

## Child Neuropsychology

A Journal on Normal and Abnormal Development in Childhood and Adolescence

ISSN: (Print) (Online) Journal homepage: [www.tandfonline.com/journals/ncny20](http://www.tandfonline.com/journals/ncny20)

# Association between working memory performance and parent and teacher ratings of working memory in 11-year-old children born preterm

Laura Haveri, Petriina Munck, Jukka M. Leppänen, Satu Korpela, Leena Haataja & Anna H. Nyman

**To cite this article:** Laura Haveri, Petriina Munck, Jukka M. Leppänen, Satu Korpela, Leena Haataja & Anna H. Nyman (14 Oct 2024): Association between working memory performance and parent and teacher ratings of working memory in 11-year-old children born preterm, *Child Neuropsychology*, DOI: [10.1080/09297049.2024.2415146](https://doi.org/10.1080/09297049.2024.2415146)

**To link to this article:** <https://doi.org/10.1080/09297049.2024.2415146>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 14 Oct 2024.



Submit your article to this journal [↗](#)



Article views: 169




View related articles [↗](#)



View Crossmark data [↗](#)

## Association between working memory performance and parent and teacher ratings of working memory in 11-year-old children born preterm

Laura Haveri <sup>a,b</sup>, Petriina Munck<sup>b,c</sup>, Jukka M. Leppänen<sup>a</sup>, Satu Korpela<sup>d</sup>,  
Leena Haataja<sup>e</sup> and Anna H. Nyman<sup>a,b</sup>

<sup>a</sup>Department of Psychology and Speech-Language Pathology, University of Turku, Turku, Finland;

<sup>b</sup>Department of Paediatrics and Adolescent Medicine, Turku University Hospital, Turku, Finland; <sup>c</sup>Faculty of Medicine, Department of Psychology and Logopedics, Unit of Psychology, University of Helsinki, Helsinki, Finland; <sup>d</sup>Child Neurology, Helsinki University Hospital, Helsinki, Finland; <sup>e</sup>Department of Paediatric Neurology, Paediatric Research Center, Children's Hospital, Helsinki University Hospital and University of Helsinki, Helsinki, Finland

### ABSTRACT

Working memory (WM) difficulties are often observed in children born preterm. We examined whether performance-based measures of WM components are associated with parent- and teacher-rated WM difficulties in the everyday life of children born very preterm and/or at very low birth weight (VPT/VLBW) at 11 years ( $n = 165$ ). The WM components as defined in the original Baddeley's model – phonological loop (PL), visuospatial sketchpad (VS), and central executive (CE) – were assessed with tasks from the Working Memory Test Battery for Children (WMTB-C) and the Wechsler Intelligence Scale for Children – Fourth edition (WISC-IV). Parents and teachers completed the WM subscale of the Behavioral Rating Inventory for Executive Functions (BRIEF). Measures of WM components were modestly associated with BRIEF scores, explaining 18.9% of the variance in parent-rated and 14.0% of teacher-rated WM difficulties. CE was the component most consistently associated with parent- and teacher-rated everyday WM. To conclude, our results suggest that tasks that utilize CE functions may best reflect WM outside of controlled test settings in the follow-up of VPT/VLBW children. However, performance and rating-scale measures provide unique information and are both needed to comprehensively assess WM skills.



### ARTICLE HISTORY

Received 14 November 2023  
Accepted 4 October 2024

### KEYWORDS

Working memory; preterm birth; performance-based measures; behavioral rating scales; WM components

Survival rates of infants born prematurely have improved, leading to an increase in the number of individuals at risk for lifelong medical and social consequences associated with preterm birth (Moster et al., 2008), as well as heightened risk for difficulties in cognitive development (e.g., Aarnoudse-Moens et al., 2009; Brydges et al., 2018; Twilhaar et al., 2018). The prevalence of cognitive difficulties in children born very preterm (VPT; <32 gestational weeks) and/or at very low birth weight (VLBW; birthweight  $\leq 1500$  g)

**CONTACT** Laura Haveri  [laura.haveri@utu.fi](mailto:laura.haveri@utu.fi)  Department of Psychology and Speech-Language Pathology, University of Turku, Turku 20014, Finland

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

remains high compared to term-born children, despite advances in care practices (Anderson, 2014; Nyman et al., 2017; Twilhaar et al., 2018). In addition, VPT/VLBW children who perform within the average range in global cognitive assessments may have specific neuropsychological vulnerabilities, for example, in working memory (WM) (Aarnoudse-Moens et al., 2012; Anderson, 2014; Mulder et al., 2009; Taylor & Clark, 2016; van Houdt et al., 2019). It is of particular concern that VPT/VLBW children seem to be vulnerable to WM difficulties when faced with heightened cognitive demands (Wehrle et al., 2016; Woodward et al., 2022). These difficulties are not necessarily captured by brief WM assessments in controlled test settings and may only manifest in everyday life. Thus, it is important to understand whether and how standard neuropsychological assessments are associated with everyday functioning in VPT/VLBW children.

WM is a fundamental aspect of human cognition and is widely considered one of the key mechanisms that mediate individual differences in everyday cognitive functions (e.g., Alloway & Alloway, 2010; Gathercole, Pickering, Knight, et al., 2004; Kyllonen & Christal, 1990; Schubert et al., 2023). WM is considered as one of the core components of executive functions, which are often defined as a set of high-level cognitive functions that enable goal-directed behavior, especially in novel situations that require problem-solving (Gioia et al., 2000; Lezak et al., 2004; Miyake et al., 2000). Baddeley and Hitch (1974) defined WM as an ability to sustain and process a limited amount of information in the mind temporarily and proposed a multi-component model of WM for processing qualitatively different information (Baddeley, 2000; Baddeley & Hitch, 1974). This model comprises two short-term memory components: a phonological loop (PL) for verbal information and a visuospatial sketchpad (VS) for visual and spatial information. These subcomponents are controlled by the central executive (CE), which is an attentional component that manages and integrates information. This model has been shown to apply to children from 6 years onwards (Gathercole, Pickering, Ambridge, et al., 2004). Baddeley has subsequently developed the original model further by introducing a fourth component, the episodic buffer, to the model and by highlighting the role of WM in integrating new and previously learned material (Baddeley, 2000). Combined with the notion of WM being a bottleneck skill for learning (Forsberg et al., 2022; Gathercole & Alloway, 2006), this could explain why WM abilities predict academic performance (Mulder et al., 2010) and do so better than the intelligence quotient (Alloway & Alloway, 2010; Titz & Karbach, 2014).

Previous research has consistently shown that difficulties in WM are more common in VPT/VLBW children than term-born children (Aarnoudse-Moens et al., 2012; Mulder et al., 2010; Mürner-Lavanchy et al., 2014; Wehrle et al., 2016). To our knowledge, only one study has applied Baddeley's theoretical framework to WM in VPT/VLBW children (Korpela et al., 2018). Korpela et al. (2018) found that the PL of VPT/VLBW children was intact, but their VS and CE were weaker compared to the test norms. In other studies, findings concerning verbal memory functions in school-aged VPT/VLBW children vary, as some studies show differences between VPT/VLBW and term-born children (Aanes et al., 2019), while others do not (Everts et al., 2019; Ni et al., 2011; Ritter et al., 2014). Studies focusing on visuospatial memory functions have provided more consistent evidence for difficulties in VPT/VLBW children compared to term-born children (Clark & Woodward, 2010; Retzler et al., 2022; Trickett et al., 2022; Woodward et al., 2022). Problems in CE may also exist, as VPT/VLBW children may be more affected by

the cognitive load and difficulty level of a task than term-born controls (Retzler et al., 2022; Woodward et al., 2022). Using a comprehensive model of WM can help to understand and differentiate the specific nature of the vulnerabilities in WM related to preterm birth.

In performance-based measures, WM is often operationalized as WM capacity, defined as the quantity of new information individuals can sustain and process in their minds in a controlled test setting. Another way to measure WM and acquire information about WM beyond test settings is to observe the behavioral manifestations of WM in everyday life. One of the few questionnaires used to assess everyday WM can be found in the Behavior Rating Inventory of Executive Functions (BRIEF; Gioia et al., 2000). This is a rating scale for parents and teachers that has been developed to assess the frequency of problematic behavior related to executive functions in everyday settings at home and school. BRIEF consists of three indexes and eight subscales for different areas of executive functions, one of which is a scale to measure WM. The authors of BRIEF describe that the WM scale measures “the capacity to hold information in mind for the purpose of completing a task” (Gioia et al., 2000, p. 19). The authors report good validity for the scale as well as convergent and divergent associations with other questionnaires (Gioia et al., 2000). The BRIEF WM scale includes questions assessing sustained attention integral to the WM function (Gioia et al., 2000) and has been criticized for overlapping with attention disorder questionnaires (e.g., see Soto et al., 2020). However, when compared with other WM assessments, BRIEF WM has demonstrated strong to moderate correlations with rating scales designed to measure WM in everyday life (see Law et al., 2021). Hereafter, we will use the term “everyday WM” to refer to WM as measured using the BRIEF WM scale.

