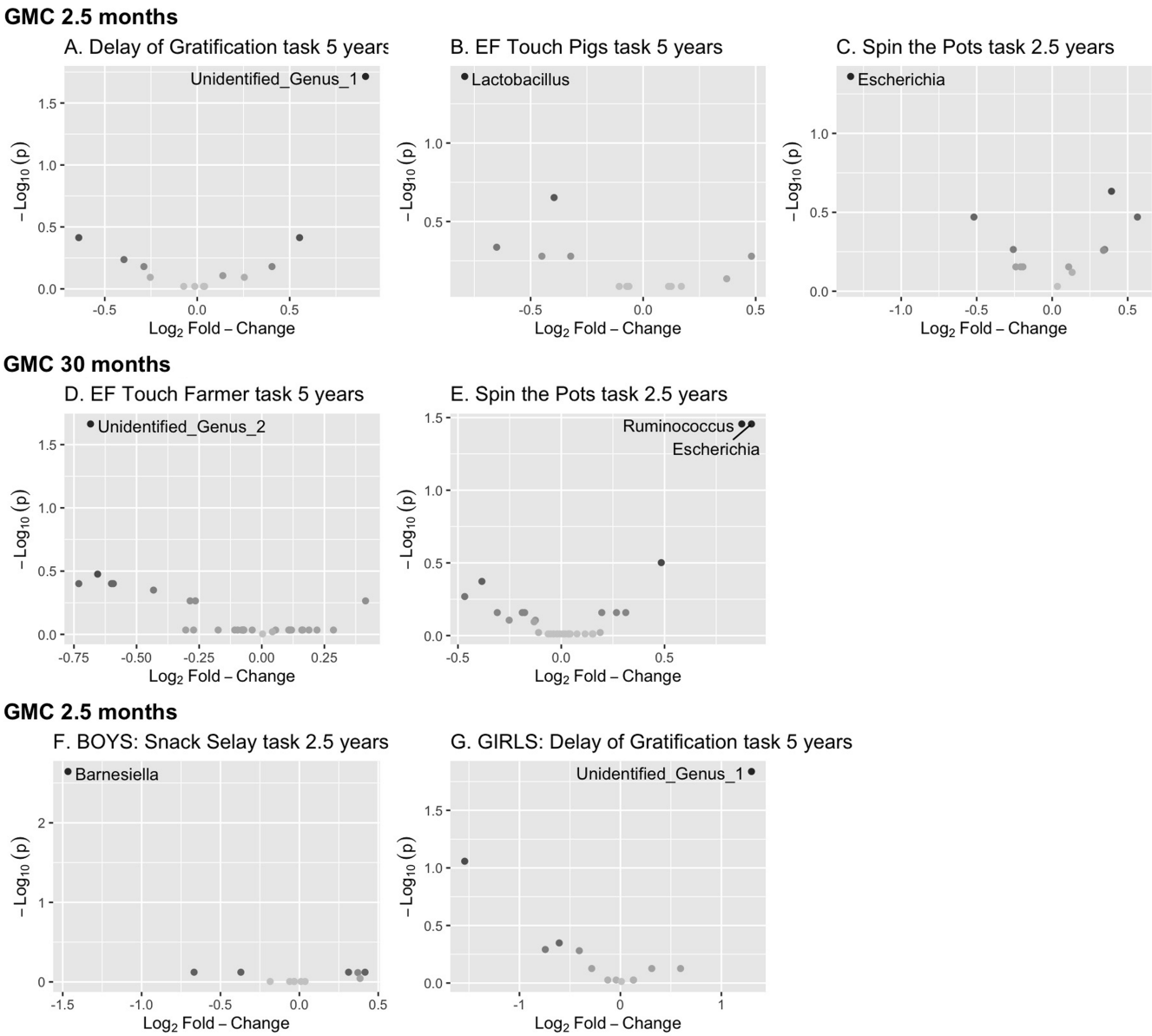


Figure 5. Volcano plots showing associations between genera and EF tasks. The statistically significant genera are labeled.

deviations from the typical *Bifidobacterium*-dominated trajectory, whereas at 2 weeks variation in diversity might reflect the initial colonization process. These differences in measurement and developmental timing may partly explain the differences in findings.

Furthermore, the rapid and nonlinear development of EF in childhood may contribute to discrepancies between studies examining associations with gut microbiota. Longitudinal models show that a large portion of EF improvement occurs between the ages of 3 and 5, with children typically showing an improvement of even more than one standard deviation over a 12-month period (Willoughby et al., 2014). As a result, EF measures at 3 versus 5 years may capture different developmental processes, with earlier assessments reflecting a less stable construct. As EF matures, individual differences become more stable and the overall variability decreases, making later assessments more reliable indicators of meaningful differences in EF performance. This developmental shift in EF performance can partly explain why associations with early GMC may differ across studies, as can been seen in our and Willemsen et al. (2023) studies. Taken together, these findings highlight the importance of considering the developmental timing of both microbiota sampling and EF assessment when interpreting associations between them.

