Parental report of vocabulary in 3- to 6-year-old Polish children: Reliable but not valid

Łuniewska-Etenkowska, Magdalena; Krysztofiak, Magdalena; Haman, Ewa

Łuniewska, M., Krysztofiak, M., & Haman, E. (2024). Parental report of vocabulary in 3- to 6-year-old Polish children: Reliable but not valid. International Journal of Language & Communication Disorders, n/a(Early View), 1–14. https://doi.org/10.1111/1460-6984.13101

DOI: https://doi.org/https://doi.org/10.1111/1460-6984.13101

Journal title: International Journal of Language and Communication Disorders
Posted at The Institutional Repository of the University of Warsaw
ReIn UW: https://repozytorium.uw.edu.pl//handle/item/160446
Unique UUID of the publication: b9d0c7de-8ba0-4563-b79c-58f3034fde57

Received: 27 October 2023 | Accepted: 26 July 2024

DOI: 10.1111/1460-6984.13101

International Journal of Language & Communication Disorders

RESEARCH REPORT

Parental report of vocabulary in 3- to 6-year-old Polish children: Reliable but not valid

Magdalena Łuniewska | **Magdalena Krysztofiak** | **Ewa Haman**

Faculty of Psychology, University of Warsaw, Warszawa, Poland

Correspondence
Magdalena Łuniewska, Faculty of Psychology, University of Warsaw, Stawki 5/7, 00–183, Warszawa, Poland.
Email: magdalena.luniewska@psych.uw.edu.pl

Abstract

Background: For over 30 years, parental reports have been used to study the vocabulary of children under 4 years of age. Research exploring parental checklists as a measure of vocabulary in older children is very limited. Typically, authors of parental checklists report the reliability of the developed tools but do not explore validity in terms of the agreement between parental assessments and the children's actual word knowledge.

Aims: We aimed to explore the reliability and validity of a parental checklist for assessing vocabulary in children aged between 3 and 6 years. Furthermore, we aimed to evaluate the agreement between indirect (parental checklist) and direct (picture naming and picture recognition tasks) assessments of children's vocabulary.

Methods and Procedures: A group of 94 typically developing monolingual Polish-speaking children aged between 3 and 6 years were first directly tested onsite with picture naming and picture recognition tasks (Cross-Linguistic Lexical Tasks). Subsequently, the participants' parents completed an online checklist containing the same set of 128 items and marked all the words that they had ever heard in their child's spontaneous speech.

Outcomes and Results: The parental checklist demonstrated very high internal consistency. The scores of the parental checklist and vocabulary tasks were moderately correlated. We compared the total number of words marked by parents and the number of items correctly identified by children in the picture naming and picture recognition tasks. In picture naming, we found no difference between the children's scores and the number of words selected by parents. However, parents selected significantly fewer words than children correctly recognised in the picture recognition task. When data were analysed at the level of individual items (i.e., whether parents selected exactly the same items that children answered correctly), we found that the level of agreement was low. The level of agreement correlated negatively with the children's vocabulary; that

is, the more words a child knew, the lower the agreement between the direct measure and the parental checklist.

Conclusions and Implications: Parental checklists should be used with caution in children aged between 3 and 6 years, especially if the assessed children have a large vocabulary and if item analysis is planned. Such checklists may be of more use in younger children or in children with limited vocabulary.

KEYWORDS
assessment, children, parents, preschool, psychometrics, vocabulary

WHAT THIS PAPER ADDS

What is already known on the subject

- Parental checklists are commonly used to assess the vocabulary of children younger than 4 years of age. Previous research has indicated that parental checklists are reliable in terms of internal consistency and valid in terms of predictive and convergent validity.

What this paper adds to the existing knowledge

- This study introduces a parental checklist designed for assessing the vocabulary of monolingual Polish-speaking children aged between 3 and 6 years. Statistical analyses reveal that while the parental checklist exhibits high reliability, and the scores on the checklist correlate with direct measures of vocabulary, the agreement between parental reports and direct vocabulary measures (i.e., validity) is notably low, particularly when examining individual test items.

What are the clinical implications of this work?

- These findings underscore the importance of exercising caution when using parental vocabulary checklists with children aged between 3 and 6 years. These checklists can serve as a replacement for direct vocabulary tests only when the general/overall score is needed. However, when specific words are the subject of interest, parental reports may not be a valid measure.

INTRODUCTION

For over three decades, parental reports have served as a reliable method for studying the vocabulary of children under 4 years (Marchman & Dale, 2023). However, research examining parental reports as a measure of vocabulary in children older than 4 years remains limited. The most widely used tool for assessing early vocabulary, available in over 100 language versions, is the MacArthur-Bates Communicative Development Inventory (MB-CDI; Marchman et al., 2023). These inventories are typically designed for children aged from 8 months to 30−36 months (Jarůšková et al., 2023; Marchman & Dale, 2023). While a version of the MB-CDI for older children, that is, 3-year-olds, was developed 2 decades ago (CDI-III; Feldman et al., 2005), it gained popularity only recently with the introduction of new language versions (see, Table 1). These versions of the CDI-III have consistently demonstrated excellent reliability, typically characterised by very high internal consistency (Cadime et al., 2021; Eriksson, 2017; Kas et al., 2022; Tulviste & Schults, 2020). Moreover, some language versions of the CDI-III also include data on the tools' validity (Cadime et al., 2021; Eriksson, 2017; Tulviste & Schults, 2020).

TABLE 1 The recently developed CDI-III versions with age ranges.

Language	Age range (years; months)	Authors
Basque	2;6 to 4;2	Garcia et al. (2014)
Croatian	2;6 to 4;0	Šmit Brleković and Kraljević (2023)
Estonian	2;10 to 3;3	Tulviste and Schults (2020)
Hungarian	2;0 to 4;2	Kas et al. (2022)
Norwegian	2;6 to 4;0	Holm et al. (2023)
Persian	2;6 to 3;1	Mokhtari et al. (2022)
Portuguese	2;6 to 4;0	Cadime et al. (2021)
Spanish	2;6 to 4;0	Jackson-Maldonado et al. (2022)
Swedish	2;6 to 4;0	Eriksson (2017)

Reliability and validity of parental reports

Reliability

The reported reliability of parental reports typically refers to the concept of internal consistency. In publications of parental vocabulary checklists, the most commonly applied measures of internal consistency include Cronbach's α (Eriksson 2017; Holm et al., 2023; Tulviste & Schults, 2020), Kuder-Richardson 20 score (Cadime et al., 2021) and the split-half method (Šmit Brleković and Kraljević, 2023). Considered together, these different measures of internal consistency indicate the individual items in parental reports of vocabulary measure the same underlying construct.

Predictive validity

The validity of parental reports is typically reported in terms of predictive and/or convergent validity. Predictive validity informs us whether, based on the scores of the parental report, we may "predict" a child's score in another measurement of, for instance, reading skills, after a given time interval. For example, expressive vocabulary size at the age of 2 years (as measured by parental reports) can be used to predict subsequent language and literacy achievements up to fifth grade (Lee, 2011). Similarly, the decoding skills and reading comprehension of sixth graders can be predicted based on the child's CDI score when they are 16 to 30 months old (Bleses et al., 2016). Recently, the educational achievements of 15-year-olds have been shown to be predicted by expressive vocabulary measured at the age of 16 to 30 months (Dale et al., 2023).

