The relationship between HIV and reading performance for children in Tanzania

Silvia Bonacina[a], Jonathan D. Lichtenstein[b,c], Christopher E. Niemczak[b,c], Albert Magohe[d], Abigail Fellows[b], Trent Nicol[a], Enica Massawe[e], Nina Kraus[a] and Jay C. Buckey[b,c]

Objective: Children with HIV (CWH) are at increased risk for cognitive and developmental delays, although HIV's influence on reading development remains unclear. Research using internationally validated reading measures with control for factors known to influence literacy outcomes is needed. The Early Grade Reading Assessment (EGRA) is a tool for assessing students' progress toward reading that has been validated across countries. It was administered to a cohort of children living with and without HIV (CWH/CWOH) and data on other factors that might affect literacy were also measured.

Design and methods: 388 children [217 children without HIV (CWOH) and 171 CWH; ages 3–8] drawn from a longitudinal study in Dar es Salaam, Tanzania completed the EGRA. EGRA performance between CWH and CWOH was compared adjusting for age, socioeconomic status, years of education, English learning, and type of school (public or private).

Results: Despite the biological and environmental confounders, CWH performed significantly worse than CWOH on the Letter Name Knowledge subtest, the Syllable subtest, the Non-Word subtest, and the Reading Comprehension subtest. The difference approached significance for the Oral Reading Fluency subtest.

Conclusions: CWH performed worse than CWOH on the EGRA, indicating literacy skill development in CWH needs early intervention. Longitudinal analyses, including electrophysiological and behavioral data, are needed to find the factors associated with poor reading and literacy performance in CWH.

AIDS 2025, **39**:49–56

Keywords: HIV, literacy development, preschoolers, reading, school-age children, Tanzania

Introduction

Children with HIV (CWH) are at risk for developmental delays and behavioral problems. A review of developmental outcomes in CWH under age 18 shows varying impacts across different developmental domains, with inconsistent evidence regarding language development. This emphasizes the urgency for well controlled studies on factors affecting language outcomes in CWH [1]. Failure to master literacy skills early in life has far-reaching implications. Children who do not learn to read in the early grades risk falling further behind in later grades as they cannot absorb printed information, follow written instructions, communicate well in writing, and actively participate in community life [2].

Recognizing the significance of prereading and reading skills, ministries of education and development experts from diverse international institutions collaboratively developed, piloted, and introduced the Early Grade

[a]Department of Communication Sciences and Disorders, Northwestern University, Chicago, IL, [b]Geisel School of Medicine at Dartmouth, Dartmouth College, [c]Dartmouth-Hitchcock Medical Center, Lebanon, NH, USA, [d]Dar Dar Programs, and [e]Muhimbili University of Health and Allied Sciences, Dar es Salaam, United Republic of Tanzania.

Correspondence to Nina Kraus, 2240 Campus Drive, Evanston, IL 60208, USA.

E-mail: nkraus@northwestern.edu

Received: 8 January 2024; revised: 22 August 2024; accepted: 16 September 2024.

DOI:10.1097/QAD.0000000000004020

Reading Assessment (EGRA) in 2006. This tool was deployed across >50 countries and used in 70 different languages to evaluate students' advancements in learning to read [3]. Since then, many children have been assessed with the EGRA around the world. Various EGRA metrics have been analyzed and revised to prove the reliability and validity of this instrument [4]. The EGRA has proven to be a strong assessment method, independent of the approach used in teaching a language.

Developing literacy skills, however, is influenced by various factors, both biological and environmental [5], particularly in low to middle income countries. Biological factors include age and gender, whereas common environmental factors are maternal educational level [6,7], socioeconomic status (SES) [8,9], child's education system, and the child–teacher relationship [10].

Data from 2021 show 96 000 children under the age of 14 are currently living with HIV in Tanzania (https://data.worldbank.org). This study examined the impact of HIV infection on reading measures provided by EGRA in children living in Dar es Salaam, Tanzania. The analysis included several environmental factors that might affect language performance measures. CWH from this cohort show worse automatized naming skills [11], intellectual functioning, attention, processing speed [12,13], and subcortical auditory processing compared to CWOH [14]. These areas are closely related with reading abilities, so our central hypothesis was that the EGRA would reveal differences between CWH and CWOH independently of biological and environmental factors that may moderate literacy skills.