In some studies, VPT/VLBW children have shown more clinically significant problems in everyday WM than term-born children according to parent (Ritter et al., 2014) and teacher ratings (Nyman et al., 2019). Studies that have used both parent and teacher ratings have shown differences between the raters, and difficulties may sometimes only be reflected in the ratings given by the parent or the teacher (Nyman et al., 2019; O’Meagher et al., 2019). The differences found between the raters and the settings highlight the importance of acquiring information from multiple sources.

Performance-based and rating measures could be expected to be associated if they provide information about the same underlying cognitive construct. However, the associations between these two methods of assessment have been shown, at best, to be inconsistent and modest in reviews of executive function measurement, and it has been suggested that they assess different aspects of executive functions (see Isquith et al., 2013; Toplak et al., 2013). In general, performance-based tasks assess specific executive functions in an examiner-led standardized setting over a short time. The measures used to quantify these functions may vary between studies and between specific executive function tasks (i.e., whether performance is measured by response accuracy or response time, or both). The rating scales, in turn, assess behaviors related to executive functions in everyday life situations through the rater’s perspective. Regarding WM in particular, the scales are intended to provide us with information about how reduced WM capacity is manifested in everyday functioning; for example, it may present as forgetfulness or as difficulties following multistep instructions (e.g., Alloway et al., 2009; Gioia et al., 2000). Previous studies have found weak

to moderate correlations between measured WM capacity and parent- and teacher-rated everyday WM across different populations (e.g., Howarth et al., 2013; Toplak et al., 2008; Vugs et al., 2014). Previously, Ritter et al. (2014) have examined the association between WM capacity and everyday WM in a sample of school-aged VPT/VLBW children. In their study, verbal WM was significantly, but weakly, correlated with the parent rating of everyday WM ( $r = -0.29$ ) (Ritter et al., 2014).

To summarize, VPT/VLBW children have been found to have difficulties in performance-based measures of WM and ratings of everyday WM, although these difficulties seem to vary depending on the specific modality of a performance-based task and the setting in which the child's functioning is assessed with ratings. Furthermore, even though there is evidence of moderate associations between the performance-based and rating measures of WM, the interpretation of the existing studies is complicated by the use of varying methodologies across different studies. Specifically, none of the studies so far have examined, within the same study, how different components of WM (such as in Baddeley's model) are associated with parent- and teacher-rated everyday WM in the preterm population. This question has theoretical value, as it links Baddeley's model with the behavior of children in everyday life. It also has value for clinicians, as investigating the commonalities and potential discrepancies between the two types of measurement may help to understand which common neuropsychological measures most informatively reflect the severity of WM problems in the everyday life of VPT/VLBW children.

### ***The aim of this study***

In the present study, we investigated the associations between WM performance measures and the frequency of problems in parent- and teacher-rated everyday WM. To our knowledge, this is the first study to comprehensively assess the association between the different WM components of Baddeley's model (Baddeley & Hitch, 1974) and everyday WM problems in the preterm population using the BRIEF parent and teacher questionnaire. The aim of the present study was to examine how the three WM components are associated with parent- and teacher-rated everyday WM. Based on prior work, our hypothesis was that poorer performance in VS and CE would be associated with more parent- and teacher-rated problems in WM in everyday life.

## **Materials and methods**

### ***Participants and demographic characteristics***

This study is part of the prospective, multidisciplinary follow-up study PIPARI (Development and Functioning of Very Low Birthweight Infants from Infancy to School Age). The sample has been previously described in, for example, Uusitalo et al. (2021) and consists of infants born preterm (gestational weeks < 37) with very low birth weight (birth weight  $\leq 1500$  g) at Turku University Hospital between January 1 2001, and December 31 2006. From January 1 2004, the inclusion criteria were expanded to include all children born at a very low gestational age (gestational weeks < 32) regardless of their birth weight. The children included were required to have at least one Finnish or Swedish-speaking parent and to live in the catchment area. The only developmental

disorders excluded were severe central nervous system malformations (e.g., congenital hydrocephalus and meningomyelocele) or diagnosed genetic syndromes affecting the development of the child. The flowchart describing the selection and retention of participants is shown in the [Appendix \(Figure A1\)](#).

The background data – consisting of prenatal, delivery, neonatal, and developmental outcome data as well as sociodemographic data – were collected systematically as part of the PIPARI study protocol. Brain magnetic resonance imaging (MRI) was performed at term age and the results were categorized using the classification of normal, minor, and major brain pathology (see Maunu et al., 2009, for more details). The children of families that withdrew from the follow-up assessment and the children in the final sample differed significantly with regards to multiple pregnancies ( $p = .04$ ). In the final sample, there were more multiple pregnancies than within the withdrawn families. No other differences between the two groups were found in the demographic variables ([Table 1](#)).

The final sample consisted of 165 VPT/VLBW children who took part in the cognitive assessment, which was conducted at 11 years of age by two clinical neuropsychologists. The assessments were completed during one uniformly structured visit to Turku University Hospital. Parents were given the parent questionnaire of Behavioral Rating Inventory of Executive Functions (BRIEF; Gioia et al., 2000), which they could complete and return either during the visit or take with them to be completed at home and returned by mail. The parents were also given a BRIEF teacher questionnaire and self-addressed envelopes that they were asked to give to their child's teachers. The BRIEF questionnaires were returned by 155

**Table 1.** Sample characteristics ( $n = 165$ ).

Variable	Values
Antenatal corticosteroids, $n$ (%)	156 (94.5)
Gestational age (week), mean (SD) [min, max]	29.05 (2.75) [23.0, 35.9]
Birth weight (g), mean (SD) [min, max]	1125.58 (317.62) [400, 2120]
Small for gestational age <sup>a</sup> , $n$ (%)	56 (33.9)
Male, $n$ (%)	89 (52.1)
Multiple birth, $n$ (%)	56 (33.9)
Cerebral palsy, $n$ (%)	9 (5.5)
Treated retinopathy of prematurity, $n$ (%)	5 (3.0)
Operated necrotizing enterocolitis, $n$ (%)	7 (4.2)
Sepsis or meningitis, $n$ (%)	31 (18.8)
Bronchopulmonary dysplasia, $n$ (%)	20 (12.1)
Ductal ligation, $n$ (%)	73 (44.2)
Brain MRI grading, $n$ (%)	
Normal	91 (55.2)
Minor	27 (16.4)
Major	42 (25.5)
Level of maternal education, $n$ (%)	
Low	19 (11.5)
Intermediate	67 (40.6)
High	68 (41.2)
Level of paternal education, $n$ (%)	
Low	20 (12.1)
Intermediate	98 (59.4)
High	35 (21.2)
5 years full-scale IQ < 70, $n$ (%)	4 (2.3)

MRI = magnetic resonance imaging; IQ = intelligence quotient.

<sup>a</sup>Defined as birth weight of < -2 SDs below the mean according to the age and gender-specific Finnish growth charts.

parents (94%) and 158 (96%) teachers. Ethical approval was obtained for the study protocol from The Ethical Committee of the Hospital District of Southwest Finland in December 2000 and January 2012. All parents and children who agreed to participate were provided with written and oral information about the study and signed a written informed consent.

### Cognitive development

General cognitive development was assessed with the Finnish version of the Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV; Wechsler, 2010). For this study, the General Ability Index (GAI) was used, consisting of six subtests of WISC-IV, excluding the WM and processing speed indexes (Wechsler, 2003). Developmental delay was defined as a GAI score of 1.0 SD below the mean of the test norms (GAI < 85).