As for community composition, for the whole sample, the *Escherichia*-dominant group at 2.5 months performed better in an inhibition task (the EF Touch Pigs) at 5 years of age compared to the *Bacteroides*-dominant group. Interestingly, in a study by Tamana et al. (2021), the *Bacteroides*-dominant GMC at one year of age was found to be beneficial for neurodevelopmental outcomes at two years of age in a previous study. Both *Bacteroides (Phoceaicola)* and *Escherichia* are functionally diverse genera, and the age points and outcome measures differ to the extent that comparisons are difficult to make, and further studies are required to make more robust conclusions.

When girls and boys were analyzed separately, for girls only the *Clostridium*-dominant group at 2.5 months of age performed

https://doi.org/10.1017/S0954579426101527 Published online by Cambridge University Press

Table 3. Differential abundances and EF adjusted

	Timepoint	EF task	Associated genus	df	Baseline mean abundance	Log2 fold −change	Adjusted *p*-value
All	2.5 months	Spin the pots 2.5 years	Escherichia	160	35,828.92	−1.33	.04
		EF touch pigs 5 years	Lactobacillus	149	1529.92	−0.79	.04
		Delay of gratification 5 years	Enterococcaceae_unidentified_genus_1	158	8,327.91	0.91	.02
	30 months	Spin the pots 2.5 years	Escherichia Ruminococcus	145 145	114,301.47 59.39	0.92 0.87	.04 .04
		EF touch Farmer 5 year	Lachnospiraceae_unidentified_genus_2	85	1878.11	−0.68	.02
Girls	2.5 months	Delay of gratification 5 years	Enterococcaceae_unidentified_genus_1	65	3,123.02	1.30	.02
Boys	2.5 months	Snack Delay 2.5 years	Barnesiella	77	1,286.00	−1.47	>.01

better in an inhibition task (the EF Touch Pigs) at 5 years of age compared to the *Bacteroides*- and *Bifidobacterium*-dominant groups. *Bacteroides* and *Bifidobacterium* in early life have been associated with beneficial outcomes such as better neurodevelopmental outcomes, adaptive skills and normal development scores (Naspolini et al., 2024). However, our finding should be interpreted with caution. The *Clostridium*-dominant group for girls was small and comprised solely of top scores, which limited overall variance. Furthermore, as noted above, EFs develop rapidly in early childhood, likely contributing to ceiling effects in this task, as some children quickly reached the maximum performance level for certain measures (Miller et al., 2023; Willoughby et al., 2014). Overall, the complexity of assessing EFs in young children may help explain why associations were observed in some EF measures but not across the entire battery. This underscores the importance of group sizes, developmental variability, and task selection when interpreting EF associations with microbiota composition.

For boys, the *Bifidobacterium*-dominant group at 2.5 years of age performed better in an inhibition task (the EF Touch Arrows) at 5 years of age compared to the *Escherichia*- and *Veillonella*-dominant groups. This observation is in line with previous research, since it has shown that *Bifidobacterium* is abundant and beneficial in both infant and toddlerhood gut (Sudo et al., 2004; Turroni et al., 2017). Overall, sex differences in these results could be partly explained by differences in EF task performance between girls and boys (Grissom & Reyes, 2019). While the between-sex differences in EF task performance are not as prominent as the within-sex differences, Grissom and Reyes (2019) suggest that differences in cognitive strategies rather than abilities may exist.

Moreover, accumulating evidence indicates that gut-brain interactions are often sex-specific. Animal studies have shown that GMC can influence neurodevelopment, stress response, and cognitive function differentially in males and females, potentially via sex hormones, immune system modulation, or divergent brain maturation trajectories (Clarke et al., 2013; Jašarević et al., 2016). A systematic review on rodent studies has shown that modulating maternal microbiota alters male offspring behavior (Hassib et al., 2023) indicating that males may be more vulnerable to the effects of these microbial alterations. This is in line with our finding on community composition and EF in boys. Importantly, early-life associations between GMC and behavioral development in early human studies have also been shown to differ by sex, as infant GMC has been linked with later behavioral development in a sex-dependent manner (Huovinen et al., 2023; Laue et al., 2022). Therefore, the associations observed between early GMC profiles and later EF performance may be influenced not only by sex-specific cognitive strategies but also by the seemingly sex-dependent characteristics of gut–brain communication, highlighting the importance of investigating the role of biological sex in future studies.