Methods

The Dartmouth College Committee for the Protection of Human Subjects and the Research Ethics Committee of the Muhimbili University of Health and Allied Sciences (MUHAS) approved the research. During the first study session a parent or guardian provided consent for the children to participate.

Participants

Participants were part of an ongoing longitudinal study in Dar es Salaam, Tanzania, including children recruited between the ages of 3–8 without major neurocognitive disease or mental illness with the target of enrolling the same number of children across the different age groups.

Four hundred and seventy-two children were originally enrolled in the study. Of these, 388 children had EGRA data and were included in the current analysis. Among the 388 children, 385 took part in the first session, 365 in two consecutive sessions, 286 in three sessions, 150 in four sessions, 62 in five sessions, 17 in six sessions, and 6 in seven sessions.

One hundred and seventy-one CWH and 217 children without HIV (CWOH) between the ages of 3 and 8 (mean 6.97 ± 1.41 years) were included for this cross-sectional study. All CWH were receiving antiretroviral treatment. CWOH were enrolled in the same way as CWH from within the city of Dar es Salaam from local pediatric programs, district hospitals, schools, or through a referral from (ongoing) participating guardians. Demographic variables of sex, age, years of education, and socioeconomic status were similar between groups. English instruction at home or school differed between CWH and CWOH (see Table 1).

Self-report questionnaire

A demographic questionnaire (Appendix 1, Supplemental Digital Content, http://links.lww.com/QAD/D328) covering early life experiences, schooling, socio-economic status (SES), English learning, and other factors was completed by the parent or guardian of each participant. Overall socio-economic status was computed as reported in [12] using a principal component analysis (PCA) on the SES questions. The PCA was based on questions that addressed three main categories of SES: household assets and access to resources, measures of income and occupation, and measures of parental education. PCA was used to reduce the dimensionality of the dataset by linearly transforming it into a noncorrelated set of principal components. The eigenvalues of the principal components were used to represent the amount of variance that each principal component captured, and thus, the majority of the total variance in the data could be represented in relatively few principal components. Positive values signify higher SES. Years of education were calculated from the answers to the questions in the education section of the questionnaire. English learning at home and at school was assessed by considering the responses to the questions related to the English Instruction section of the demographic questionnaire (see Appendix 1, Supplemental Digital Content, http://links.lww.com/QAD/D328 for specific questions). The kind of school (public or private) attended by the children was documented in the Child's education section of the demographic questionnaire.

Early Grade Reading Assessment

The EGRA is the product of collaboration among scholars, practitioners, government officials, and education development professionals to advance early reading assessment among primary school children in low-income countries [4]. In this study, we focused on five subtests: Letter Name Knowledge, Syllables, Nonwords, Oral Reading Passage, and Reading Comprehension. The EGRA was administered in Kiswahili.

Letter name knowledge

The subtest Letter Name Knowledge is one of the most basic assessments of reading preparedness (and risk for reading difficulties). In this task, the children are asked to

Table 1. Demographic information for children with HIV (CWH) and children without HIV (CWOH) and provides statistical results of tests for group differences.

	CWH	CWOH	Group differences χ^2 or 95% CI, *F* statistic, df, and p-value
Sex	81 males – 90 females	107 males – 110 females	$\chi^2 = 0.144$, df = 1, $P = 0.704$
Age (mean ± SD)	6.89 ± 1.37 years	7.07 ± 1.46 years	[−0.464, +0.104], $F = 1.551$, df = 1, $P = 0.214$
Years of education (mean ± SD)	2.79 ± 1.31 yrs.	2.99 ± 1.37 years	[−0.075, +0.474], $F = 2.034$, df = 1, $P = 0.155$
Socioeconomic Status (mean ± SD)	−0.040 ± .954	0.110 ± 1.05	[−0.57, +0.358], $F = 2.032$, df = 1, $P = 0.155$
English learning School and home (% yes)	38.9%	53.5%	$\chi^2 = 7.769$, df = 1, $P = 0.005$**
School system (private or public)	120 Private 70.2% 43 Public 25.1% 8 n/a 4.6%	162 Private 74.6% 50 Public 23.04% 5 n/a 2.3%	$\chi^2 = 0.386$, df = 1, $P = 0.534$