### Working memory components

As described in Korpela et al. (2018), the WM domains of the Baddeley and Hitch (1974) model including PL, VS, and CE were assessed with tests selected from the Finnish translation of WISC-IV (Digit Recall Forward & Backward) and Working Memory Test Battery for Children (WMTB-C, Pickering & Gathercole, 2001, the Finnish version of Nonword Recall, Counting Recall, Block Recall, & Maze Recall). The episodic buffer of Baddeley's updated model (2000) is not as well validated as the other components, and its operationalization has proven to be challenging (e.g., see Gray et al., 2017; Nobre et al., 2013). Therefore, it was not included in this study. For all tests, a deficit was defined as a 1.0 SD below the mean of the test norms (standard score less than 85). A Finnish data sample of 68 children was collected to ensure that the British test design could be applied to the Finnish population and the British norms were used for comparisons (see Korpela et al., 2018, for more details). The descriptives of WM component and subtest raw scores and standard scores are shown in Table 2.

**Table 2.** WM performance and everyday WM raw and standard scores compared to normative mean.

	<i>n</i>	Raw scores		Standard scores		<i>t</i>	<i>p</i>
		Mean ( <i>SD</i> )	[min, max]	Mean ( <i>SD</i> )	[min, max]		
<b>WM component/Subtest</b>							
Phonological loop	159			104.13 (14.14)	[60, 142]	3.67	<.001
Digit Recall Forward <sup>a</sup>	165	6.85 (1.63)	[2, 12]	100.91 (14.35)	[55, 145]	0.81	.42
Nonword Recall <sup>b</sup>	159	16.00 (3.36)	[7, 24]	107.64 (17.86)	[59, 145]	5.39	<.001
Visuospatial sketchpad	164			89.10 (12.16)	[57, 112]	-11.48	<.001
Maze Recall <sup>b</sup>	164	19.21 (6.03)	[5, 31]	89.63 (16.94)	[57, 124]	-7.83	<.001
Block Recall <sup>b</sup>	165	25.79 (3.66)	[12, 35]	88.38 (14.06)	[57, 125]	-10.62	<.001
Central executive	165			81.17 (19.02)	[55, 133]	-12.72	<.001
Digit Recall Backward <sup>a</sup>	165	6.09 (1.57)	[0, 10]	78.88 (31.02)	[55, 145]	-8.75	<.001
Counting Recall <sup>b</sup>	165	19.47 (4.66)	[6, 31]	83.45 (14.29)	[55, 121]	-14.87	<.001
<b>Everyday WM<sup>c</sup></b>							
Parent rating	155	16.06 (4.95)	[10, 30]	52.97 (11.31)	[38, 89]	3.27	.001
Teacher rating	158	14.68 (4.83)	[10, 28]	56.15 (13.01)	[43, 99]	5.93	<.001

WM = Working Memory. <sup>a</sup>Standard scores derived from WISC-IV Finnish normative sample. <sup>b</sup>Standard scores derived from WMTB-C British normative sample. <sup>c</sup>Scores in WM subscale of Behavioral Rating Inventory of Executive Functions. For WM components normative group has  $M = 100$  ( $SD = 15$ ), and for everyday WM,  $M = 50$  ( $SD = 10$ ).

### *Phonological loop (PL)*

The Digit Recall Forward subtest (WISC-IV; Wechsler, 2010) and the Nonword Recall subtest (Finnish adaptation, based on the task from WMTB-C; Pickering & Gathercole, 2001) were used to assess the PL. In Digit Recall, the task involves listening to and recalling digits in the correct order from a spoken list of digits presented with a 1-s interval between the digits. The digit lists are randomly generated sequences of numbers, and the length of the list increases after at least one correct recall of two lists with the same length. Nonword Recall requires listening to and recalling nonwords from lists of nonwords constructed by randomly combining two common Finnish syllables. The task was divided into blocks, and each block consisted of six trials of the same difficulty level (i.e., six lists with the same number of nonwords). The task was initiated with practice trials consisting of one- to three-nonword lists. The number of nonwords in the starting block of the actual task was determined by the length of the correctly recalled list in the practice trials. Upon four correct recalls within a block, the task proceeded to the next block in which the list length increased by one. The test continued until three errors were made within a block.

### *Visuospatial sketchpad (VS)*

The Block Recall and Maze Recall subtests of the WMTB-C (Pickering & Gathercole, 2001) were used to assess the VS, and followed the same protocol as Nonword Recall. In Block Recall, the test administrator tapped a sequence of blocks on a block board and asked the child to repeat the order in which the blocks were tapped. The blocks were arranged in a random configuration. The same level of difficulty was presented up to a maximum of six times. Upon four correct recalls, the difficulty level was increased by adding one block to the sequence of tapped blocks. In Maze Recall, the child was shown two-dimensional routes marked on pictures of mazes. The test administrator demonstrated drawing the route with a finger and asked the child to show the same route on a picture of an empty maze. A maximum of six mazes with the same level of difficulty were shown. After the child had achieved four correct recalls, the complexity of the mazes was increased.

### *Central executive (CE)*

CE was assessed using the Digit Recall Backward task of WISC-IV (Wechsler, 2010) and the Counting Recall task of WMTB-C (Pickering & Gathercole, 2001). In Digit Recall Backward, the task was to listen to and recall digits in reverse order from a spoken list of digits presented at 1-s intervals. As in Digit Recall Forward, the lists were random, and the length increased if a list was recalled correctly. In Counting Recall, the test administrator showed arrays of dots on a series of cards to the child and asked them to point and count aloud the number of dots on the cards. After the cards were shown, the child was asked to recall the number of dots on the cards in the order in which they were presented. The number of dots varied between the cards. The Counting Recall had the same protocol as the other WMTB-C tasks presented above. A maximum of six trials at the same level of difficulty were presented and after four correct recalls, the difficulty level was increased by adding one card to the series.

### *Parent and teacher ratings of everyday working memory*

To measure everyday WM, the Working Memory subscale of the BRIEF questionnaire was chosen (Gioia et al., 2000). Finnish translations of both the parent and teacher forms were used (translation by three clinical neuropsychologists A. Nyman, M. Kuusisto, and J. Torsti, approved by PAR Psychological Assessment Resources, Inc). The BRIEF questionnaire consists of statements that utilize a 3-point Likert scale to rate the frequency of the problem on a scale of (1) “Never”, (2) “Sometimes,” and (3) “Often.” The WM subscale consists of 10 statements (such as “when given three things to do, remembers only the first or the last one”). The sum of the scores is transferred into T-scores ( $M = 50$ ,  $SD = 10$ ), which are standardized by age and sex. Higher T-scores represent more reported difficulties. A cutoff score of 64 is pre-established in the manual of the questionnaire, and scores above this indicate clinically significant difficulties in WM in everyday life.

### *Statistical analysis*

The data encompassing pre- and neonatal background factors, cognitive assessment scores, and questionnaire answers were numerically encoded. To identify differences in background characteristics between the participating families and those who withdrew, we used independent sample t-tests for continuous variables and chi-square tests for categorical variables. Differences between the children’s WM component scores and the normative sample means were examined using one-sample t-tests. Preliminary analyses were conducted for the total sample ( $n = 165$ ) and a subsample including only children with GAI scores of 85 or above ( $n = 120$ ). In both samples, the distribution of variables was approximately normal, apart from the positively skewed rating in everyday WM by teachers and parents. The correlations between the WM components and everyday WM were investigated by calculating pairwise correlations (Spearman’s rho). The level of agreement between groups with deficits based on clinical cutoff points was analyzed with Cohen’s Kappa.

To assess whether performance-based measures of WM are associated with parent- and teacher-rated problems in WM, we constructed two separate linear mixed models (LMMs). Applying LMMs allowed the clustering caused by twin and triplet children from the same family to be taken into account. The first model included fixed effects for PL, VS, and CE as independent variables and parent-rated WM scores as the dependent variable, with GAI, sex, and maternal and paternal education as covariates. MRI status was not added to the models because we wanted to focus on background factors at 11 years and limit the number of covariates in the models to avoid overfitting as well as multicollinearity with GAI. The second model mirrored the first, substituting the dependent variable with teacher-rated WM scores. These models were expanded to incorporate random intercepts for families, acknowledging the potential dependency of outcomes within family clusters. In the teachers’ model, initial diagnostics revealed two influential cases with Cook’s distance values greater than one. To assess the robustness of our findings, a sensitivity analysis was conducted. A second model was also conducted, excluding the influential cases. A comparison of the two models using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) indicated

a better fit for the second model ( $\Delta AIC = 35.25$ ,  $\Delta BIC = 35.39$ ). Results from both teachers' models – with and without the influential cases – are presented. P-values of  $< .05$  were considered statistically significant. All statistical analyses, with the exception of the LMMs, were performed using IBM SPSS Statistics (version 27). The LMMs were implemented in R (R Core Team, 2023) using the lme4 package (Bates et al., 2015).