Convergent validity

Convergent validity, on the other hand, evaluates the validity of parental reports by correlating the scores reported by parents with an independent measure, often administered by a specialist using a different tool. This approach is referred to as convergent or congruent validity. Assessing convergent validity is generally more straightforward than predictive validity, as it involves measurements at a single time point. Consequently, there is a substantial body of research demonstrating the convergent validity of parental reports on vocabulary.

For example, in children aged 16 to 21 months, CDI scores have been correlated with a direct measure of receptive vocabulary (Friend et al., 2012). In children aged 2.5 to 4 years, the total vocabulary scores in CDI-III were found to be correlated with scores on the language subscale of the Griffiths Mental Development Scale (Cadime et al., 2021). Likewise, scores on the Developmental Vocabulary Assessment for Parents (DVAP), a parental checklist based on items from the Peabody Picture Vocabulary Test 4 (PPVT-4; Libertus et al., 2015), were correlated with raw PPVT-4 scores in children aged 2;10 to 6;2 ($r = 0.69$). In summary, the number of words selected by parents as produced or understood by their children is consistently reported to correlate with standardised tests of receptive or expressive vocabulary.

Why might correlations not be enough?

Studies on the validity of parental reports indicate that there is a positive correlation between the number of words a parent indicates and the number of pictures a child names or recognises in standardised tests. However, this correlation alone does not provide a complete picture of the agreement between parental reports and direct, test-based measures of vocabulary. In essence, we lack insight into whether the specific words selected from a list by parents accurately represent the words the child actually knows. To put it differently, we lack a comprehensive understanding of the level of agreement between parental reports and alternative vocabulary assessment methods.

So far, the only study that has delved into the agreement between parental assessments of children's vocabulary and children's actual performance on language tasks using the same set of words was conducted over 2 decades ago (Ring & Fenson, 2000) and as far as we are aware, there have been no replication studies of these findings. In this particular study, parents of 40 children aged 19 to 31 months were first asked to complete a CDI: Words and Sentences to assess the child's total vocabulary and then to evaluate two sets of pictures. For one set of 35 items, parents were asked "whether they believed their child comprehended each target word" (Ring & Fenson, 2000, p. 145). For the other set of 35 pictures, parents were asked to "indicate which pictures their child would be able to label verbally" (Ring & Fenson, 2000, p. 145). Following the parental assessments, children were directly tested using a picture-book technique, employing the very same items as those judged by parents. This included the same pairs of pictures for the recognition task and the same set of pictures for the production task. The results from this study revealed significant correlations between parental assessments and children's performance, with a correlation coefficient of 0.55 for comprehension and 0.67 for production. Interestingly, the strongest relationship was found between the children's production score and the total score (the number of words that the parents marked as spoken by the children out of the 680 items on the list) on the CDI: Words and Sentences completed before the testing session. This finding may indicate that when given a longer list of words, parents can more accurately estimate children's overall vocabulary, even though they may not be able to report the specific words that children know.

What is crucial from the perspective of agreement between parental reports and other measures of language skills is that, despite a significant correlation between the two, parents tended to overestimate their children's accuracy in both comprehension and production tasks. The accuracy estimated by parents exceeded children's actual performance by approximately 10% (in comprehension) to 20% (in production). These findings suggest that while there is a moderate correlation between parental assessments and children's performance (indicating convergent validity), there may be a lack of precise agreement between the two measures. This lack of agreement is particularly noteworthy in the context of analyses of individual word knowledge as reported by parents and aggregated in WordBank (Frank et al., 2021). Specifically, the potential lack of concordance between parental reports and the actual set of words children know could impact experiments that rely on parental reports to indicate that children know specific words.

Parental reports at preschool age

To date, only one tool has been developed for parental reporting of the vocabulary of children older than 4 years—specifically those aged 2 to 7 years (Libertus et al., 2015). This tool is known as the DVAP. The DVAP was created based on the first 204 items of the PPVT-4 (Dunn, & Dunn, 2007), encompassing nouns, verbs, and adjectives arranged in order of increasing difficulty. Parents completing the DVAP are tasked with indicating the words they have ever heard their children use. The total score is the number of words selected by the parents. DVAP scores have been shown to correlate with scores from the CDI-III (in children aged 2;2 to 4 years; $r = 0.79$) and PPVT-4 (Libertus et al., 2015).

Recently, the DVAP has been employed in various studies in which the vocabulary size of preschool-age children served as a control variable (Elliott et al., 2022). Additionally, a similar parental checklist, based on the French version of the PPVT-4, has been developed for the French language (Byers-Heinlein et al., 2022). Although DVAP results are correlated strongly with children's scores on PPVT-4, the authors did not compare the average scores on the two tools using statistical analyses (Libertus et al., 2015), nor did they analyse the differences between the children's scores and the parental reports at the level of the individual words included in the two tools.

In summary, parental reports can serve as a valuable source of information about the vocabulary size of children older than 3 to 4 years; however, the agreement between the parental report and the child's performance in a vocabulary task may not go down to the single-item level. Moreover, to date, no study has undertaken an in-depth analysis of the accuracy of parents in selecting specific words known to their children over the age of 3 years.

Current study

In this study, we assess the psychometric properties of a parental checklist based on items from a standardised vocabulary test known as the Cross-Linguistic Lexical Task (CLT-PL; Haman et al., 2015). Specifically, we focus on examining the reliability and validity of the parental checklist, as well as verifying the agreement between direct vocabulary assessment using the CLT-PL and indirect assessment via the parental checklist.

Our study was pre-registered with the Open Science Framework (OSF), and our hypotheses were initially formulated in the preregistration, which can be accessed here: https://osf.io/tusm6.

1. Reliability of the parental checklist: we anticipate that the parental checklist will demonstrate high reliability, as indicated by high Cronbach's α and McDonald's ω coefficients, as well as strong split-half reliability.
2. Validity of the parental checklist: we expect the total scores (i.e., the number of words selected by parents) on the checklist to significantly correlate with standardised measures of vocabulary. This expectation aligns with previous studies demonstrating convergent validity of parental reports (Cadime et al., 2021; Libertus et al., 2015; Smolík & Bytešníková, 2021). Given that parents are asked to select words produced by children in spontaneous speech, we anticipate stronger correlations for production tasks compared to comprehension tasks. This pattern is consistent with findings in younger children (Ring & Fenson, 2000).
3. Agreement between methods: we also expect to observe agreement between children's accuracy on the CLT-PL and the parental checklist, both at the participant level and the individual item level. Notably, we predict higher agreement for production items than for the items of the comprehension tasks, primarily due to the specific wording used in the parental checklist. However, it is important to note that this expectation contrasts with previous data from younger children, where parents were more accurate in predicting scores for comprehension than for production (Ring & Fenson, 2000).