CI, confidence interval; SD, standard deviation.

provide the names (not the sounds) of all of the letters of the alphabet they can identify within one minute. It assesses the ability to recognize the graphemic features of each letter and accurately associate it to its name. The full set of letters of the alphabet is listed in random order, 10 letters to a row, using a clear, large, and familiar font. The child's score for this subtest is calculated as the number of correct letters identified aloud per minute. For this subtest the English version of the EGRA was used.

Syllables

The Syllables subtest assesses the ability to identify common syllables easily and automatically as an essential step to developing reading fluency. In this task the child is asked to read aloud a list of 100 syllables presented in random order on a page within one minute. The child's score for this subtest is calculated as the number of correct syllables read aloud per minute.

Nonwords

The subtest Nonwords assesses the ability to decode nonwords based on the knowledge of grapheme-phoneme mappings. In this task, the child is given one minute to read aloud quickly and carefully a list of 50 nonwords presented in random order on a page. Words were constructed from correct Swahili letter combinations but were nonsensical. The child's score for this subtest is calculated as the number of correct nonwords read aloud per minute.

Oral reading passage

The subtest Oral Reading Passage assesses fluency and reading accuracy. In this task, the child has one minute to read aloud a grade-level short story printed on a page. The child's score for this subtest is calculated as the number of correct words read aloud per minute.

Reading comprehension

The subtest Reading Comprehension assesses the ability to understand the meaning of what is read. In this task, after children read the passage aloud, they are asked three to five comprehension questions about the passage, both explicit and inferential. The child's score for this subtest is calculated as the number of correct answers.

Statistical analysis

Before assessing the difference in literacy performance between CWH and CWOH, we tested for sex differences between the HIV groups using the χ^2 statistic and we performed three distinct univariate ANOVAs to compare means of demographic variables (age, years of education, and SES at time of EGRA assessment) and evaluate whether the HIV groups were matched with respect to them. We also ran descriptive statistics to test whether the HIV groups were balanced with respect to school type prevalence and to English learning exposition both at home and at school.

To examine the factors associated with completion of the five EGRA subtasks, logistic regression was used. Test completion (0 = yes/1 = no) was used as the outcome variable, with HIV status (0 = CWH O; 1 = CWH), age, SES; English instruction (1 = yes; 0 = no), and school system (0 = public; 1 = private) as predictors.

To test the relationship between each EGRA subtest and age, we correlated age and EGRA performance individually within each group (CWH and CWOH) and then we tested for HIV status*Age interaction effect within each subtest. To specifically test the difference between children with and without HIV with respect to each EGRA subtest performance, we computed five univariate ANOVA considering HIV status as fixed factor and sex (0 = male; 1 = female), age, SES, years of education, the English instruction variable (1 = yes; 0 = no) and the kind of school (0 = public or 1 = private) as covariates. Each ANOVA was related to one single EGRA's subtest so as to include only the children that were able to complete the specific subtest entirely.

Finally, to support the validity of the data collected, we qualitatively compared them with the data available from the Tanzania National assessments, which were conducted in 2013 and 2016 to understand how students in

Tanzania were developing literacy, basic mathematics, and life skills. The National assessment involved only public schools [15,16]. See Appendix 2 Table A2, Supplemental Digital Content, http://links.lww.com/QAD/D329 for more information about this comparison.

Results

Participant characteristics and demographics

CWH and CWOH did not differ significantly on sex, age, years of education, SES, and kind of school. They did differ significantly on English instruction at school or at home, with a higher percentage of CWOH exposed to English instruction than CWH (Table 1).

Data from the first visit when the child was able to perform at least one EGRA subtest successfully were used in the analysis. 41.5% of the EGRA data were collected during the 1st visit, 36% during the 2nd visit, 13.4% during the 3rd visit, 5.2% during the 4th visit, and 3.1% during the 5th visit. This range of EGRA availability was driven mainly by the age of the child at their first visit.