## Results

### Preliminary analysis

The mean age of the children at the time of assessment was 11.14 years ( $SD = 0.25$ , range 10.57–11.66), and the mean GAI was 90.64 ( $SD = 15.23$ , range 40–120). Forty-five children (27.3%) had a cognitive delay ( $GAI < 85$ ). GAI was moderately correlated with parent-rated ( $r = -.27$ ,  $p < .001$ ), and teacher-rated WM ( $r = -.31$ ,  $p < .001$ ) scores. No correlations were found between the WM measures and gestational age or birth weight and thus, these were not examined in further analyses.

### Association between the WM components and everyday WM

Table 3 presents the correlations between the WM components and everyday WM for both the total sample and a subgroup of children with a GAI of 85 or above. For the total sample, significant negative correlations between WM components and everyday WM were observed, with Spearman's rho ranging from weak to moderate in the parent ( $r_s = -.34$  to  $-.21$ ) and teacher ratings ( $r_s = -.34$  to  $-.21$ ). The negative correlation indicates that a higher score in the WM component was associated with fewer parent- and teacher-rated WM problems. In the subgroup of children with a GAI of 85 or above, the correlation between CE and parent-rated everyday WM was statistically significant ( $r_s = -.21$ ,  $p = .025$ ), as were the correlations between teacher-rated WM scores and CE ( $r_s = -.23$ ,  $p = .012$ ) and VS ( $r_s = -.24$ ,  $p = .011$ ).

Further examination of the association was carried out based on categorical groups, using clinical cutoff points for WM components and everyday WM. A deficit in WM was defined as a WM component standard score of less than 85 and, for everyday WM, as T-scores equal to or above 65. The proportion of children in these groups is shown in Table 4. A slight agreement was found between parent-rated WM scores and VS. Moreover, agreements were slight between teacher-rated WM scores and the WM components.

**Table 3.** Correlations between WM components and everyday WM by rater and GAI.

WM Component	Everyday WM <sup>a</sup>			
	Parent		Teacher	
	All	GAI $\geq 85$	All	GAI $\geq 85$
Phonological loop	-.21**	-.17	-.21**	-.16
Visuospatial sketchpad	-.34**	-.15	-.34**	-.24*
Central executive	-.34**	-.21*	-.31**	-.23*

One-tailed Spearman's correlation coefficients are presented. WM = working memory; GAI = general ability index.

<sup>a</sup>Scores in WM subscale of Behavioral Rating Inventory of Executive Functions.

\* $p < .05$ ; \*\* $p < .01$ .

**Table 4.** Cross-tabulation for WM component and everyday WM by cutoff points.

WM component	Parent-rated everyday WM <sup>a</sup>			Teacher-rated everyday WM <sup>a</sup>		
	<65	≥65	Kappa (95% CI)	<65	≥65	Kappa (95% CI)
Phonological loop, <i>n</i> (%)						
≥85	121 (87.1)	18 (12.9)	0.17 (−0.03 to 0.38)	111 (78.2)	31 (21.8)	0.17 (0.01 to 0.33)
<85	6 (60.0)	4 (40.0)		4 (40.0)	6 (60.0)	
Visuospatial sketchpad, <i>n</i> (%)						
≥85	94 (90.4)	10 (9.6)	0.19 (0.04 to 0.34)	86 (80.4)	21 (19.6)	0.17 (0.01 to 0.33)
<85	37 (74.0)	13 (26.0)		32 (64.0)	18 (36.0)	
Central executive, <i>n</i> (%)						
≥85	46 (90.2)	5 (9.8)	0.05 (−0.03 to 0.13)	42 (85.7)	7 (14.3)	0.11 (0.01 to 0.21)
<85	86 (82.7)	18 (17.3)		77 (70.6)	32 (29.4)	

WM = working memory; CI = confidence interval. <sup>a</sup>Scores in WM subscale of Behavioral Rating Inventory of Executive Functions.

Two separate LMMs were created to examine whether CE, VS, and PL, incorporated in the same model, were associated with parent- and teacher-rated WM scores. Additionally, the models explored how controlling for covariates (GAI, sex, and maternal and paternal education) influenced these associations. We examined the effects in a stepwise manner, with the first step including only the WM components and the second step adding the covariates. The covariates were selected *a priori*, as they were expected to be related to everyday WM based on previous literature. In the models, family was included as a random effect. The random effects are shown in the Appendix (Table A1). The WM components explained 18.9% of the variance in parent-rated WM scores, and the VS and CE showed a significant effect on the parent-rated WM scores (Table 5). The effects of VS and CE remained significant after adjusting for covariates. In the model predicting teacher-rated WM scores, the WM components explained 14.0% of the variance, and only the effect of CE was significant (Table 6). However, when the covariates were added to the model, the effect of CE was weakened, and GAI and paternal education were significant contributors to the model. The model that excluded influential cases showed similar results (Table 6, Model 4B).

**Table 5.** Linear mixed model estimates predicting parent-rated everyday WM.

Fixed effect	Parent-rated everyday WM <sup>a</sup>					
	Model 1: without covariates			Model 2: with covariates		
	Estimate	95% CI	<i>p</i>	Estimate	95% CI	<i>p</i>
Intercept	53.10	51.40–54.81	<.001	53.55	50.84–56.25	<.001
Phonological loop	−0.13	−2.41–2.15	>.9	0.39	−1.88–2.66	.7
Visuospatial sketchpad	−2.85	−5.25 – −0.44	.021	−2.61	−5.09 – −0.14	.039
Central executive	−3.79	−6.30 – 1.28	.003	−3.70	−6.26 – −1.14	.005
GAI				−1.03	−3.15–1.09	.3
Gender <sup>b</sup>				−0.43	−3.66–2.80	.8
Maternal education <sup>c</sup>				−2.48	−6.00–1.05	.2
Paternal education <sup>c</sup>				1.81	−2.44–6.05	.4

WM = Working Memory; GAI = General Ability Index; CI = Confidence Interval. <sup>a</sup>Scores in WM subscale of Behavioral Rating Inventory of Executive Functions. <sup>b</sup>Reference category: male. <sup>c</sup>Reference category: low or intermediate education level.

Table 6. Linear mixed model estimates predicting teacher-rated everyday WM.

Fixed effect	Teacher-rated everyday WM <sup>a</sup>											
	Model 3: without covariates				Model 4A: with covariates				Model 4B: influential cases excluded			
	Estimate	95% CI	<i>p</i>	Estimate	95% CI	<i>p</i>	Estimate	95% CI	<i>p</i>			
Intercept	55.81	53.83–57.80	<.001	57.69	54.44–60.94	<.001	56.91	53.88–59.93	<.001			
Phonological loop	–1.07	–3.54–1.41	.4	–0.13	–2.64–2.38	>.9	0.84	–1.54–3.21	.5			
Visuospatial sketchpad	–2.46	–5.09–0.18	.068	–0.84	–3.62–1.94	.5	–1.32	–3.93–1.29	.3			
Central executive	–3.36	–6.13 – –0.59	.018	–2.23	–5.11–0.65	.13	–1.85	–4.54–0.84	.2			
GAI				–2.97	–5.37 – –0.57	.016	–3.54	–5.79 – –1.28	.002			
Gender <sup>b</sup>				–2.00	–5.68–1.69	.3	–2.56	–6.00–0.88	.14			
Maternal education <sup>c</sup>				1.28	–2.96–5.52	.6	2.15	–1.78–6.08	.3			
Paternal education <sup>c</sup>				–5.57	–10.68 – –0.47	.033	–5.11	–9.83 – –0.39	.034			

WM = Working Memory; GAI = General Ability Index; CI = Confidence Interval. <sup>a</sup>Scores in WM subscale of Behavioral Rating Inventory of Executive Functions. <sup>b</sup>Reference category: male. <sup>c</sup>Reference category: low or intermediate education level.

## Discussion

Our findings showed weak to moderate associations between the WM components and parent- and teacher-rated WM scores in 11-year-old VPT/VLBW children. Overall, better performance in WM components was associated with fewer problems in everyday WM as rated by parents and teachers. We found a partial overlap in the groups identified to be at risk for WM component and everyday WM difficulties. The explanatory value of WM components in relation to everyday WM was significant but weak, as WM components explained 14.0% to 18.9% of the variability in everyday WM. Of the WM components, only CE explained both parent- and teacher-rated WM scores, although the association with teacher rating disappeared after adjustment for covariates.