METHOD

Participants

The families recruited for this study initially registered to a larger PolkaNorski project, which includes a norming study of the Polish CLT-PL. The CLT-PL norming study in the PolkaNorski project received approval from the Research Ethics Committee at the Faculty of Psychology, University of Warsaw. Independently, the parental checklist study (the focus of the current paper) received approval from the Rector's Committee for the Ethics of Research Involving Human Participants at the University of Warsaw.

We analyse data from 94 children, ranging in age from 2.97 to 6.15 years (M = 4.79, SD = 0.84, Me = 4.88). The group consisted of 47 girls and 47 boys, all attending preschools in 10 different towns across four regions of Poland. The sample size for this study was determined through power analysis, as outlined in our preregistration: https://osf.io/tusm6.

Participants were recruited through posters and leaflets in preschools. Before entering the PolkaNorski project, parents gave written informed consent and children gave oral consent. Children were tested in preschools and received books and certificates as tokens of appreciation. After the testing sessions, the parents who agreed to further contact and whose children fulfilled the participation criteria for the parental checklist study (discussed later) were sent email invitations to the parental checklist. The parents consented to participate in the parental checklist study (the focus of the current paper) independently of the PolkaNorski project. Upon completion of the parental checklist, the parents received a voucher for an online bookstore.

We sent email invitations to parents of all children who were tested between October 2022 and March 2023 and met the following criteria: the child was a monolingual Polish speaker and completed the CLT-PL assessment; the parents had previously filled out all required questionnaires; and the parents had agreed to be contacted for recruitment to research projects (N = 142). We ceased inviting parents after reaching the 100th participant. The original sample consisted of 103 children, as the parents of 39 children did not respond to the invitation. Data from nine children were excluded due to data loss in one of the CLT-PL tasks (n = 8) or the parental checklist (n = 1).

Regarding the children's language environment, the parents of eight children reported that their children had some exposure to English, primarily in preschool during short English lessons (none of the children attended a bilingual preschool) or while watching television, listening to children's songs or playing on smartphones or tablets. The remaining 86 children were reported to not have regular contact with languages other than Polish. None of the parents reported any serious health or neurodevelopmental disorders in their children.

The sample of parents consisted of 84 mothers and 6 fathers (the number of parents was lower than the number of children as 86 caregivers completed reports about one child, while four mothers provided data for two children, i.e., younger and older siblings). Caregivers in our study were highly educated: 73% had a university degree, 8% started education at a university and 19% had completed high school.

MATERIALS AND PROCEDURES

Procedure

The procedure consisted of three steps. First, parents filled in an online demographic and language environment questionnaire (Mieszkowska et al., 2021) and a Polish revised adaptation of the Parents of Bilingual Children Questionnaire (PaBiQ; Tuller, 2015). Second, children were individually tested with the CLT-PL in a preschool (as

a part of the CLT-PL norming study in the PolkaNorski project). Third, parents (one parent per child) of the children completed the parental checklist online (the study we focus on in this manuscript).

Background questionnaire

The PaBiQ questionnaire collects information about the child's language environment (e.g., the number of languages used with the child, the frequency of their use, contexts in which each language is used, and the child's skills as reported by parents), and general information about the child and their family (e.g., date of birth, early development, parental education and difficulties at school experienced by family members). In the current project, we used data from PABiQ related to parental education, the children's development and health and language environment.

Direct assessment of vocabulary: CLT-PL

To directly assess the children's vocabulary, we used the Polish version of the LITMUS (Haman et al., 2015). CLT-PL comprises both word comprehension (picture recognition) and word production (picture naming) subtasks. In the comprehension subtasks, children hear prerecorded questions (e.g., 'Where is the flag?' in the noun comprehension task and 'Who is swimming?' in the verb comprehension task) and select one of four pictures displayed on a tablet by touching the screen. In the production subtasks, children view a single picture and respond to prerecorded questions (e.g., 'What is this?' in the noun production task and 'What is she doing?' in the verb production task). Each CLT-PL subtask (noun comprehension, verb comprehension, noun production, verb production) consists of 32 items, resulting in a score range of 0–64 for both receptive and expressive vocabulary, totalling 128 items.

The target words in the comprehension and production subtasks of the CLT-PL were carefully selected, following the procedure for creating CLTs (Haman et al., 2015), based on the age of acquisition (AoA) of the words (Łuniewska et al., 2016) and their morphological and phonological complexity (Haman et al., 2015; Hansen et al., 2017; van Wonderen & Unsworth, 2021). One of the steps in creating the CLT-PL was to match the words that would be used as the targets in the production and the comprehension subtasks into pairs of the same semantic category and similar AoA and morpho-phonological complexity (Haman et al., 2015). From each pair of words, one word was used as the target in the comprehension subtask, and the other served as the target in the production task. This step is taken to ensure a comparable level of difficulty of the words used as targets in the comprehension and production subtasks. Thanks to the systematic selection of words, the target words in the comprehension and production tasks did not differ in terms of AoA ($t(125) = 1.02$, $p = 0.32$), complexity index ($t(124.1) = 0.71$, $p = 0.48$), or imageability ($t(125.7) = 0.25$, $p = 0.80$). The complete list of CLT-PL items and the psycholinguistics characteristics of these words (values of AoA, complexity index and imageability; Wolna et al., 2023) are available from the OSF archive: https://osf.io/tyw3g/.

The CLT-PL testing was run according to the manual for the norming study, and the procedure was very close to previous studies (Hansen et al., 2019; Łuniewska et al., 2022). We administered the CLT-PL using a mobile app (Child Lexicon CLT-PL app; Sobota et al. 2022), compatible with Android tablets and iPads, the links to the app are shared in the OSF archive: https://osf.io/tyw3g/. Testing occurred individually in a quiet room within the children's preschools. The testing sequence began with comprehension (nouns followed by verbs) and then moved to production tasks (nouns followed by verbs). In the comprehension part, children heard prerecorded questions, followed by a brief 100 ms display of a blank screen before viewing four pictures on the tablet and making their selection. In the production part, pictures were presented sequentially, each accompanied by a relevant question, and the experimenter tapped the screen to proceed to the next item. The entire CLT-PL testing session typically lasted about 15–20 min. Parts with production tasks were audio recorded to enable offline scoring of the children's answers.

Indirect assessment of vocabulary: Parental checklist

An indirect assessment of the child's vocabulary was conducted using a parental checklist based on the CLT-PL. After the child completed the CLT-PL testing in preschool, we invited one parent (the one who registered the child for the study, typically the mother) to complete the online checklist. We sent an individually generated Qualtrics survey link via email. The checklist included two demographic questions regarding the parent's relationship to the child and the child's date of birth. These questions were followed by questions about the child's vocabulary.