Factors playing a role in determining the ability to fully complete an Early Grade Reading Assessment subtask

Age was a significant predictor across all five EGRA subtasks. As children age their probability of completing a subtask fully increased. HIV status appeared to significantly predict task completion for the Nonwords and Reading Comprehension subtests, revealing that children with HIV are less likely to complete these two subtests. SES predicted task completion for the Nonwords and Oral Reading Passage subtests, showing that having high SES increases the probability of fully completing the task. Finally, the English learning factor predicted task completion for the Oral Reading Passage subtest, revealing the presence of English instruction can increase the probability of completing the task. See Appendix 2 Table A1, Supplemental Digital Content, http://links.lww.com/QAD/D329 for regressions results.

Early Grade Reading Assessment performance relates to age similarly between groups

Across both groups, age was positively related with EGRA performance except for the Oral Reading Passage subtest in CWOH where the relationship was not significant. For each subtest, we assessed the difference in the slopes between the groups. The interaction of age and HIV status was significant only for the Oral Reading Passage subtest ($F = 4.224$; $P = .047$). See Table 2 for summary of results. We also reported, in Table S1, Supplemental Digital Content, http://links.lww.com/QAD/D330, the results of the ANOVAs obtained considering "0" score when the subtask was not completed and, in Table S2, Supplemental Digital Content, http://links.lww.com/QAD/D330, the result of correlations between Age and each subtest performance computed using the nonparametric Spearman statistic, instead of Pearson. The results are consistent with those reported in Table 2. Figure 1 displays the relationship between age and EGRA performance on each subtest for CWH and CWOH.

Early Grade Reading Assessment performance at each subtest between groups

Children with HIV performed significantly worse than children without HIV on all EGRA subtasks considering sex, age, SES, years of education, English instruction at school or at home, and school type as covariates. See Table 3 for summary of results. Figure 2 displays mean performance ± 95% confidence interval (CI) for each EGRA subtest for CWH and CWOH.

Discussion

This study describes performance on EGRA subtests in children living with and without HIV in a low- and middle-income country (LMIC). It is the first to explore the effects of HIV on this extensively employed early reading assessment tool. Given the role of environmental factors on language development, in the main analysis we controlled for these variables to disentangle the effect of HIV infection from other potential confounding factors. When considering all covariates, we found that CWH performed worse than children without HIV on the Letter Name Knowledge, Syllable, Nonword, and Reading Comprehension subtests. The difference approached significance for the Oral Reading Passage subtest. These differences were small to moderate in effect size, and they were significant when controlling for age, sex, SES, variables related to the instruction of written English at school, and the kind of school attended by each child. Therefore, we assume an HIV related mechanism underlies the significant differences between groups. The results were only marginally different when no covariates were included or only age was considered. The main objective of the current study was to consider as many external contributing factors as possible so the results with these confounders are presented.

Recent studies have highlighted auditory central nervous system dysfunction in individuals with HIV, from childhood to adulthood [17,14]. These studies showed a less robust, more variable, and less accurate neural encoding of speech stimuli when compared with individuals without HIV. Knowing the relationship between the integrity of neural encoding of speech sound and literacy skills [17–20], we hypothesize that the difference we are seeing in EGRA reflect underlying brain dysfunction, which is revealed in the central auditory tests. Studies exploring electrophysiological and behavioral data together may be able to confirm this link.

Table 2. (i) R Pearson correlation values of the relationship between each subtest performance and age for children without HIV (CWOH) and with HIV (CWH) separately; (ii) mean and SD of EGRA subtest's performance within each HIV group; (iii) ANOVA's results of EGRA performance's comparison between HIV groups without any covariate; (iv) ANOVA's results of EGRA performance's comparison between HIV groups with age only as covariate.