Additionally, differential abundance analysis showed a set of associations between different gut bacterial genera and EF scores. When analyzed for the whole sample, there was a time-dependent association between the genus *Escherichia* and working memory: a decrease at 2.5 months but an increase at 2.5 years of age was associated with an increase in the Spin the Pots task at 2.5 years of age indicating that there is no consistent positive association between the genus *Escherichia* and EF performance. The shift from breastfeeding to solid foods that usually occurs after 4-6 months leads to alterations in the composition of the bacterial community (Aatsinki et al., 2025). The cross-feeding with other bacteria and different dietary sources may modulate the metabolism of species in the *Escherichia* genus, potentially underlying the difference in direction (Seki et al., 2025).

In the whole sample, a higher abundance of the genus *Lactobacillus* at 2.5 months was negatively associated with performance in the inhibition task (EF Touch Pigs) at 5 years of age. This finding contrasts with previous research, as *Lactobacillus* is generally considered a beneficial component of the infant gut microbiota due to its association with breastfeeding (Zhang et al., 2020), and probiotic treatments have been linked with improved neurodevelopment in animal models. However, due to the resolution of our data, we cannot determine whether the *Lactobacillus* strains represent resident microbes or originate from probiotic formulations, and we cannot rule out potential confounding effects of probiotic treatments.

When analyzed for boys and girls separately, for boys only, an increase in the Snack Delay task at 30 months of age was associated with a decrease in the genus *Barnesiella* at 2.5 months. Willemsen et al. (2023) found that *Barnesiella* at 3 years of age had a positive association with internalizing and externalizing behaviors at the same age. These findings suggest that *Barnesiella* may play a role in early behavioral development. However, it is important to note the differential age points again, as well as the distinction between EF and externalizing behavior. Although both are related to self-regulation, they represent different behavioral constructs. Interestingly, we found no statistically significant associations between alpha diversity, community composition, or differential abundances and BRIEF-2 questionnaire. BRIEF-2 measures parent-reported EF in everyday life and captures different aspects

https://doi.org/10.1017/S0954579426101527 Published online by Cambridge University Press

of EF compared to performance-based tasks. The lack of associations with this measure may be explained by the complex nature of EF, which is a heterogeneous phenotype shaped by various biological and environmental influences.

Lastly, as mentioned, we found no statistically significant associations between metabolites - spanning polar metabolites, bile acids, short chain fatty acids - and EF. However, we observed that generally EF tasks negatively correlated with secondary bile acids and short-chain fatty acids at the nominal significance level. Gut microbial metabolites are important mediators of the effects on host neurodevelopment (Ahmed et al., 2022). Our results highlight the potential role of bile acids in self-regulatory capacities such as EF functioning. Bile acids can bind to Farnesoid X receptor and G-protein coupled bile acid receptors that are expressed in numerous tissues, including the brain (Sabahat et al., 2024). Moreover, BAs can influence neurotransmitter production as well as neuroinflammation. Of note, even though our metabolite panel was comprehensive, it did not include potential neurotransmitters, such as GABA, tryptophan or their precursors. Future studies should consider assessing how gut microbiota functional potential associates with EFs, and study targeted metabolomics, where the mechanistic relevance for EF might be clearer.

The EF tasks in this study included Spin the Pots and Snack Delay tasks at 2.5 years of age and Spin the Pots, Delay of Gratification, and EF Touch tasks as well as the BRIEF-2 questionnaire at 5 years of age. Different EF tasks did not correlate with each other in this study, which reflects the complexity of this phenotype. EF skills can be defined as dynamic and context dependent. For instance, previous research has shown that rank-order stability of EFs in childhood is moderate at best (Willoughby et al., 2014), which can be seen in our study as well.

The strengths of this study include extensive EF measurements, including both task- and questionnaire data as well as multiple time points regarding GMC and EF. One limitation of this study is the 16S rRNA sequencing, which does not allow for detailed information on bacterial species and functional potential. However, we added metabolite data to decrease the impact of that limitation. In our knowledge, this is one of the first few studies to investigate the potential associations between GMC and stool metabolites and later EF.

Taken together, these results add to the novel research literature on the associations between GMC and EF in early childhood when neurodevelopment, including the gut-brain axis, is rapid and malleable to environmental influence. Our results indicate that both infant and toddler gut microbiota features, such as the genus *Escherichia*, relate to inhibition and working memory. However, given the current heterogeneity of findings as well as the methodology applied in this field, drawing clear conclusions is challenging. One potential explanation for the disparate results is the dynamic and evolving nature of both GMC and EF, which are not fully stable during early development. Although it is suggested that GMC is one of the contributing factors to neurodevelopment, its impact is still largely unclear. Given these challenges, there is a need for more research on this topic. Further studies could help to confirm the existing findings, including the findings of the current study, and identify the mechanisms underlying the associations between GMC and EF in childhood.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0954579426101527.