In particular, we asked parents to select from a list all the words that the parent had ever heard in spontaneous speech from their child (the exact instruction in Polish was as follows: 'Na poniższej liście prosimy o zaznaczenie tylko tych słów, które kiedykolwiek słyszeli Państwo w spontanicznych wypowiedziach Państwa dziecka'). The checklist,

International Journal of Language & Communication Disorders

containing all 128 CLT-PL items, presented words in a randomly generated order for each participant. This order included a mix of nouns and verbs. Completing the checklist typically took between 4 and 7 min ($M = 6.0$, SD = 2.8, $Me = 5.3$), and the time interval between the direct and indirect assessments varied from 0 to 27 days ($M = 6.6$, SD = 5.0, $Me = 6.0$).

DATA PREPARATION AND ANALYSES

CLT-PL scoring

We coded the accuracy for the picture choice and picture naming tasks. In the picture choice tasks, accuracy was automatically scored using the app. Children received 1 point for selecting the correct picture and 0 points for choosing any other picture (or none).

For the picture naming tasks, we implemented strict accuracy scoring criteria, following the approach used in prior research with the CLT-PL (Haman et al., 2017; Łuniewska et al., 2022). Only responses that included a form of the target word, such as inflections, derivations or mispronunciations, were considered correct. Any other responses, including synonyms, regional variations or definitions, were categorised as incorrect. We employed these strict criteria to assess whether the child knew a particular word form (rather than a synonymical/regional word), as our parental assessment focused on specific word forms used by the child in speech. However, synonymic responses were exceptionally rare in the collected data, accounting for only 0.6% of responses in the production tasks. For each child, we calculated three CLT-PL scores: one based on overall accuracy across all 128 items, one for production (64 items) and one for comprehension (64 items).

Plan of the analyses

First, to assess the parental checklist's reliability, we employed two measures of internal consistency: Cronbach's α and McDonald's ω. Additionally, we calculated split-half reliability by randomly dividing the items into two halves.

Second, in evaluating the validity of the parental checklist, we examined non-parametric Spearman correlations among CLT-PL scores (total scores and scores for production and comprehension tasks), parental checklist scores and the child's age. In addition to the pre-registered correlations, we calculated partial correlations between CLT-PL scores and the parental checklist with the child's age partialled out.

Third, we analysed the agreement between the direct and indirect vocabulary measures on three levels: group, participant and item. We began by comparing total scores between the two measures using paired *t* tests, considering the entire set of items as well as subsets for comprehension and production items. Next, we estimated intraclass correlation coefficients (ICCs) to assess agreement at the individual participant level. This involved calculating three separate ICCs in two-way mixed models for the entire dataset, as well as for production and comprehension items separately. Finally, we evaluated the item-level agreement between parental reports and child accuracy for each participant by computing three Cohen's kappas: one for the complete test and one for the comprehension and production parts. All analyses were conducted using R, and the data analysis scripts can be found in the OSF archive: https://osf.io/tyw3g/.

RESULTS

Reliability analyses

The internal consistency of the parental checklist was high: Cronbach's $\alpha = 0.99$, McDonald's $\omega_h = 0.81$ and $\omega_t = 0.99$. Similarly, the calculated split-half reliability for randomly divided items of the checklist indicated high reliability $r_{SB} = 0.96$.

Correlation analyses

The CLT-PL scores were moderately correlated with the children's ages (rho between 0.59 and 0.64; Table 2, Figure 1a). The scores of the parental checklist were moderately correlated with the CLT-PL scores (rho between 0.32 and 0.38; Table 2, Figure 1b) but not correlated with participants' age (rho <0.15). The partial correlations between CLT-PL scores and the scores on the parental checklist with age partialled out were moderate as well (rho between 0.30 and 0.41).

Agreement

Group level

In general, in the parental checklist, parents selected significantly fewer words ($M = 100.60$) than the children answered correctly during the direct assessment with CLT-PL ($M = 110.41$), as revealed with a series of paired *t* tests (Table 3 and Figure 2). When the comparison was limited to the words that were target items in the

TABLE 2 The correlations between the participants' age, score on the direct assessment of vocabulary with CLT-PL and score on the indirect assessment of vocabulary with the parental checklist.

		CLT-PL			Parental checklist		
		Total	**Production**	**Comprehension**	**Total**	**Production**	**Comprehension**
		correlations (rho values)			correlations (rho values)		
	Age	0.64***	0.60***	0.59***	0.13	0.09	0.14
CLT-PL	**Total**				0.38***	0.36***	0.38***
	Production				0.38***	0.36***	0.38***
	Comprehension				0.33**	0.32**	0.33**
		Partial correlations with age partialled out					
Parental Checklist	**Total**	0.41***	0.39***	0.32**			
	Production	0.40***	0.38***	0.34**			
	Comprehension	0.38***	0.38**	0.30**			

Note: Below diagonal are partial correlations between the measures with age partialled out.
**$p < 0.01$, ***$p < 0.001$ Abbreviation: CLT-PT, Cross-Linguistic Lexical Task.

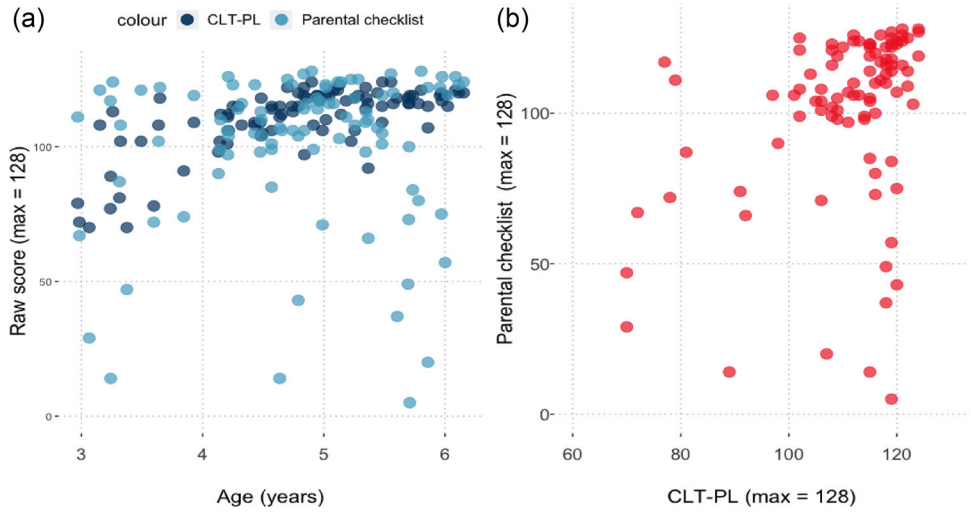

FIGURE 1 The relationships between participants' ages and scores in the direct (CLT-PL) and indirect (parental checklist) assessments of overall vocabulary (a), and between the two vocabulary measures (b). Abbreviation: CLT-PL, Cross-Linguistic Lexical Task.

production task, there was no significant difference between the number of words selected by parents (M = 51.93) and the number of pictures named correctly by the children (M = 50.68). Conversely, when only the comprehension items were taken into account, parents selected significantly fewer (M = 48.67) words than children solved correctly (M = 59.73).