	Age relationship within each HIV group CWOH CWH	Descriptive statistics, mean (std. deviation) CWOH CWH	Univariate ANOVA 95% CI for HIV group difference [lower bound − upper bound], *F* statistic, *P*-value; μ^2	Univariate ANOVA with age as unique covariate 95% CI for HIV group difference [lower bound − upper bound], *F* statistic, *P*-value; μ^2	
Letter name	$R=0.463$; $P<0.001$	33.30 (23.992)	[−0.539 − 9.146]	HIV	[1.358 − 10.080] $F=1.044$; $P=0.308$; $\mu^2=0.003$
Knowledge	$R=0.409$; $P<0.001$	28.99 (20.400)	$F_{(1,342)}=3.055$, $P=0.081$; $\mu^2=0.009$	Age	$F=78.510$; $P<.001$; $\mu^2=0.188$
				Age*HIV	$F=2.383$; $P=0.124$; $\mu^2=0.007$
Syllables	$R=0.505$; $P<0.001$	40.74 (27.444)	[3.387 − 16.119]	HIV	[7.763 − 19.158] $F=21.639$; $P<0.001$; $\mu^2=0.075$
	$R=0.424$; $P<0.001$	30.99 (23.806)	$F_{(1,267)}=9.099$, $P=0.003$; $\mu^2=0.033$	Age	$F=66.795$; $P<0.001$; $\mu^2=0.201$
				Age*HIV	$F=0.863$; $P=0.353$; $\mu^2=0.003$
Nonwords	$R=0.421$; $P<0.001$	18.41 (11.277)	[0.371 − 6.276]	HIV	[1.469 − 6.969] $F=0.343$; $P=0.559$; $\mu^2=0.002$
	$R=0.325$; $P=0.003$	15.09 (9.174)	$F_{(1,207)}=4.924$, $P=0.028$; $\mu^2=0.023$	Age	$F=29.831$; $P<.001$; $\mu^2=0.127$
				Age*HIV	$F=1.094$; $P=0.297$; $\mu^2=.005$
Oral Reading	$R=0.074$; $P=0.742$	47.00 (14.458)	[4.554 − 25.867]	HIV	[5.600 − 24.829] $F=5.786$; $P=0.021$; $\mu^2=0.217$
Passage	$R=0.616$; $P=0.005$	31.79 (19.217)	$F_{(1,39)}=8.335$, $P=0.006$; $\mu^2=0.176$	Age	$F=6.375$; $P=0.016$; $\mu^2=0.147$
				Age*HIV	$F=4.224$; $P=0.047$; $\mu^2=0.102$
Reading	$R=0.421$; $P<0.001$	2.33 (1.365)	[−0.153 − 0.737]	HIV	[0.058 − 0.891] $F=0.020$; $P=0.888$; $\mu^2=0.000$
Comprehension	$R=0.380$; $P=0.004$	2.04 (1.388)	$F_{(1,162)}=1.682$, $P=0.197$; $\mu^2=0.010$	Age	$F=27.740$; $P<0.001$; $\mu^2=0.148$
				Age*HIV	$F=0.034$; $P=0.855$; $\mu^2=0.000$

CI, confidence interval; EGRA, Early Grade Reading Assessmentl SD, standard deviation.

Among the biological factors, age was positively correlated with EGRA performance across most subtests, which was expected as formal reading instruction is introduced as children become older. The only exception was the Oral Reading Passage, specifically for children without HIV. This result was not surprising given the age range of the children involved in this study [22]. Only a subset of children over 6 could complete the Oral Reading Passage test, and the absence of relationship for the Oral Reading Passage subtest is likely due to the lower number of children attempting this specific subtest. Additionally, there is a ceiling effect for this subtest with about one-quarter of the children across both groups finishing at the highest score.

Among the external variables, we examined the kind of school children attended by asking parents the name of the school and whether it was a private or public institution. Even though the majority (72.7%) of the children in our cohort attended private schools, there was no relationship between the kind of school and HIV status. In Tanzania, the main difference between these two options is the number of children in each classroom and, consequently, the kind of relationship that each child can establish with the teacher. Another difference is the language of instruction. In private school, English is most often the language of instruction whereas Kiswahili is taught as a subject. In public school it is the opposite. Given the amount of time children spend at school, we thought that this variable would have been crucial to EGRA performance. Nevertheless, the results revealed HIV to be the strongest predictor of literacy performance, after age.

A similar trend was observed with SES. The measure of SES we used for the analysis was a result of a principal component analysis (PCA), previously described [13,11], that includes variables related to both household assets and access to resources, and parental education levels. In our cohort, no relationship was found between SES and HIV status, which further supports the influence of HIV on literacy performance.