Previous studies have not examined the association between WM components of Baddeley's model and everyday WM in the preterm population in such detail. Nonetheless, our findings are in line with previous findings that have shown a modest negative correlation between WM capacity, measured with the maximum digit backward span, and parent-rated everyday WM scores in school-aged VPT/VLBW children (Ritter et al., 2014). Our findings add to these findings by showing how different WM components are associated with everyday WM, while also expanding the analysis to examine the association between WM components and teacher-rated WM scores. In heterogeneous non-preterm populations, associations between WM components and everyday WM range from modest (Howarth et al., 2013; Toplak et al., 2008) to non-significant (Conklin et al., 2008).

Studies using parent and teacher ratings have emphasized the importance of multiple raters, since the agreement between the raters has been moderate at best (Gioia et al., 2000; Nyman et al., 2019; O'Meagher et al., 2019). Aside from potential rater effects, the inconsistencies may be caused by different WM demands in school and home settings. When examining the groups at risk based on cutoff points, the agreement was poor. It appeared that this was largely due to parent and teacher ratings being unable to detect most of the children at risk for WM component difficulties. Overall, our findings support the use of multiple sources of information, as this appears to help capture behaviors unique to the setting in which the assessment is performed.

Our results suggest that PL is not associated with everyday WM in VPT/VLBW children. In line with the previously reported results for a subsample of the participants in this study (Korpela et al., 2018), our comparison of WM component scores suggests that PL could remain more intact than other components of Baddeley's WM model in VPT/VLBW children, making the associations between PL and problems with everyday WM minimal. Based on these findings, we speculate that PL tasks, as implemented in this study, may be the least informative of the WM components in terms of identifying existing everyday WM difficulties in VPT/VLBW children. WM is often measured with a Digit Recall task, which combines information from both forward and backward recall (Wechsler, 2003). It has been recommended to look at the two types of recall separately because the forward condition requires sustaining information in the mind, while the backward condition requires the processing and transformation of sustained material (Reynolds, 1997). Previously, Omizzolo et al. (2014) examined these two types of recall separately in both VPT/VLBW and term-born groups and found a difference between the groups in Backward Digit Recall, but not in Forward Digit Recall, when children with an

intelligence quotient below 70 were excluded. Our findings support this differentiation in VPT/VLBW children, suggesting that it is important not to over emphasize the PL functions in clinical WM assessment, as this subtest may underestimate the WM capacity difficulties of VPT/VLBW children related to everyday life.

An association between VS and everyday WM was found. Surprisingly, the VS contributed only to the parent-rated and not to the teacher-rated WM scores. A difference seems to exist between the modalities of the WM components, since an association was found with VS but not with PL. Consistent findings of vulnerability in visuospatial memory functions in VPT/VLBW children (Clark & Woodward, 2010; Retzler et al., 2022; Trickett et al., 2022; Woodward et al., 2022) could be reflected in the children's behavior and explain the association between VS and everyday WM. However, it remains unclear why VS did not contribute to teacher-rated everyday WM. As visuospatial memory functions have previously been associated with academic outcomes in the VPT/VLBW population (Trickett et al., 2022; Twilhaar et al., 2020; Woodward et al., 2022), VS might be related to other everyday domains of performance at school than those assessed by the BRIEF WM scales. In addition, it is plausible that parent ratings might more accurately identify everyday WM difficulties than teacher ratings, as parents typically spend more one-on-one time with the child.

We found an association between CE and both parent- and teacher-rated WM scores. However, only the association between CE and parent-rated everyday WM remained after controlling for WM components, GAI, sex, and the parents' education. The tasks that were used to measure CE require simultaneous sustaining and processing of information and involve a higher cognitive load than PL or VS. VPT/VLBW children have shown a vulnerability to decreasing performance in situations where the cognitive load increases (Retzler et al., 2022; Woodward et al., 2022). Thus, the link between CE and everyday WM may be explained by the CE task better capturing performance in everyday life settings where demands are usually higher and distractions more common. Another mechanism that could mediate the observed association with WM problems is the attentional element of CE (Baddeley, 2000; Baddeley & Hitch, 1974). Behavioral manifestations of poor attention and poor WM can be difficult to separate, and the WM subscale of BRIEF has been proven to be useful in identifying inattentive behavior (Gioia et al., 2000). Previously, an association between visuospatial WM and parent-rated inattention has been observed in VPT/VLBW children by Retzler et al. (2019). Since our study did not include direct measures of attention, we were not able to evaluate its role as a mediating mechanism, but this hypothesis can be addressed in future studies.

Global cognitive development did not contribute to parent-rated everyday WM, which means the associations found can be generalized to VPT/VLBW children with varying levels of cognitive development. However, in teacher-rated WM scores, the global cognitive development level was shown to have a stronger effect than the WM components. The effect was negative, i.e., higher global cognitive performance was associated with fewer problems in everyday WM. Speculatively, this could be due to the heightened cognitive requirements in school settings that overrule the effect of WM on behavior through, for example, compensatory strategies. When rating the children, teachers might be sensitive to their overall cognitive development, but this should be examined further in future studies. In addition, the father's education contributed to the teacher-rated everyday WM, as teachers reported fewer problems with everyday WM

when the paternal education level was higher. Previously, a positive association between cognitive development and maternal education has been observed (Sentenac et al., 2021). This could suggest higher socioeconomic status having a protective effect.

In our models, the total explanatory power was weak, and thus, in the preterm population, performance-based measures of WM are not strongly indicative of ratings of WM in everyday settings. We showed that only a minority of children performed poorly (i.e., beyond the cutoff points) in both measurements. Considering the vast number of factors that can affect the behavioral manifestation and operationalization of cognitive constructs such as WM, previous research seems to agree that, at best, modest associations are expected between the two levels of measurement (Isquith et al., 2013; Soto et al., 2020; Toplak et al., 2013). According to Toplak et al. (2013), this is because the two levels of measurement seem to evaluate different characteristics of behavioral and cognitive functioning. Performance-based measurement is often criticized for capturing the optimal performance of an individual, while impairments impacting everyday life may remain hidden (see Toplak et al., 2013). Our findings support this idea, since we found groups of children that performed within the average range in WM components yet showed difficulties in everyday WM. These groups might describe the proportion of VPT/VLBW children who perform optimally in controlled test settings but whose WM difficulties surface outside of these settings. Additionally, elevated BRIEF WM scores may reflect other aspects than WM difficulties in everyday life, for example, the child's motivation, attentional resources, and compensatory strategies (Alloway et al., 2009; Wallisch et al., 2018). Furthermore, as a questionnaire, the BRIEF is affected by the respondent's perspective, qualities, and experiences (e.g., see Roth et al., 2014). Overall, multiple levels of WM measurement are needed to understand WM deficits better, as their meaning to everyday functioning varies in the preterm population.

According to our results, it is probable that disagreement will appear in the classification of deficits when using performance-based and rating measures, but as Silver (2014) has discussed, examining the reasons behind these disagreements might reveal issues relevant to the individual's life and guide the planning of efficient support measures. Diverse interventions that can be applied in everyday contexts have been shown to be beneficial in supporting WM (see Rowe et al., 2019, for a review). In addition, interventions aiming to provide students with attentional and memory strategies in a classroom context (see Colmar & Double, 2017) could benefit not only the children with WM difficulties but all kinds of learners. Most VPT/VLBW children did not show clinically significant everyday WM difficulties, even if they performed below average in WM components. In the future, it would be interesting to investigate the possible protective factors behind this finding.

To our knowledge, this is the first study to investigate the association between WM components and everyday WM by applying a comprehensive theoretical approach to the assessment of WM and multi-informant reports of everyday WM in the preterm population. Thus, one of the strengths of this study is the use of different sources and levels of information to assess WM. This allowed us to examine and gain an understanding of the interrelations and informational value of the measurements that are commonly used in WM assessment in clinical contexts. Another strength is the data collected from a clearly defined regional cohort of VPT/VLBW children that describes, to a certain extent, the long-term WM effects in a population treated in

a university hospital. In addition to the parents' reports, the teachers' reports permitted the generalization of our findings to school settings. Moreover, the narrow age range allowed us to control for age effects and focus on a developmental time window in which the children have already had approximately 4 years of formal schooling; during this time the development of WM functions is slower than in the earlier stages of development (Ahmed et al., 2022).