Data availability statement. Due to Finnish national legislation on personal data protection and ethics consent, the original data cannot be made available online, but data can potentially be shared with Research Agreement as part of research collaboration. Requests for collaboration can be sent to the Board of the FinnBrain Birth Cohort Study; please contact Linnea Karlsson (linnea.karlsson@utu.fi).

The R scripts for data analyses can be found in Zenodo doi: 10.5281/zenodo.15198449.

The metabolomics datasets generated in this study is available at the NIH Common Fund's National Metabolomics Data Repository (NMDR) website, the Metabolomics Workbench (https://www.metabolomicsworkbench.org) where it has been assigned with DataTrackID 4779. The data can be accessed directly at Metabolomics Workbench: http://doi.org/10.21228/M8BB26. The 16s rRNA gene sequencing data with key metadata for this study are available in the European Nucleotide Archive with accession code PRJEB73815. The R scripts for data analyses can be found in Zenodo doi:10.5281/zenodo.11222841.

The metabolomics datasets generated in this study is available at the NIH Common Fund's National Metabolomics Data Repository (NMDR) website, the Metabolomics Workbench (https://www.metabolomicsworkbench.org) where it has been assigned with DataTrackID 4779. The data can be accessed directly at Metabolomics Workbench: http://doi.org/10.21228/M8BB26. The 16s rRNA gene sequencing data with key metadata for this study are available in the European Nucleotide Archive with accession code PRJEB73815. The R scripts for data analyses can be found in Zenodo doi:10.5281/zenodo.11222841.

Acknowledgements. We acknowledge the contributions of all participating families, as well as the FinnBrain staff and assisting personnel. This work was supported by the "Inflammation in human early life: targeting impacts on life-course health" (INITIALISE) consortium funded by the Horizon Europe Program of the European Union under Grant Agreement 101094099. Artificial intelligence (AI) tools (e.g., ChatGPT by OpenAI) were used to assist in language editing and improving clarity of phrasing in the revision of this manuscript. The authors take full responsibility for the content of the manuscript.

Funding statement. This research was funded by Finnish State Grants for Clinical Research (ERVA) Research Council of Finland (308252, 308589, 308176) and Waterloo Foundation.

VH was funded by Emil Aaltonen foundation, OLVI Foundation and Signe & Ane Gyllenberg foundation. LK was funded by the Research Council of Finland (grant number 308176 and 325292), Yrjö Jahnsson Foundation (6847, 6976), Signe and Ane Gyllenberg Foundation, Finnish State Grants for Clinical Research (P3654), Jalmari and Rauha Ahokas Foundation, and Waterloo Foundation (2110-3601). AKA was supported by Yrjö Jahnsson Foundation, Psychiatry Research Foundation, Emil Aaltonen Foundation, Brain Foundation, Instrumentarium Science Foundation, Signe and Ane Gyllenberg Foundation, Duodecim Finnish Medical Society, Juho Vainio Foundation and Research Council of Finland (grant number 347640). LL was supported by Research Council of Finland (grant number 330887). EM was supported by the government research grant awarded to Turku University Hospital. AD has been funded by the Waterloo foundation and Research Council of Finland (347924). "Inflammation in human early life: targeting impacts on life-course health" (INITIALISE) consortium funded by the Horizon Europe Program of the European Union under Grant Agreement 101094099 (to MO HK AD).

Competing interests. All authors declare that they have no conflicts of interest.

Pre-registration statement. This study was not pre-registered. The analyses included exploratory components aimed at examining associations between early gut microbiota composition and later executive functioning.

References

Aatsinki, A.-K., Lahti, L., Uusitupa, H.-M., Munukka, E., Keskitalo, A., Nolvi, S., O'Mahony, S., Pietilä, S., Elo, L. L., Eerola, E., Karlsson, H., & Karlsson, L. (2019). Gut microbiota composition is associated with temperament traits in infants. *Brain, Behavior, and Immunity*, *80*, 849–858. https://doi.org/10.1016/j.bbi.2019.05.035

Aatsinki, A.-K., Lamichhane, S., Isokääntä, H., Sen, P., Kråkström, M., Alves, M. A., Keskitalo, A., Munukka, E., Karlsson, H., Perasto, L., Lukkarinen, M., Oresic, M., Kailanto, H.-M., Karlsson, L., Lahti, L., &

Dickens, A. M. (2025). Dynamics of gut metabolome and microbiome maturation during early life. *iScience*, *28*(11), 113596. https://doi.org/10.1016/j.isci.2025.113596

Ahmed, H., Leyrolle, Q., Koistinen, V., Kärkkäinen, O., Layé, S., Delzenne, N., & Hanhineva, K. (2022). Microbiota-derived metabolites as drivers of gut–brain communication. *Gut Microbes*, *14*(1), 2102878. https://doi.org/10.1080/19490976.2022.2102878

Bäckhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., Li, Y., Xia, Y., Xie, H., Zhong, H., Khan, M. T., Zhang, J., Li, J., Xiao, L., Al-Aama, J., Zhang, D., Lee, Y. S., Kotowska, D., Colding, C., . . . , Wang, J. (2015). Dynamics and stabilization of the human gut microbiome during the first year of life. *Cell Host & Microbe*, *17*(5), 690–703. https://doi.org/10.1016/j.chom.2015.04.004

Beck, D. M., Schaefer, C., Pang, K., & Carlson, S. M. (2011). Executive function in preschool children: Test–retest reliability. *Journal of Cognition and Development*, *12*(2), 169–193. https://doi.org/10.1080/15248372.2011.563485

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Borman, T., Ernst, F., Shetty, S., & Lahti, L. (2025). mia: Microbiome analysis. R package version 1.18.0. https://bioconductor.org/packages/mia

Bridgett, D. J., Burt, N. M., Edwards, E. S., & Deater-Deckard, K. (2015). Intergenerational transmission of self-regulation: A multidisciplinary review and integrative conceptual framework. *Psychological Bulletin*, *141*(3), 602–654. https://doi.org/10.1037/a0038662

Bridgett, D. J., Oddi, K. B., Laake, L. M., Murdock, K. W., & Bachmann, M. N. (2013). Integrating and differentiating aspects of self-regulation: Effortful control, executive functioning, and links to negative affectivity. *Emotion*, *13*(1), 47–63. https://doi.org/10.1037/a0029536

Broomell, A. P. R., & Bell, M. A. (2022). Longitudinal development of executive function from infancy to late childhood. *Cognitive Development*, *63*, 101229. https://doi.org/10.1016/j.cogdev.2022.101229

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P.2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 7. https://doi.org/10.1038/nmeth.3869

Carlson, A. L., Xia, K., Azcarate-Peril, M. A., Goldman, B. D., Ahn, M., Styner, M. A., Thompson, A. L., Geng, X., Gilmore, J. H., & Knickmeyer, R. C. (2018). Infant gut microbiome associated with cognitive development. *Biological Psychiatry, Early Life Environments and Later Vulnerabilities*, *83*(2), 148–159. https://doi.org/10.1016/j.biopsych.2017.06.021

Clarke, G., Grenham, S., Scully, P., Fitzgerald, P., Moloney, R. D., Shanahan, F., Dinan, T. G., & Cryan, J. F. (2013). The microbiome-gut-brain axis during early life regulates the hippocampal serotonergic system in a sex-dependent manner. *Molecular Psychiatry*, *18*(6), 6. https://doi.org/10.1038/mp.2012.77

Cryan, J. F., O'Riordan, K. J., Cowan, C. S. M., Sandhu, K. V., Bastiaanssen, T. F. S., Boehme, M., Codagnone, M. G., Cussotto, S., Fulling, C., Golubeva, A. V., Guzzetta, K. E., Jaggar, M., Long-Smith, C. M., Lyte, J. M., Martin, J. A., Molinero-Perez, A., Moloney, G., Morelli, E., Morillas, E., . . . , Dinan, T. G. (2019). The microbiota-gut-brain axis. *Physiological Reviews*, *99*(4), 1877–2013. https://doi.org/10.1152/physrev.00018.2018

Dalile, B., Van Oudenhove, L., Vervliet, B., & Verbeke, K. (2019). The role of short-chain fatty acids in microbiota–gut–brain communication. *Nature Reviews Gastroenterology & Hepatology*, *16*(8), 8. https://doi.org/10.1038/s41575-019-0157-3

Derrien, M., Alvarez, A.-S., & De Vos, W. M. (2019). The gut microbiota in the first decade of life. *Trends in Microbiology*, *27*(12), 997–1010. https://doi.org/10.1016/j.tim.2019.08.001

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Eckermann, H. A., Ou, Y., Lahti, L., & de Weerth, C. (2022). Can gut microbiota throughout the first 10 years of life predict executive functioning in childhood? *Developmental Psychobiology*, *64*(3), e22226. https://doi.org/10.1002/dev.22226

Fujihara, H., Matsunaga, M., Ueda, E., Kajiwara, T., Takeda, A. K., Watanabe, S., Baba, K., Hagihara, K., & Myowa, M. (2023). Altered gut microbiota composition is associated with difficulty in explicit emotion regulation in young children. *Microorganisms*, *11*(9), 2245. https://doi.org/10.3390/microorganisms11092245

Gilmore, J. H., Knickmeyer, R. C., & Gao, W. (2018). Imaging structural and functional brain development in early childhood. *Nature Reviews Neuroscience*, *19*(3), 123–137. https://doi.org/10.1038/nrn.2018.1

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2015). *Behavior Rating Inventory of Executive Function* (2nd ed.) (BRIEF®2). PAR Inc.