Given the impact simultaneously learning two languages has on language and literacy development [21], we asked whether children were learning to speak and write English at school and at home. The data collected revealed that learning English was less likely in CWH even though the groups did not differ in SES. We may be seeing this relationship because CWH are living with adults with HIV, which can increase stressors in the home [23,24]. As such, learning English may be less of a priority in the context of other more immediate needs.

Analyses performed to explore the factors playing a role in the ability of the children to entirely complete a subtest contributed to a better understanding of the data. Across all subtests, age was a predictor of successful subtest completion. SES and HIV status appeared to have a role only with respect to some subtests. English instruction revealed a role only for the Oral Reading Passage subtest. The kind of school did not predict the ability to fully perform a subtest. These results are in line with the fact that EGRA is aimed at tracking learning progress and therefore its completion success relates with age. Moreover, its completion success looks only partially influenced by the other environmental factors considered.

The current study has some limitations to. While we accounted for a variety of environmental variables, some potentially important elements were not included in the analysis (e.g., objective value for the quality of education experienced by each child; parental mortality; nutritional, growth, and other health concerns; English performance

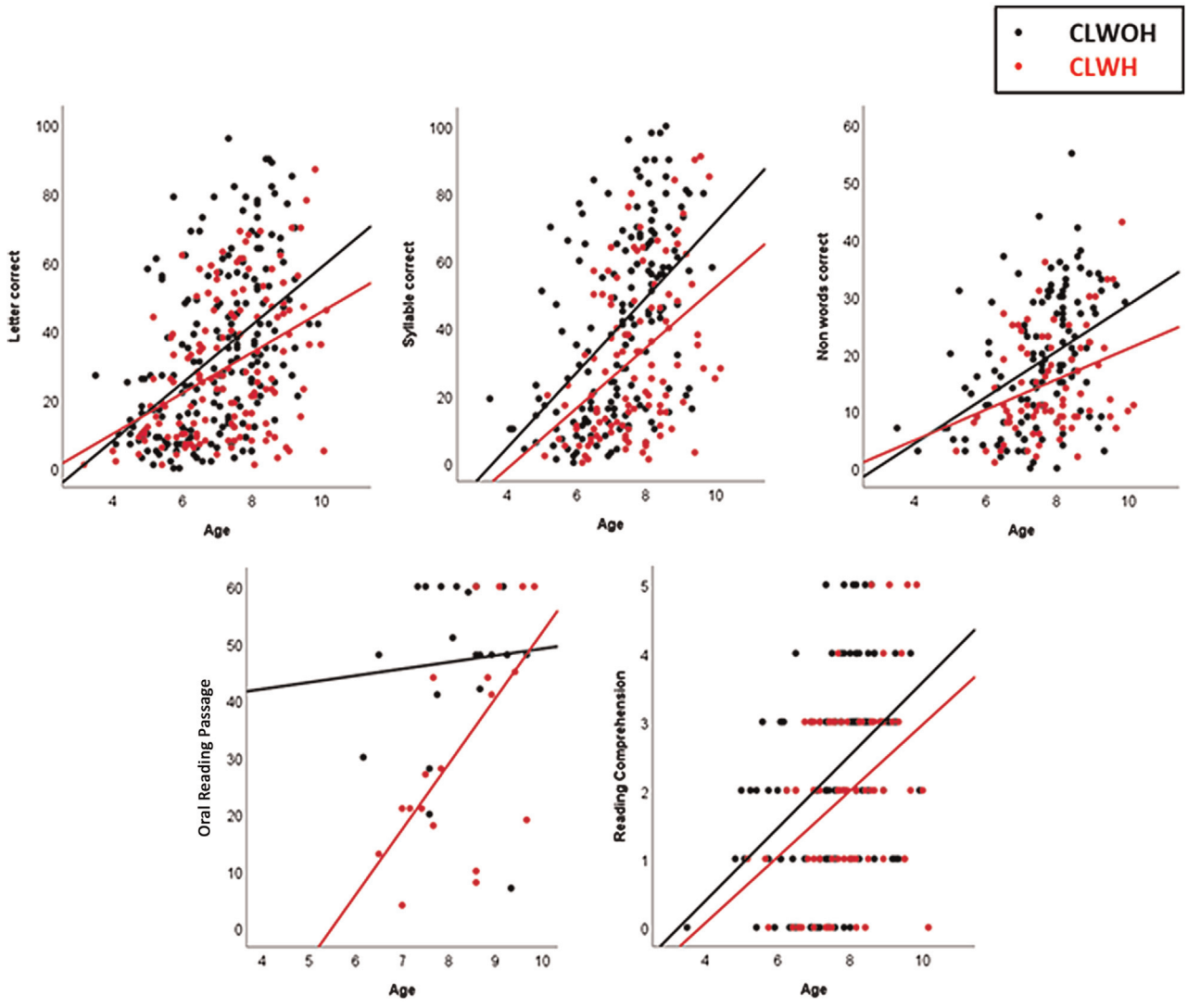