Participant level

Next, we calculated the ICCs to estimate the agreement between the indirect and direct assessments at the level of individual participants. Overall, the agreement was poor, as we obtained the following ICCs: for the whole set of items — ICC = 0.41, p = 0.006 (95% confidence interval [CI]: 0.11−0.61); for production items — ICC = 0.46, p = 0.002 (95% CI: 0.19−0.64); and for comprehension items — ICC = 0.30, p = 0.045 (95% CI: −0.06−0.53).

Item level

The values of the calculated Cohen's kappa coefficients suggested only a slight agreement between the direct and indirect assessments of vocabulary (Table 4). The maximal calculated kappa coefficients were all below 0.60, indicating that agreement on the item level was low or moderate, even in the case of the most accurate parents.

Given the low values of Cohen's kappas, we performed additional exploratory analyses (not included in the preregistration) to check whether the level of agreement

TABLE 3 The scores in the direct (CLT-PL) and indirect (parental checklist) assessments of vocabulary for total tasks (128 items) and divided into production (64 items) and comprehension (64 items) parts; mean (SD) and score ranges are reported.

	CLT-PL score	Parental checklist score	Comparison
Total	110.41 (12.49) 70–124	100.60 (28.82) 5–128	$t(93) = 3.51$, $p < 0.001$
Production	50.68 (8.23) 22–61	51.93 (13.99) 3–64	$t(93) = -0.89$, $p = 0.38$
Comprehension	59.73 (4.81) 40–64	48.67 (14.97) 2–64	$t(93) = 7.51$, $p < 0.001$

Abbreviation: CLT-PL, Cross-Linguistic Lexical Task.

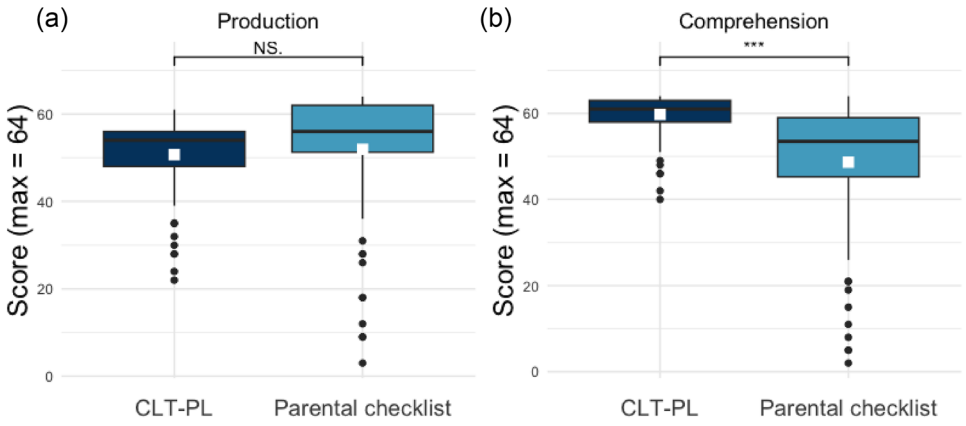

FIGURE 2 The comparison of the scores from the direct (CLT-PL) and indirect (parental checklist) assessments of vocabulary, separately for the items of the production (a) and the comprehension tasks (b). Abbreviation: CLT-PL, Cross-Linguistic Lexical Task.

TABLE 4 The values (mean for all participants and range) of Cohen's kappa as calculated for the whole set of items, and for production and comprehension items separately.

All participants ($N = 94$)	Total	Production	Comprehension
Mean	0.13	0.18	0.10
Range	−0.11 to 0.42	−0.11 to 0.55	−0.11 to 0.56

between the direct and indirect assessments was related to the children's ages or vocabulary scores (Table 5 and Figure 3). We found a pattern of moderate negative correlations between the level of agreement and the participants' scores and vocabulary: specifically, the greater the participant's age and the larger their vocabulary, the lower the agreement between the scores of the direct and indirect assessments of vocabulary.

DISCUSSION

Our study aimed to present a parental checklist for vocabulary assessment in children between 3 and 6 years of age, along with its psychometric properties. The presented parental checklist was constructed based on exactly the same set of items used in the Polish version of a standardised tool for receptive and expressive vocabulary assessment in children, namely the CLT-PL (Haman et al., 2015). It is worth noting that, especially since the COVID-19 pandemic, a significant portion of psycholinguistic research has shifted from preschools and labs to virtual reality. Recently, parental reports of vocabulary have been extensively utilised. For example, a vocabulary checklist partly based on the CLT has been employed as a measure of vocabulary (McGKaillion et al., 2023) without providing information about the tool's reliability or validity.

We found that the presented parental checklist exhibited high reliability in terms of internal consistency, similar to other parental reports of vocabulary in children older than 2.5 years (Šmit Brleković and Kraljević, 2023; Cadime et al.,

TABLE 5 Spearman correlations between the level of agreement between direct and indirect assessment of vocabulary, children's age and vocabulary size.

	Age	CLT-PL total score	Parental checklist total score
Level of agreement (Cohen's kappa)	−0.23*	−0.40***	−0.25*

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Abbreviation: CLT-PL, Cross-Linguistic Lexical Task.

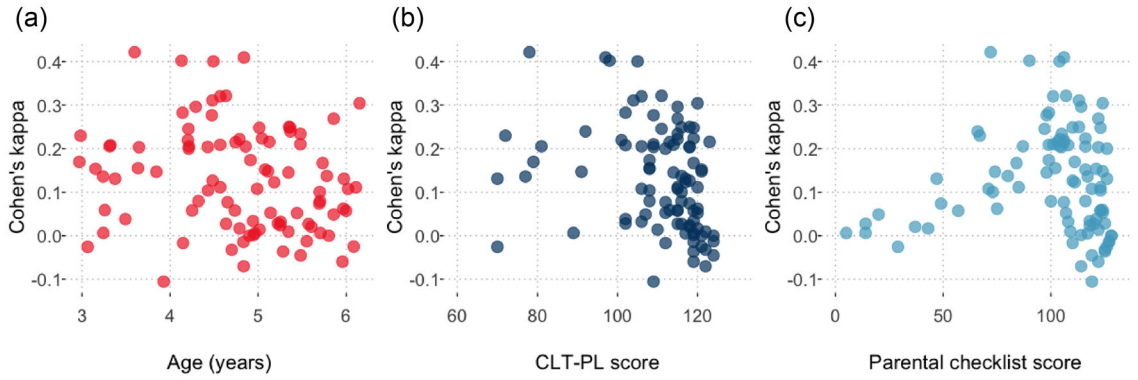

FIGURE 3 The relationships between the level of agreement of the direct and indirect assessment scores (Cohen's kappa) in terms of participants' ages (a) and vocabulary sizes (b) and (c). Abbreviation: CLT-PL, Cross-Linguistic Lexical Task.