One limitation of this study was the expansion of the inclusion criteria during the cohort recruitment, which added 10 children to the sample. Our final sample included a higher percentage of multiple pregnancies (33.9%) than the group of withdrawn families. However, the proportion of multiple pregnancies in the final sample is comparable to the rate of multiple pregnancies in the VPT/VLBW population in Finland, which has varied from 19% to 33% between 2005 and 2022 (Finnish Institute for Health and Welfare, 2023), suggesting that our sample is representative of the VPT/VLBW population. Another limitation was that the tasks used to assess WM were chosen from two different test batteries with different norm samples. Z-transformations of raw scores were mainly used instead of standard scores to avoid the effects caused by different norm samples. The use of z-scores was justified because the interest was in the associations between the variables and the sample, which consisted of children from a narrow age group. Standard scores were used to examine groups by cutoff points used in clinical practice, which demonstrated how the association may appear in clinical practice. A Finnish normative sample was previously collected to ensure the reliability of using the British norms of WMTB-C (see Korpela et al., 2018). Our aim was to assess WM in a versatile way and to combine information from different viewpoints, leading to a better understanding of the association between these WM measurement levels. In this study, we did not have a control group and could not compare the associations with those of term-born children.

## Conclusions

In summary, we found that the WM components differ in their association with everyday WM at home and at school in VPT/VLBW children, with CE showing the most consistent associations with everyday WM. Thus, in the follow-up of children born preterm, a comprehensive, multi-informant approach is recommended to identify vulnerabilities in the WM profile, as challenges might vary depending on the setting, the modality, or the demand of the task. If possible, greater emphasis should be given to the tasks that pose higher demands on CE since they may reveal difficulties that are reflected in everyday behavior. It is essential to identify the children at risk for WM difficulties using the most efficient tools and to understand the properties of the tools that are used; this will ensure that the implications are as reliable as possible, and support can be targeted accordingly.

## Acknowledgments

We would like to thank the members of the PIPARI (Development and Functioning of Very Low Birthweight Infants from Infancy to School Age) Study Group: Mikael Ekblad; Satu Ekblad; Eeva Ekholm; Linda Grönroos; Leena Haataja; Laura Haveri; Minttu Helin;

Mira Huhtala; Eveliina Joensuu; Max Karukivi; Riikka Korja; Katri Lahti; Helena Lapinleimu; Liisa Lehtonen; Tuomo Lehtonen; Marika Leppänen; Annika Lind; Jonna Maunu; Petriina Munck; Eeva Mäkilä; Laura Määttänen; Anna Nyman; Helena Ollila; Riitta Parkkola; Päivi Rautava; Katriina Saarinen; Tiina Saarinen; Susanna Salomäki; Virva Saunavaara; Sirkku Setänen; Matti Sillanpää; Suvi Stolt; Karoliina Uusitalo; Milla Ylijoki. We would also like to thank Ali Moazami Goodarzi for statistical support and Elizabeth Nyman for language editing.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work was supported by the Otto A. Malm foundation, the Foundation for Pediatric research, and the OLVI Foundation.

## ORCID

Laura Haveri  <http://orcid.org/0009-0004-0883-6432>

## References

- Aanes, S., Bjuland, K. J., Sripada, K., Sølsnes, A. E., Grunewaldt, K. H., Håberg, A., Løhaugen, G. C., & Skranes, J. (2019). Reduced hippocampal subfield volumes and memory function in school-aged children born preterm with very low birthweight (VLBW). *NeuroImage Clinical*, 23, 101857. <https://doi.org/10.1016/j.nicl.2019.101857>
- Aarnoudse-Moens, C. S. H., Duivenvoorden, H. J., Weisglas-Kuperus, N., Van Goudoever, J. B., & Oosterlaan, J. (2012). The profile of executive function in very preterm children at 4–12 years. *Developmental Medicine & Child Neurology*, 54(3), 247–253. <https://doi.org/10.1111/j.1469-8749.2011.04150.x>
- Aarnoudse-Moens, C. S. H., Weisglas-Kuperus, N., van Goudoever, J. B., & Oosterlaan, J. (2009). Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. *Pediatrics*, 124(2), 717–728. <https://doi.org/10.1542/peds.2008-2816>
- Ahmed, S. F., Ellis, A., Ward, K. P., Chaku, N., & Davis-Kean, P. E. (2022). Working memory development from early childhood to adolescence using two nationally representative samples. *Developmental Psychology*, 58(10), 1962–1973. <https://doi.org/10.1037/dev0001396>
- Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, 106(1), 20–29. <https://doi.org/10.1016/j.jecp.2009.11.003>
- Alloway, T. P., Gathercole, S. E., Holmes, J., Place, M., Elliott, J. G., & Hilton, K. (2009). The diagnostic utility of behavioral checklists in identifying children with ADHD and children with working memory deficits. *Child Psychiatry and Human Development*, 40(3), 353–366. <https://doi.org/10.1007/s10578-009-0131-3>
- Anderson, P. J. (2014). Neuropsychological outcomes of children born very preterm. *Seminars in Fetal & Neonatal Medicine*, 19(2), 90–96. <https://doi.org/10.1016/j.siny.2013.11.012>
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417–423. [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2)
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). Elsevier. [https://doi.org/10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1)