Fig. 1. Scatterplots display the relationship between age and EGRA performance on each subtest for CWH (red dots and red linear fit) and CWOH (black dots and black linear fit). CWH, children with HIV; CWOH, children without HIV. EGRA, Early Grade Reading Assessment. Age*HIV group interaction resulted significant only for the Oral Reading Passage subtest ($P < 0.05$).

Table 3. Mean and SE within each HIV group (CWH – CWOH) adjusted for all covariates considered (sex [1 = female; 0 = male]; age; socio-economic status (SES); years of education; English instruction at school or at home [1 = yes; 2 = no], and school type [0 = public; 1 = private]) and ANOVA results related to HIV group comparisons with respect to EGRA performance considering sex, age, SES, years of education, English instruction, and school type as covariates.

	CWH Mean (Std. error) *N*	CWOH Mean (Std. error) *N*	HIV group differences 95% CI for difference [lower bound – upper bound], *F* statistic, *P*-value; μ^2
Letter Name Knowledge	29.017 (1.673) $N = 139$	34.040 (1.518) $N = 168$	[0.513 – 9.532], $F_{(1,307)} = 4.805$, $P = 0.029$; $\mu^2 = 0.016$
Syllables	29.845 (2.284) $N = 103$	42.474 (1.964) $N = 137$	[6.520 –18.737], $F_{(1,240)} = 16.590$, $P < 0.001$; $\mu^2 = 0.067$
Nonwords	15.031 (1.140) $N = 75$	18.664 (0.927) $N = 111$	[0.658 – 6.607], $F_{(1,186)} = 5.808$, $P < 0.017$; $\mu^2 = 0.032$
Oral Reading Passage	33.160 (3.854) $N = 18$	44.618 (3.854) $N = 18$	[−0.271 – 23.187], $F_{(1,36)} = 4.004$, $P = 0.055$; $\mu^2 = 0.125$
Reading Comprehension	1.933 (0.178) $N = 53$	2.39 (0.130) $N = 96$	[0.012 – 0.903], $F_{(1,149)} = 4.115$, $P = 0.044$; $\mu^2 = 0.028$

CI, confidence interval; EGRA, Early Grade Reading Assessment; SD, standard deviation.