2021; Eriksson, 2017; Holm et al., 2023; Kas et al., 2022). The scores of the parental checklist were moderately correlated with the scores of the standardised vocabulary assessment using the CLT-PL. Although the obtained correlation coefficients may appear low at first glance, they are similar in magnitude to correlations reported in previous studies that employed parental reports of vocabulary in children older than 2.5 years. For instance, Smolík and Bytešníková introduced a new parental vocabulary questionnaire for Czech-speaking children aged 15 to 42 months (the Stručný Dotazník Dětského Slovníku [Short Questionnaire of Children's Vocabulary]; SDDS), and validated the tool against the Czech version of the CLT. They obtained correlation coefficients between SDDS and CLT scores ranging from 0.32 to 0.53 (Smolík & Bytešníková, 2021, p. 6). Similarly, when validating the European Portuguese CDI-III, the authors calculated correlations with scores from a standardised language test, reporting coefficients between 0.36 and 0.53, depending on the specific vocabulary list included in the analysis (Cadime et al., 2021). Additionally, when analysing the validity of the DVAP, a parental checklist based on a subset of PPVT-4 items, by correlating the scores of the tool with PPVT-4 scores, Libertus and colleagues reported a standardised regression coefficient of 0.42 (Libertus et al., 2015), which is close to the value obtained in our study. This is particularly interesting as, similarly to our study, Libertus and colleagues used overlapping sets of items in both direct and indirect assessments of vocabulary. Nevertheless, the correlation between the two measures was of moderate size. Interestingly, the authors of all of the aforementioned studies were satisfied with the obtained correlation scores and argued for the validity of their tools.

Typically, the introduction of a new parental checklist concludes at this point: if the tool demonstrates reliability in terms of internal consistency and shows correlations with another measure of vocabulary, it is considered reliable and valid. Such a tool may then proceed to a norming study or be directly applied in research. Our study suggests that the image of using parental checklists to measure vocabulary in children older than 3 years is more complicated.

First, we observed that the total score on the parental checklist (the number of words selected by parents as being part of their child's vocabulary) was lower than the number of words understood and correctly identified by children in direct assessment with the CLT-PL. This discrepancy primarily affected the comprehension items. In the case of the production items, the overall number of words selected by parents was not different from the overall number of pictures correctly named by the children. The difference between the children's scores in the comprehension task and the number of words marked by parents as part of their child's spontaneous speech can be explained in several ways. One possibility is that the children knew more words receptively than expressively; they understood and correctly identified some words in the comprehension task but did not use them in spontaneous speech. This asymmetry between expressive and receptive vocabulary, both in children and adults, has been

well-documented in research (Gershkoff-Stowe & Hahn, 2013). Alternatively, it is possible that participants responded by guessing when they did not know certain words. Since there are four pictures on each picture board, guessing could potentially increase scores by 25%, equivalent to 16 words (given a maximum score of 64). In reality, the difference between the number of words selected by parents and the number of correctly chosen pictures in the comprehension task was smaller, around 11 words. This suggests that parents might even overestimate their children's vocabulary scores, which aligns with previously reported results (Ring & Fenson, 2000). Finally, the absence of a significant difference between the direct and indirect assessments of expressive vocabulary at the group level may suggest that, for obtaining a general insight into children's vocabulary at a certain age, parents can serve as a reliable source of information. For instance, if a researcher needs to ascertain how many words from a specific set are typically known by preschool-age children, consulting parents of preschoolers may yield valid data.

However, as we delve deeper from the group level to individual participants and items, the image becomes more complex. While at the group level, the number of words selected by parents aligns with the number of items correctly named by children, at the individual participant level, there is low agreement between the two measurements. Through the analysis of ICCs, we uncovered poor agreement between the direct and indirect assessments of vocabulary. This poor agreement indicates that the parental checklist and CLT-PL should not be used interchangeably at least for children in the age range of 3 to 6 years. It also underscores the need for caution when comparing scores from direct and indirect vocabulary measurements. In the past, indirect parental measures have sometimes been used in longitudinal studies examining patterns of associations between vocabulary size and other variables at different ages. In many of these studies, children's vocabulary size was measured by parental report at an earlier stage (usually under the age of 3), while direct measures of other variables were used at a later time (which could be several or even a dozen years after the first measurement; e.g., Bleses et al., 2016; Dale et al., 2023; Lee, 2011). It is a matter for further investigation (or reanalysis of previous data) whether in such cases agreement between the two measurements is present at both group and individual levels, or only at the group level, as in our study.

In another longitudinal study both indirect and direct measures of vocabulary size were obtained when children were younger than 3 years of age (16 months old at the first time point and 22 months old at the second one; Legacy et al., 2018). The authors assumed that the two measures reflected the same latent variable. While our results suggest that such an assumption may be misleading for children above the age of 3 years, it can be that for younger children (below the age of 3 or even 2 years) the indirect and direct measures are more accordant with each other also at the individual participant level due to the overall smaller children's vocabulary size. Again, although indirect measures are so far considered to be valid for toddlers and children under 3 years of age, in light of our findings and for future studies we suggest going beyond group-level comparisons and assessing agreement at the individual participant and/or item level as well. This may be of particular importance when the aim of the research is not only to estimate the overall vocabulary size but also to characterise the content of children's lexicons.

In our study, the agreement between the direct and indirect assessments was notably poor when individual items were analysed. At the group level, parents generally selected the same number of words from the list as the children knew (either words they could name or understand); however, during direct measurement with a standardised tool, the words chosen by parents did not match those correctly identified by the children. The overall agreement between parental word selection and the children's accuracy was only slightly above the level expected by chance. Even the responses of the most accurate parents exhibited only moderate agreement with the results of the direct assessment.

The low agreement between the two tools could be attributed to the possibility that children above the age of 3 years already knew and used too many words for parents to accurately remember all of them. This is supported by the observation that the level of agreement between parental word selection as used in the children's spontaneous speech and children's accuracy was negatively correlated with the child's age and vocabulary. In other words, the older the child and the more extensive their vocabulary, the lower the agreement between the direct and indirect assessments. This decline in agreement with age and vocabulary suggests that the parental checklist may be more useful for younger children than for older ones. A similar decrease in the amount of variance explained in the direct assessment of vocabulary by a parental checklist was observed for the DVAP (Libertus et al., 2015). These findings suggest that as children's vocabulary grows, parents may not be fully aware of all the words their children use or understand, limiting their role as a source of information about their children's vocabulary.

The limited accuracy of parents in selecting words known by children has significant implications for research in which certain words are assumed to be known or unknown to children. Studies that, based on WordBank data, assume that certain words have already been acquired by a certain proportion of children (Conboy &

Mills, 2006; Kuhl et al., 2013) may need to incorporate additional measures to validate these assumptions. Similarly, this potential limitation of parental reports may have significant implications for speech and language therapy practice. While questionnaires such as the CDI may be useful for screening purposes (where only a total score is needed), speech-language pathologists should be cautioned that using the exact words reported by parents on the questionnaire to characterise the child's lexicon or to plan therapy may not be useful because we cannot determine how accurately parents report the exact words a child understands or produces.