Grissom, N. M., & Reyes, T. M. (2019). Let's call the whole thing off: Evaluating gender and sex differences in executive function. *Neuropsychopharmacology*, *44*(1), 86–96. https://doi.org/10.1038/s41386-018-0179-5

Gronier, B., Savignac, H. M., Di Miceli, M., Idriss, S. M., Tzortzis, G., Anthony, D., & Burnet, P. W. J. (2018). Increased cortical neuronal responses to NMDA and improved attentional set-shifting performance in rats following prebiotic (B-GOS®) ingestion. *European Neuropsychopharmacology*, *28*(1), 211–224. https://doi.org/10.1016/j.euroneuro.2017.11.001

Hassib, L., de Oliveira, C. L., Rouvier, G. A., Kanashiro, A., Guimarães, F. S., & Ferreira, F. R. (2023). Maternal microbiome disturbance induces deficits in the offspring's behaviors: A systematic review and meta-analysis. *Gut Microbes*, *15*(1), 2226282. https://doi.org/10.1080/19490976.2023.2226282

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, R., Soneson, C., Ernst, F. G. M., Rue-Albrecht, K. C., Yu, G., Hicks, S. C., & Robinson, M. D. (2021). *TreeSummarizedExperiment: A S4 class for data with hierarchical structure* (9: 1246). F1000Research. https://doi.org/10.12688/f1000research.26669.2

Hughes, C., & Ensor, R. (2005). Executive function and theory of mind in 2 year olds: A family affair? *Developmental Neuropsychology*, *28*(2), 645–668. https://doi.org/10.1207/s15326942dn2802_5

Huovinen, V., Aatsinki, A.-K., Kataja, E.-L., Munukka, E., Keskitalo, A., Lamichhane, S., Raunioniemi, P., Bridgett, D. J., Lahti, L., O'Mahony, S. M., Dickens, A., Korja, R., Karlsson, H., Nolvi, S., & Karlsson, L. (2023). Infant gut microbiota and negative and fear reactivity. *Development and Psychopathology*, *36*(4), 2016–2031. https://doi.org/10.1017/S0954579423001396

Jašarević, E., Morrison, K. E., & Bale, T. L. (2016). Sex differences in the gut microbiome–brain axis across the lifespan. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1688), 20150122. https://doi.org/10.1098/rstb.2015.0122

Karlsson, L., Tolvanen, M., Scheinin, N. M., Uusitupa, H.-M., Korja, R., Ekholm, E., Tuulari, J. J., Pajulo, M., Huotilainen, M., Paunio, T., Karlsson, H., & FinnBrain Birth Cohort Study Group. (2018). Cohort profile: The FinnBrain birth cohort study (FinnBrain). *International Journal of Epidemiology*, *47*(1), 15–16j. https://doi.org/10.1093/ije/dyx173

Keskitalo, A., Aatsinki, A.-K., Kortesluoma, S., Pelto, J., Korhonen, L., Lahti, L., Lukkarinen, M., Munukka, E., Karlsson, H., & Karlsson, L. (2021). Gut microbiota diversity but not composition is related to saliva cortisol stress response at the age of 2.5 months. *Stress-the International Journal on the Biology of Stress*, *24*(5), 551–560. https://doi.org/10.1080/10253890.2021.1895110

Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology*, *36*, 220–232. https://doi.org/10.1037/0012-1649.36.2.220

Laue, H. E., Karagas, M. R., Coker, M. O., Bellinger, D. C., Baker, E. R., Korrick, S. A., & Madan, J. C. (2022). Sex-specific relationships of the infant microbiome and early-childhood behavioral outcomes. *Pediatric Research*, *92*(2), 580–591. https://doi.org/10.1038/s41390-021-01785-z

Miller, S. E., Galvagno, L. G., & Elgier, Á. (2023). Universality and context-specificity in early executive function development. *Infant Behavior and Development*, *71*, 101841. https://doi.org/10.1016/j.infbeh.2023.101841

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Monk, C., Lugo-Candelas, C., & Trumpff, C. (2019). Prenatal developmental origins of future psychopathology: Mechanisms and pathways. *Annual*

https://doi.org/10.1017/S0954579426101527 Published online by Cambridge University Press

Review of Clinical Psychology, *15*(1), 317–344. https://doi.org/10.1146/annurev-clinpsy-050718-095539

Naspolini, N. F., Schüroff, P. A., Figueiredo, M. J., Sbardellotto, G. E., Ferreira, F. R., Fatori, D., Polanczyk, G. V., Campos, A. C., & Taddei, C. R. (2024). The gut microbiome in the first one Thousand days of neurodevelopment: a systematic review from the microbiome Perspective.. *Microorganisms*, *12*(3), 424. https://doi.org/10.3390/microorganisms12030424