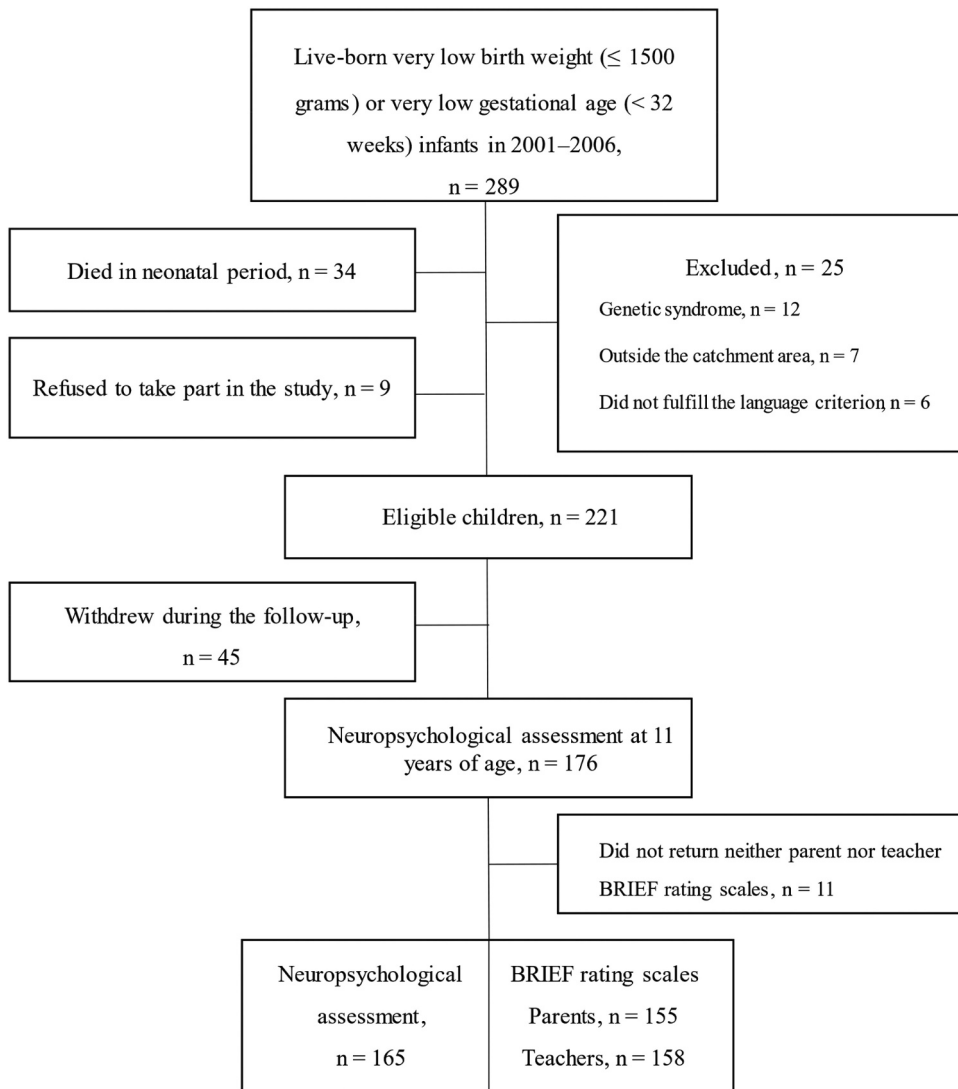
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Brydges, C. R., Landes, J. K., Reid, C. L., Campbell, C., French, N., & Anderson, M. (2018). Cognitive outcomes in children and adolescents born very preterm: A meta-analysis. *Developmental Medicine & Child Neurology*, 60(5), 452–468. <https://doi.org/10.1111/dmcn.13685>
- Clark, C. A. C., & Woodward, L. J. (2010). Neonatal cerebral abnormalities and later verbal and visuospatial working memory abilities of children born very preterm. *Developmental Neuropsychology*, 35(6), 622–642. <https://doi.org/10.1080/87565641.2010.508669>
- Colmar, S., & Double, K. (2017). Working memory interventions with children: Classrooms or computers? *Journal of Psychologists and Counsellors in Schools*, 27(2), 264–277. <https://doi.org/10.1017/jgc.2017.11>
- Conklin, H. M., Salorio, C. F., & Slomine, B. S. (2008). Working memory performance following paediatric traumatic brain injury. *Brain Injury*, 22(11), 847–857. <https://doi.org/10.1080/02699050802403565>
- Everts, R., Schöne, C. G., Mürner-Lavanchy, I., & Steinlin, M. (2019). Development of executive functions from childhood to adolescence in very preterm-born individuals – a longitudinal study. *Early Human Development*, 129, 45–51. <https://doi.org/10.1016/j.earlhumdev.2018.12.012>
- Finnish Institute of Health and Welfare. (2023). *Pienten keskusten tilastokuu 2005–2022* [Database cube of small premature infants 2002–2005]. Retrieved July 5, 2024, from <https://thl.fi/tilastot-ja-data/tilastot-aiheittain/seksuaali-ja-lisaantymisterveys/synnyttajat-synnytykset-ja-vastasyntyneet/pienten-keskusten-rekisterin-tilastokuu>
- Forsberg, A., Guitard, D., Adams, E. J., Pattanakul, D., & Cowan, N. (2022). Children’s long-term retention is directly constrained by their working memory capacity limitations. *Developmental Science*, 25(2), Article e13164. <https://doi.org/10.1111/desc.13164>
- Gathercole, S. E., & Alloway, T. P. (2006). Practitioner review: Short-term and working memory impairments in neurodevelopmental disorders: Diagnosis and remedial support. *Journal of Child Psychology and Psychiatry*, 47(1), 4–15. <https://doi.org/10.1111/j.1469-7610.2005.01446.x>
- Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, 40(2), 177–190. <https://doi.org/10.1037/0012-1649.40.2.177>
- Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, 18(1), 1–16. <https://doi.org/10.1002/acp.934>
- Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior rating inventory of executive function: BRIEF*. Psychological Assessment Resources.
- Gray, S., Green, S., Alt, M., Hogan, T., Kuo, T., Brinkley, S., & Cowan, N. (2017). The structure of working memory in young children and its relation to intelligence. *Journal of Memory and Language*, 92, 183–201. <https://doi.org/10.1016/j.jml.2016.06.004>
- Howarth, R. A., Ashford, J. M., Merchant, T. E., Ogg, R. J., Santana, V., Wu, S., Xiong, X., & Conklin, H. M. (2013). The utility of parent report in the assessment of working memory among childhood brain tumor survivors. *Journal of the International Neuropsychological Society*, 19(4), 380–389. <https://doi.org/10.1017/S1355617712001567>
- Isquith, P. K., Roth, R. M., & Gioia, G. (2013). Contribution of rating scales to the assessment of executive functions. *Applied Neuropsychology: Child*, 2(2), 125–132. <https://doi.org/10.1080/21622965.2013.748389>
- Korpela, S., Nyman, A., Munck, P., Ahtola, A., Matomäki, J., Korhonen, T., Parkkola, R., & Haataja, L. (2018). Working memory in very-low-birthweight children at the age of 11 years. *Child Neuropsychology*, 24(3), 338–353. <https://doi.org/10.1080/09297049.2016.1260101>
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, 14(4), 389–433. [https://doi.org/10.1016/S0160-2896\(05\)80012-1](https://doi.org/10.1016/S0160-2896(05)80012-1)
- Law, C., Juraskova, I., & Lah, S. (2021). Systematic review of pediatric memory questionnaires. *Child Neuropsychology*, 27(6), 734–781. <https://doi.org/10.1080/09297049.2021.1888908>

- Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, H. J., & Fischer, J. S. (2004). *Neuropsychological assessment* (4th ed.). Oxford University Press.
- Maunu, J., Parkkola, R., Rikalainen, H., Lehtonen, L., Haataja, L., Lapinleimu, H., & Pipari Group. (2009). Brain and ventricles in very low birth weight infants at term: A comparison among head circumference, ultrasound, and magnetic resonance imaging. *Pediatrics*, *123*(2), 617–626.
- 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>
- Moster, D., Lie, R. T., & Markestad, T. (2008). Long-term medical and social consequences of preterm birth. *New England Journal of Medicine*, *359*(3), 262–273. <https://doi.org/10.1056/NEJMoa0706475>
- Mulder, H., Pitchford, N. J., Hagger, M. S., & Marlow, N. (2009). Development of executive function and attention in preterm children: A systematic review. *Developmental Neuropsychology*, *34*(4), 393–421. <https://doi.org/10.1080/87565640902964524>
- Mulder, H., Pitchford, N. J., & Marlow, N. (2010). Processing speed and working memory underlie academic attainment in very preterm children. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *95*(4), F267–F272. <https://doi.org/10.1136/adc.2009.167965>
- Mürner-Lavanchy, I., Ritter, B. C., Spencer-Smith, M. M., Perrig, W. J., Schroth, G., Steinlin, M., & Everts, R. (2014). Visuospatial working memory in very preterm and term born children—impact of age and performance. *Developmental Cognitive Neuroscience*, *9*, 106–116. <https://doi.org/10.1016/j.dcn.2014.02.004>
- Ni, T.-L., Huang, C.-C., & Guo, N.-W. (2011). Executive function deficit in preschool children born very low birth weight with normal early development. *Early Human Development*, *87*(2), 137–141. <https://doi.org/10.1016/j.earlhumdev.2010.11.013>
- Nobre, A. D. P., Rodrigues, J. D. C., Sbicigo, J. B., Piccolo, L. D. R., Zortea, M., Duarte Junior, S., & Salles, J. F. D. (2013). Tasks for assessment of the episodic buffer: A systematic review. *Psychology & Neuroscience*, *6*(3), 331–343. <https://doi.org/10.3922/j.psns.2013.3.10>
- Nyman, A., Korhonen, T., Munck, P., Parkkola, R., Lehtonen, L., & Haataja, L. (2017). Factors affecting the cognitive profile of 11-year-old children born very preterm. *Pediatric Research*, *82*(2), 324–332. <https://doi.org/10.1038/pr.2017.64>
- Nyman, A., Munck, P., Koivisto, M., Hagelstam, C., Korhonen, T., Lehtonen, L., & Haataja, L. (2019). Executive function profiles at home and at school in 11-year-old very low birth weight or very low gestational age children. *Journal of Developmental & Behavioral Pediatrics*, *40*(7), 547–554. <https://doi.org/10.1097/DBP.0000000000000689>
- O’Meagher, S., Norris, K., Kemp, N., & Anderson, P. (2019). Examining the relationship between performance-based and questionnaire assessments of executive function in young preterm children: Implications for clinical practice. *Child Neuropsychology*, *25*(7), 899–913. <https://doi.org/10.1080/09297049.2018.1531981>
- Omizolo, C., Scratch, S. E., Stargatt, R., Kidokoro, H., Thompson, D. K., Lee, K. J., Cheong, J., Neil, J., Inder, T. E., Doyle, L. W., & Anderson, P. J. (2014). Neonatal brain abnormalities and memory and learning outcomes at 7 years in children born very preterm. Neonatal brain abnormalities and memory and learning outcomes at 7 years in children born very preterm. *Memory (Hove, England)*, *22*(6), 605–615. <https://doi.org/10.1080/09658211.2013.809765>
- Pickering, S., & Gathercole, S. (2001). *Working memory test battery for children (WMTB-C) Manual*. Pearson Educational Ltd.
- R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Retzler, J., Johnson, S., Groom, M. J., & Cragg, L. (2022). A comparison of simultaneous and sequential visuo-spatial memory in children born very preterm. *Child Neuropsychology*, *28*(4), 496–509. <https://doi.org/10.1080/09297049.2021.1993808>
- Retzler, J., Johnson, S., Groom, M., Hollis, C., Budge, H., & Cragg, L. (2019). Cognitive predictors of parent-rated inattention in very preterm children: The role of working memory and