Nonetheless, it is important to highlight that our study is the first to investigate the agreement rates between parental word selection and actual word knowledge in the same group of children. While a dataset for a similar analysis was collected during the validation of the DVAP, the authors of the DVAP did not confirm whether the words selected by caregivers in the tool corresponded to the items that children solved in the PPVT-4 (Libertus et al., 2015). Partial evidence suggests that the overlap was not complete, as, on average, caregivers reported that children said 93 words from the DVAP (out of a total of 204), while in direct assessment, the average score for the same children was 85 words (Libertus et al., 2015, p. 8). The authors did not provide an explanation for this difference. The low agreement between direct and indirect assessments of vocabulary in children aged 3 to 6 years could also motivate further studies focusing on younger children. To date, only one study has explored whether the accuracy of toddlers (2-year-olds) in picture naming and picture recognition tasks corresponds to parental reports (Ring & Fenson, 2000). While the results of this study were promising and supported the use of parental reports as a measure of vocabulary, it has not been replicated.

SUMMARY

In summary, we introduce a new parental checklist for assessing vocabulary in children aged 3 to 6 years. The analysis of reliability and validity demonstrates that the tool exhibits strong psychometric properties, similar to other parental reports of vocabulary in children older than 2.5 years of age. Based on these psychometric properties, we conclude that the presented parental checklist may serve as a valuable screening tool, particularly for younger children or those with limited vocabulary. This may be especially useful for children at risk of language problems or disorders and/or when parents are anxious about their children's slow development of lexical knowledge. However, the low agreement rates between direct and indirect measures of vocabulary, especially at the level of individual items, indicate that parents may no longer be a reliable source of information regarding specific word knowledge in children older than 3 years of age. Therefore, when experiments are based on the assumption that specific words are known by children, alternative sources of data on word knowledge should be employed.

ACKNOWLEDGEMENTS

The direct assessment of children's vocabulary with the CLT-PL was possible thanks to a grant from the Norwegian Financial Mechanism for 2014–2021 via the National Science Centre, project no 2019/34/H/HS6/00615 in the GRIEG programme. The development and usage of the parental questionnaire were possible thanks to an internal grant of the University of Warsaw POB IV for Parents—New Ideas awarded to Magdalena Łuniewska and funded by subvention from the Polish Ministry of Science and Higher Education in the Excellence Initiative—Research University (2020–2026). We are grateful to Weronika Białek, Martyna Burdach, Dorota Orzeszek and Judyta Pacewicz who conducted the testing sessions with the children and helped with coding the data.

CONFLICT OF INTEREST STATEMENT

The authors declare no competing interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in OSF at https://osf.io/tyw3g/

ORCID

Magdalena Łuniewska https://orcid.org/0000-0001-5504-9766
Magdalena Krysztofiak https://orcid.org/0000-0002-3160-9148
Ewa Haman https://orcid.org/0000-0003-1615-711X

REFERENCES

Bleses, D., Makransky, G., Dale, P.S., Højen, A. & Ari, B.A. (2016) Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, 37(6), 1461–1476. https://doi.org/10.1017/S0142716416000060

Byers-Heinlein, K., Jardak, A., Fourakis, E. & Lew-Williams, C. (2022) Effects of language mixing on bilingual children's word learning. *Bilingualism*, 25(1), 55–69. https://doi.org/10.1017/S1366728921000699

Cadime, I., Santos, A.L., Ribeiro, I. & Viana, F.L. (2021) Parental reports of preschoolers' lexical and syntactic development: Validation of the CDI-III for European Portuguese. *Frontiers in Psychology*, 12, 677575. https://doi.org/10.3389/fpsyg.2021.677575

Conboy, B.T. & Mills, D.L. (2006) Two languages, one developing brain: event-related potentials to words in bilingual toddlers.

Developmental Science, 9(1), F1–12. https://doi.org/10.1111/j.1467-7687.2005.00453.x

Dale, P.S., Paul, A., Rosholm, M. & Bleses, D. (2023) Prediction from early childhood vocabulary to academic achievement at the end of compulsory schooling in Denmark. *International Journal of Behavioral Development*, 47(2), 123–134. https://doi.org/10.1177/01650254221116878

Dunn, L.M. & Dunn, M. (2007) *PPVT-4: Peabody Picture Vocabulary Test*. Pearson Assessments.

Elliott, L., Silver, A.M., Imbeah, A. & Libertus, M. (2022) Actions may speak louder than words: comparing methods of assessing children's spontaneous focusing on number. *Journal of Experimental Child Psychology*, 214(February), 105301. https://doi.org/10.1016/j.jecp.2021.105301

Eriksson, M. (2017) The Swedish communicative development inventory III: parent reports on language in preschool children. *International Journal of Behavioral Development*, 41(5), 647–654. https://doi.org/10.1177/0165025416644078

Feldman, H.M., Dale, P.S., Campbell, T.F., Colborn, D.K., Kurs-Lasky, M., Rockette, H.E. & Paradise, J.L. (2005) Concurrent and predictive validity of parent reports of child language at ages 2 and 3 years. *Child Development*, 76(4), 856–868. https://doi.org/10.1111/j.1467-8624.2005.00882.x

Frank, M.C., Braginsky, M., Yurovsky, D. & Marchman, A.. (2021) *Variability and consistency in early language learning: The Wordbank Project*. The MIT Press.

Friend, M., Schmitt, S.A. & Simpson, A.M. (2012) Evaluating the predictive validity of the computerized comprehension task: comprehension predicts production. *Developmental Psychology*, 48(1), 136–148. https://doi.org/10.1037/a0025511

Garcia, I., Barreña, A., Ezeizabarrena, M.J., Almgren, M., Arratibel, N. & Barnes, J. (2014) "Haur Euskaldunen Komunikazio- Garapena Neurtzen 30–50 Hila- Bete Bitartean: MacArthur-Bates CDI-III Tresnaren Euskal Bertsioa." *Psikologia*. https://doi.org/10.26876/uztaro.88.2014.3

Gershkoff-Stowe, L. & Hahn, R. (2013) Word comprehension and production asymmetries in children and adults. *Journal of Experimental Child Psychology*, 114(4), 489–509 https://doi.org/10.1016/j.jecp.2012.11.005

Haman, E., Łuniewska, M., Hansen, P., Simonsen, H.G., Chiat, S., Bjekić, J., Blažienė, A., et al. (2017) Noun and verb knowledge in monolingual preschool children across 17 languages: data from cross-linguistic lexical tasks (LITMUS-CLT). *Clinical Linguistics & Phonetics*, 31(11-12), 818–843. https://doi.org/10.1080/02699206.2017.1308553

Haman, E., Łuniewska, M. & Pomiechowska, B. (2015) Designing cross-linguistic lexical tasks (CLTs) for bilingual preschool children. In S. Armon-Lotem, J. Jong, & N. Meir (Eds.), *Methods for assessing multilingual children: disentangling bilingualism from language impairment* (pp. 196–240). Multilingual Matters.