Nearing, J. T., Douglas, G. M., Hayes, M. G., MacDonald, J., Desai, D. K., Allward, N., Jones, C. M. A., Wright, R. J., Dhanani, A. S., Comeau, A. M., & Langille, M. G. I. (2022). Microbiome differential abundance methods produce different results across 38 datasets. *Nature Communications*, *13*, 342. https://doi.org/10.1038/s41467-022-28034-z

Nelson, T. D., Kidwell, K. M., Nelson, J. M., Tomaso, C. C., Hankey, M., & Espy, K. A. (2018). Preschool executive control and internalizing symptoms in elementary school. *Journal of Abnormal Child Psychology*, *46*(7), 1509–1520. https://doi.org/10.1007/s10802-017-0395-1

Posner, M. I., Rothbart, M. K., Sheese, B. E., & Voelker, P. (2014). Developing attention: Behavioral and brain mechanisms. *Advances in Neuroscience*, *2014*(1), 405094. https://doi.org/10.1155/2014/405094

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*(Database issue), D590–D596. https://doi.org/10.1093/nar/gks1219

Ramos, M., Schiffer, L., Re, A., Azhar, R., Basunia, A., Rodriguez, C., Chan, T., Chapman, P., Davis, S. R., Gomez-Cabrero, D., Culhane, A. C., Haibe-Kains, B., Hansen, K. D., Kodali, H., Louis, M. S., Mer, A. S., Riester, M., Morgan, M., Carey, V., & Waldron, L. (2017). Software for the integration of multi-omics experiments in Bioconductor. *Cancer Research*, *77*(21), e39–e42. https://doi.org/10.1158/0008-5472.CAN-17-0344

Rintala, A., Riikonen, I., Toivonen, A., Pietilä, S., Munukka, E., Pursiheimo, J.-P., Elo, L. L., Arikoski, P., Luopajärvi, K., Schwab, U., Uusitupa, M., Heinonen, S., Savilahti, E., Eerola, E., & Ilonen, J. (2018). Early fecal microbiota composition in children who later develop celiac disease and associated autoimmunity. *Scandinavian Journal of Gastroenterology*, *53*(4), 403–409. https://doi.org/10.1080/00365521.2018.1444788

Sabahat, S. E., Saqib, M., Talib, M., Shaikh, T. G., Khan, T., & Kailash, S. J. (2024). Bile acid modulation by gut microbiota: A bridge to understanding cognitive health. *Annals of Medicine & Surgery*, *86*(9), 5410–5415. https://doi.org/10.1097/MS9.0000000000002433

Salosensaari, A., Laitinen, V., Havulinna, A. S., Meric, G., Cheng, S., Perola, M., Valsta, L., Alfthan, G., Inouye, M., Watrous, J. D., Long, T., Salido, R. A., Sanders, K., Brennan, C., Humphrey, G. C., Sanders, J. G., Jain, M., Jousilahti, P., Salomaa, V., . . . **, Niiranen, T.** (2021). Taxonomic signatures of cause-specific mortality risk in human gut microbiome. *Nature Communications*, *12*(1), 2671. https://doi.org/10.1038/s41467-021-22962-y

Seki, D., Pollak, S., Kujawska, M., Kiu, R., Acuna-Gonzales, A., Crouch, L., Bakshani, C., Chivers, P., Mommers, M., Van Best, N., Penders, J., & Hall, L. J. (2025). Mutualistic interactions between Escherichia coli and Bifidobacterium bifidum enable degradation of human milk oligosaccharides in healthy infants. https://doi.org/10.1101/2025.03.24.645027

Sordillo, J. E., Korrick, S., Laranjo, N., Carey, V., Weinstock, G. M., Gold, D. R., O'Connor, G., Sandel, M., Bacharier, L. B., Beigelman, A., Zeiger, R., Litonjua, A. A., & Weiss, S. T. (2019). Association of the infant gut microbiome with early childhood neurodevelopmental outcomes: An ancillary study to the VDAART randomized clinical trial. *JAMA Network Open*, *2*(3), e190905. https://doi.org/10.1001/jamanetworkopen.2019.0905

Spinrad, T. L., Eisenberg, N., & Gaertner, B. M. (2007). Measures of effortful regulation for young children. *Infant Mental Health Journal*, *28*(6), 606–626. https://doi.org/10.1002/imhj.20156

Stewart, C. J., Ajami, N. J., O'Brien, J. L., Hutchinson, D. S., Smith, D. P., Wong, M. C., Ross, M. C., Lloyd, R. E., Doddapaneni, H., Metcalf, G. A., Muzny, D., Gibbs, R. A., Vatanen, T., Huttenhower, C., Xavier, R. J., Rewers, M., Hagopian, W., Toppari, J., Ziegler, A.-G., . . . **, Petrosino, J. F.** (2018). Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature*, *562*(7728), 583–588. https://doi.org/10.1038/s41586-018-0617-x