- processing speed. *Child Neuropsychology*, 25(5), 617–635. <https://doi.org/10.1080/09297049.2018.1510908>
- Reynolds, C. R. (1997). Forward and backward memory span should not be combined for clinical analysis. *Archives of Clinical Neuropsychology*, 12(1), 29–40. <https://doi.org/10.1093/arclin/12.1.29>
- Ritter, B. C., Perrig, W., Steinlin, M., & Everts, R. (2014). Cognitive and behavioral aspects of executive functions in children born very preterm. *Child Neuropsychology*, 20(2), 129–144. <https://doi.org/10.1080/09297049.2013.773968>
- Roth, R. M., Isquith, P. K., & Gioia, G. A. (2014). Assessment of executive functioning using the behavior rating inventory of executive function (BRIEF). In S. Goldstein & J. A. Naglieri (Eds.), *Handbook of executive functioning* (pp. 301–331). Springer New York. [https://doi.org/10.1007/978-1-4614-8106-5\\_18](https://doi.org/10.1007/978-1-4614-8106-5_18)
- Rowe, A., Titterington, J., Holmes, J., Henry, L., & Taggart, L. (2019). Interventions targeting working memory in 4–11 year olds within their everyday contexts: A systematic review. *Developmental Review*, 52, 1–23. <https://doi.org/10.1016/j.dr.2019.02.001>
- Schubert, A.-L., Löffler, C., Sadus, K., Göttmann, J., Hein, J., Schröer, P., Teuber, A., & Hagemann, D. (2023). Working memory load affects intelligence test performance by reducing the strength of relational item bindings and impairing the filtering of irrelevant information. *Cognition*, 236, Article 105438. <https://doi.org/10.1016/j.cognition.2023.105438>
- Sentenac, M., Benhammou, V., Aden, U., Ancel, P.-Y., Bakker, L. A., Bakoy, H., Barros, H., Baumann, N., Bilsteen, J. F., Boerch, K., Croci, I., Cuttini, M., Draper, E., Halvorsen, T., Johnson, S., Källén, K., Land, T., Lebeer, J., Lehtonen, L. ... Zeitlin, J. (2021). Maternal education and cognitive development in 15 European very-preterm birth cohorts from the RECAP preterm platform. *International Journal of Epidemiology*, 50(6), 1824–1839. <https://doi.org/10.1093/ije/dyab170>
- Silver, C. H. (2014). Sources of data about children's executive functioning: Review and commentary. *Child Neuropsychology*, 20(1), 1–13. <https://doi.org/10.1080/09297049.2012.727793>
- Soto, E. F., Kofler, M. J., Singh, L. J., Wells, E. L., Irwin, L. N., Groves, N. B., & Miller, C. E. (2020). Executive functioning rating scales: Ecologically valid or construct invalid? *Neuropsychology*, 34(6), 605–619. <https://doi.org/10.1037/neu0000681>
- Taylor, H. G., & Clark, C. A. C. (2016). Executive function in children born preterm: Risk factors and implications for outcome. *Seminars in Perinatology*, 40(8), 520–529. <https://doi.org/10.1053/j.semperi.2016.09.004>
- Titz, C., & Karbach, J. (2014). Working memory and executive functions: Effects of training on academic achievement. *Psychological Research*, 78(6), 852–868. <https://doi.org/10.1007/s00426-013-0537-1>
- Toplak, M. E., Bucciarelli, S. M., Jain, U., & Tannock, R. (2008). Executive functions: Performance-based measures and the behavior rating inventory of executive function (BRIEF) in adolescents with attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*, 15(1), 53–72. <https://doi.org/10.1080/09297040802070929>
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 54(2), 131–143. <https://doi.org/10.1111/jcpp.12001>
- Trickett, J., Bernardi, M., Fahy, A., Lancaster, R., Larsen, J., Ni, Y., Suonpera, E., Wolke, D., Marlow, N., & Johnson, S. (2022). Neuropsychological abilities underpinning academic attainment in children born extremely preterm. *Child Neuropsychology*, 28(6), 746–767. <https://doi.org/10.1080/09297049.2021.2014433>
- Twilhaar, E. S., De Kieviet, J. F., Van Elburg, R. M., & Oosterlaan, J. (2020). Neurocognitive processes underlying academic difficulties in very preterm born adolescents. *Child Neuropsychology*, 26(2), 274–287. <https://doi.org/10.1080/09297049.2019.1639652>
- Twilhaar, E. S., Wade, R. M., de Kieviet, J. F., van Goudoever, J. B., van Elburg, R. M., & Oosterlaan, J. (2018). Cognitive outcomes of children born extremely or very preterm since the 1990s and associated risk factors: A meta-analysis and meta-regression. *Obstetrical & Gynecological Survey*, 73(10), 562–563. <https://doi.org/10.1097/01.ogx.0000547169.91377.70>

- Uusitalo, K., Haataja, L., Nyman, A., Lehtonen, T., Setänen, S., & The Pipari Study Group. (2021). Hammersmith infant neurological examination and long-term cognitive outcome in children born very preterm. *Developmental Medicine & Child Neurology*, 63(8), 947–953.
- van Houdt, C. A., Oosterlaan, J., van Wassenauer-Leemhuis, A. G., van Kaam, A. H., & Aarnoudse-Moens, C. S. (2019). Executive function deficits in children born preterm or at low birthweight: A meta-analysis. *Developmental Medicine & Child Neurology*, 61(9), 1015–1024. <https://doi.org/10.1111/dmcn.14213>
- Vugs, B., Hendriks, M., Cuperus, J., & Verhoeven, L. (2014). Working memory performance and executive function behaviors in young children with SLI. *Research in Developmental Disabilities*, 35(1), 62–74. <https://doi.org/10.1016/j.ridd.2013.10.022>
- Wallisch, A., Little, L. M., Dean, E., & Dunn, W. (2018). Executive function measures for children: A scoping review of ecological validity. *OTJR: Occupation, Participation & Health*, 38(1), 6–14. <https://doi.org/10.1177/1539449217727118>
- Wechsler, D. (2003). *WISC-IV, technical and interpretive manual*. Pearson Assessment.
- Wechsler, D. (2010). *WISC-IV käsikirja I ja II (WISC-IV manual I and II)*. Psykologien kustannus Oy.
- Wehrle, F. M., Kaufmann, L., Benz, L. D., Huber, R., O’Gorman, R. L., Latal, B., & Hagmann, C. F. (2016). Very preterm adolescents show impaired performance with increasing demands in executive function tasks. *Early Human Development*, 92, 37–43. <https://doi.org/10.1016/j.earlhumdev.2015.10.021>
- Woodward, L. J., Horwood, L. J., Darlow, B. A., & Bora, S. (2022). Visuospatial working memory of children and adults born very preterm and/or very low birth weight. *Pediatric Research*, 91(6), 1436–1444. <https://doi.org/10.1038/s41390-021-01869-w>

## Appendix



**Figure A1.** Flow chart of participants in PIPARI study. BRIEF = Behavioral Rating Inventory of Executive Functions.

**Table A1.** Random effects of linear mixed models predicting parent- and teacher-rated everyday WM.

Random effects	Parent-rated everyday WM <sup>a</sup>		Teacher-rated everyday WM <sup>a</sup>		
	Model 1 <sup>b</sup>	Model 2 <sup>c</sup>	Model 3 <sup>b</sup>	Model 4A <sup>c</sup>	Model 4B <sup>d</sup>
$\sigma^2$	73.43	67.51	65.76	58.09	49.16
$\tau_{00}$ Family	31.72	29.34	73.78	78.84	61.25
<i>N</i> families	127	121	129	122	120
Observations	148	139	151	140	138
Marginal R <sup>2</sup>	0.19	0.23	0.14	0.21	0.25

WM = Working Memory. <sup>a</sup>Scores in WM subscale of Behavioral Rating Inventory of Executive Functions. <sup>b</sup>Without covariates. <sup>c</sup>With covariates. <sup>d</sup>Influential cases excluded.