Hansen, P., Łuniewska, M., Simonsen, H.G., Haman, E., Mieszkowska, K., Kołak, J. & Wodniecka, Z. (2019) Picture-based vocabulary assessment versus parental questionnaires: a cross-linguistic study of bilingual assessment methods. *International Journal of Bilingualism*, 23(2), 437–456. https://doi.org/10.1177/1367006917733067

Hansen, P., Simonsen, H.G., Łuniewska, M. & Haman, E. (2017) Validating the psycholinguistic aspects of LITMUS-CLT: evidence from polish and norwegian. *Clinical Linguistics & Phonetics*, 31(11-12), 910–930. https://doi.org/10.1080/02699206.2017.1307455

Holm, E., Hansen, P.B., Romøren, A.S.H. & Gram Garmann, N. (2023) The norwegian CDI-III as an assessment tool for lexical and grammatical development in preschoolers. *Frontiers in Psychology*, 14(July), 1175658. https://doi.org/10.3389/fpsyg.2023.1175658

Jackson-Maldonado, D., Marchman, V.A., Dale, P. & Rubio-Codina, M. (2022) "The MacArthur-Bates CDI-III for Spanish-Speaking Children." *PsyArXiv*. https://doi.org/10.31234/osf.io/7cw4v

Jarůšková, L., Smolík, F., Chládková, K., Oceláková, Z. & Paillereau, N. (2023) How to build a communicative development inventory: insights from 43 adaptations. *Journal of Speech, Language, and Hearing Research: JSLHR*, 66(6), 2095–2117. https://doi.org/10.1044/2023_JSLHR-22-00591

Kas, B., Jakab, Z. & Lőrik, J. (2022) Development and norming of the Hungarian CDI-III: a screening tool for language delay. *International Journal of Language & Communication Disorders /Royal College of Speech & Language Therapists*, 57(2), 252–273. https://doi.org/10.1111/1460-6984.12686

Kuhl, P.K., Coffey-Corina, S., Padden, D., Munson, J., Estes, A. & Dawson, G. (2013) Brain responses to words in 2-year-olds with Autism predict developmental outcomes at age 6. *PLoS ONE*, 8(5), e64967. https://doi.org/10.1371/journal.pone.0064967

Lee, J. (2011) Size matters: early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics*, 32(1), 69–92. https://doi.org/10.1017/S0142716410000299

Legacy, J., Zesiger, P., Friend, M. & Poulin-Dubois, D. (2018) Vocabulary size and speed of word recognition in very young French–English bilinguals: a longitudinal study*. *Bilingualism: Language and Cognition*, 21(1), 137–149. https://doi.org/10.1017/S1366728916000833

Libertus, M.E., Odic, D., Feigenson, L. & Halberda, J. (2015) A developmental vocabulary assessment for parents (DVAP): validating parental report of vocabulary size in 2- to 7-year-old children. *Journal of Cognition and Development: Official Journal of the Cognitive Development Society*, 16(3), 442–454. https://doi.org/10.1080/15248372.2013.835312

Łuniewska, M., Haman, E., Armon-Lotem, S., Etenkowski, B., Southwood, F., Anđelković, D., Blom, E., et al. (2016) Ratings of age of acquisition of 299 words across 25 languages: is there a cross-linguistic order of words? *Behavior Research Methods*, 48(3), 1154–1177. https://doi.org/10.3758/s13428-015-0636-6

Łuniewska, M., Wójcik, M., Kołak, J., Mieszkowska, K., Wodniecka, Z. & Haman, E. (2022) Word knowledge and lexical access in monolingual and bilingual migrant children: impact of word properties. *Language Acquisition*, 29(2), 135–164. https://doi.org/10.1080/10489223.2021.1973475

Marchman, V.A. & Dale, P.S. (2023) The Macarthur-bates communicative development inventories: updates from the CDI advisory board. *Frontiers in Psychology*, 14(May), 1170303. https://doi.org/10.3389/fpsyg.2023.1170303

Marchman, V.A., Dale, P.S. & Fenson, L. (2023) *MacArthur-Bates Communicative Development Inventories User's Guide and Technical Manual* (3rd Ed.). Brookes Publishing.

Mcgillion, M., Davies, C., Kong, S.P., Hendry, A. & Gonzalez-Gomez, N. (2023) Caregiver sensitivity supported young children's vocabulary development during the covid-19 UK lockdowns. *Journal of Child Language*, March, 1–17. https://doi.org/10.1017/S0305000923000211

Mieszkowska, K., Krajewski, G., Łuniewska, M., Orzeszek, D., Banasik, N., Kuś, K., Otwinowska, A. & Haman, E.. (2021) "Kwestionariusz Dla Rodziców Dzieci Jedno- I Wielojęzycznych: Dzieci Przedszkolne I Starsze." University of Warsaw.

Mokhtari, F., Kazemi, Y., Feizi, A. & Dale, P.S. (2022) Psychometric properties of the Macarthur-bates communicative development inventories-III (CDI-III) in 30 to 37 months old Persian-speaking children. *Journal of Rehabilitation*, 23(3), 372–91. https://doi.org/10.32598/RJ.23.3.3166.3

Ring, E.D. & Fenson, L. (2000) The correspondence between parent report and child performance for receptive and expressive vocabulary beyond infancy. *First Language*, 20(59), 141–159. https://doi.org/10.1177/014272370002005902

Šmit Brleković, L. & Kuvač Kraljević, J. (2023) Parental reports on language development in toddlers and preschoolers based on the Croatian Version of communicative development inventories III. *Frontiers in Psychology*, 14, 1188550. https://doi.org/10.3389/fpsyg.2023.1188550

Smolík, F. & Bytešníková, I. (2021) Validity of the SDDS: a 40-Item vocabulary screening tool for 18- to 42-month olds in czech. *Journal of Communication Disorders*, 93(August), 106146. https://doi.org/10.1016/j.jcomdis.2021.106146

Sobota, K., Łuniewska, M., Krysztofiak, M. & Haman, E., (2022) Child Lexicon CLT (version 1.0.3). University of Warsaw. https://play.google.com/store/apps/details?id=pl.edu.uw.childlexiconclt

Tuller, L. (2015) Clinical use of parental questionnaires in multilingual contexts. In: S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children: disentangling bilingualism from language impairment*. Bristol: Multilingual Matters, pp. 299–328.

Tulviste, T. & Schults, A. (2020) Parental reports of communicative development at the age of 36 months: the Estonian CDI-III. *First Language*, 40(1), 64–83. https://doi.org/10.1177/0142723719887313

Wolna, A., Łuniewska, M., Haman, E. & Wodniecka, Z. (2023) Polish norms for a set of colored drawings of 168 objects and 146 actions with predictors of naming performance. *Behavior Research Methods*, 55(5), 2706–2732. https://doi.org/10.3758/s13428-022-01923-3

van Wonderen, E. & Unsworth, S. (2021) Testing the validity of the cross-linguistic lexical task as a measure of language proficiency in bilingual children. *Journal of Child Language*, 48(6), 1101–1125. https://doi.org/10.1017/S030500092000063X

How to cite this article: Łuniewska, M., Krysztofiak, M. & Haman, E. (2024) Parental report of vocabulary in 3- to 6-year-old Polish children: Reliable but not valid. *International Journal of Language & Communication Disorders*, 1–14. https://doi.org/10.1111/1460-6984.13101