Sudo, N., Chida, Y., Aiba, Y., Sonoda, J., Oyama, N., Yu, X.-N., Kubo, C., & Koga, Y. (2004). Postnatal microbial colonization programs the hypothalamic-pituitary-adrenal system for stress response in mice: Commensal microbiota and stress response. *Journal of Physiology*, *558*(1), 263–275. https://doi.org/10.1113/jphysiol.2004.063388

Tamana, S. K., Tun, H. M., Konya, T., Chari, R. S., Field, C. J., Guttman, D. S., Becker, A. B., Moraes, T. J., Turvey, S. E., Subbarao, P., Sears, M. R., Pei, J., Scott, J. A., Mandhane, P. J., & Kozyrskyj, A. L. (2021). Bacteroides-dominant gut microbiome of late infancy is associated with enhanced neurodevelopment. *Gut Microbes*, *13*(1), 1–17. https://doi.org/10.1080/19490976.2021.1930875

Trimigno, A., Khakimov, B., Mejia, J. L. C., Mikkelsen, M. S., Kristensen, M., Jespersen, B. M., & Engelsen, S. B. (2017). Identification of weak and gender specific effects in a short 3 weeks intervention study using barley and oat mixed linkage β-glucan dietary supplements: A human fecal metabolome study by GC-MS. *Metabolomics*, *13*(10), 108. https://doi.org/10.1007/s11306-017-1247-2

Turroni, F., Milani, C., Duranti, S., Ferrario, C., Lugli, G. A., Mancabelli, L., van Sinderen, D., & Ventura, M. (2017). Bifidobacteria and the infant gut: An example of co-evolution and natural selection. *Cellular and Molecular Life Sciences: CMLS*, *75*(1), 103–118. https://doi.org/10.1007/s00018-017-2672-0

Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J. R., Vázquez-Baeza, Y., Birmingham, A., Hyde, E. R., & Knight, R. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome*, *5*(1), 27. https://doi.org/10.1186/s40168-017-0237-y

Willemsen, Y., Ou, Y., Belzer, C., Arias Vásquez, A., Smidt, H., Beijers, R., & De Weerth, C. (2023). A longitudinal study of the gut microbiota during the first three years of life: Links with problem behavior and executive functions at preschool age. *Development and Psychopathology*, *36*(4), 2032–2048. https://doi.org/10.1017/S0954579423001402

Willoughby, M., Holochwost, S. J., Blanton, Z. E., & Blair, C. B. (2014). Executive functions: Formative versus reflective measurement. *Measurement: Interdisciplinary Research and Perspectives*, *12*(3), 69–95. https://doi.org/10.1080/15366367.2014.929453

Willoughby, M. T., Blair, C. B., Wirth, R. J., & Greenberg, M. (2012). The measurement of executive function at age 5: Psychometric properties and relationship to academic achievement. *Psychological Assessment*, *24*(1), 226–239. https://doi.org/10.1037/a0025361

Willoughby, M. T., Piper, B., Kwayumba, D., & McCune, M. (2019). Measuring executive function skills in young children in Kenya. *Child Neuropsychology*, *25*(4), 425–444. https://doi.org/10.1080/09297049.2018.1486395

Yada, A., Deater-Deckard, K., Takio, F., Nordenswan, E., Eskola, E., Karlsson, H., Karlsson, L., Korja, R., Nolvi, S., & Tolvanen, A. (2025). Using item response theory to address challenges with measuring executive function in infancy and early childhood [Preprint]. *PsyArXiv*. https://doi.org/10.31234/osf.io/v96ea_v1

Yang, Y., Shields, G. S., Zhang, Y., Wu, H., Chen, H., & Romer, A. L. (2022). Child executive function and future externalizing and internalizing problems: A meta-analysis of prospective longitudinal studies. *Clinical Psychology Review*, *97*, 102194. https://doi.org/10.1016/j.cpr.2022.102194

Zhang, X., Mushajiang, S., Luo, B., Tian, F., Ni, Y., & Yan, W. (2020). The composition and concordance of lactobacillus populations of infant gut and the corresponding breast-milk and maternal gut. *Frontiers in Microbiology*, *11*, 597911. https://doi.org/10.3389/fmicb.2020.597911

Zhou, H., He, K., Chen, J., & Zhang, X. (2022). LinDA: Linear models for differential abundance analysis of microbiome compositional data. *Genome Biology*, *23*(1), 95. https://doi.org/10.1186/s13059-022-02655-5

https://doi.org/10.1017/S0954579426101527 Published online by Cambridge University Press