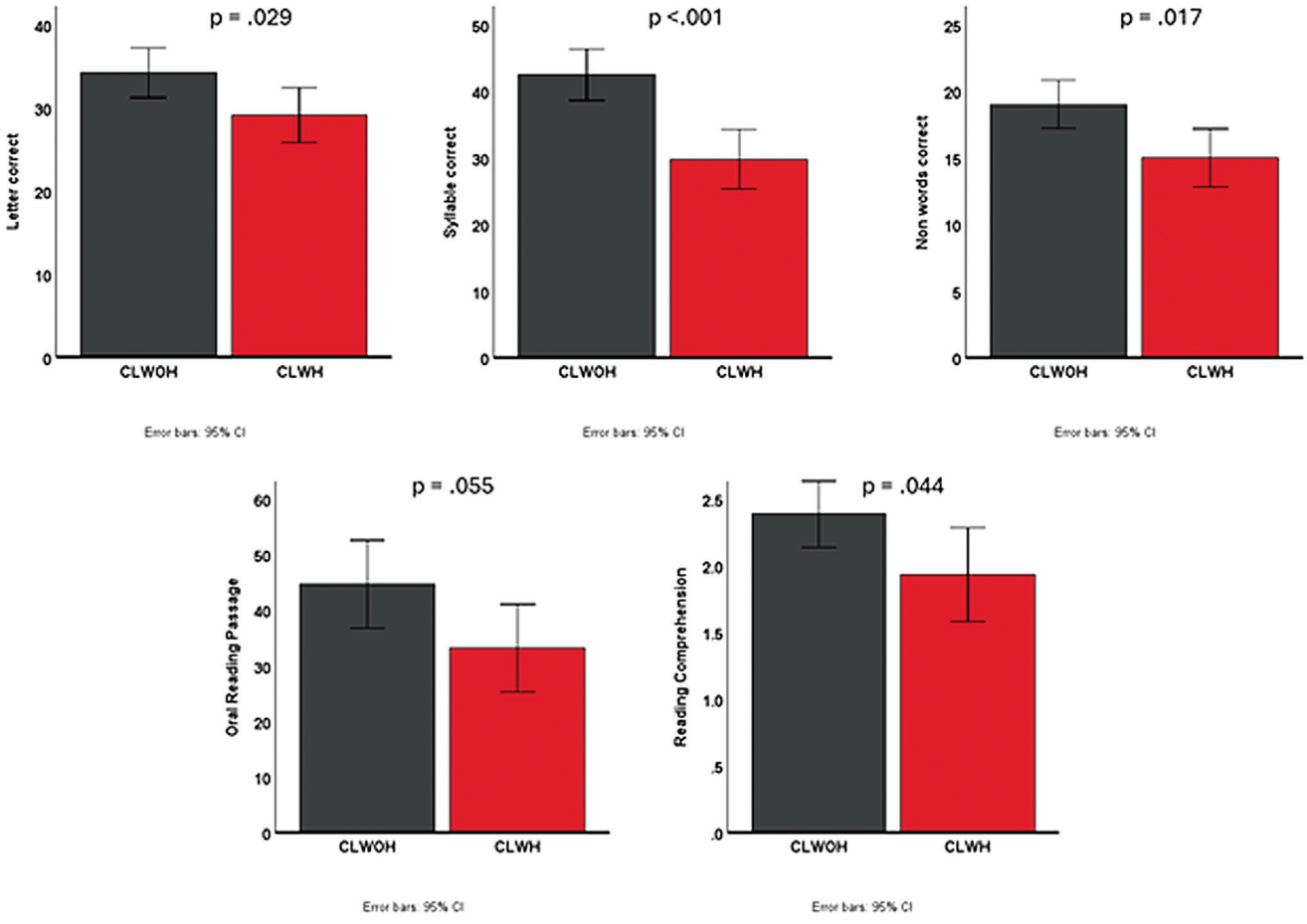

Fig. 2. Bar graph of mean performance ± 95% CI for each EGRA subtest shows that children with HIV performed worse than children without HIV for all subtests. CI, confidence interval; CWH, children with HIV; CWOH, children without HIV; EGRA, Early Grade Reading Assessment. The mean and CI reported in the figures are adjusted for all covariates considered in the analysis (sex, age, SES, years of education, English instruction at school or at home, and school type). Black bar refers to CWOH and red bars refer to CWH.

level; factors and psychological distress related to parental health [23]). This study also did not differentiate between CWOH born from mothers with HIV at their birth and between different levels of illness severity across CWH, which may yield different literacy profiles. Moreover, it did not take into consideration the HIV situation of the parents, whether they are alive, treated, and/or in good health condition. Finally, as the age of our participants was restricted from 3 to 10, this does not allow us to generalize our results at different ages.

This study also has important strengths. While our findings may be best related to a certain age group at the Dar es Salaam research site, the results of the study confirmed the hypothesis that HIV infection impairs literacy skills in children, using an instrument that was previously used in Tanzania. This potentially motivates early intervention in CWH to support their literacy development as early as possible. This study also supported the importance of considering environmental factors during analysis. Further studies should follow this methodology to provide supporting evidence, ideally within longitudinal study designs, to monitor language and reading development over time, including the factors not accessible to this study that may influence literacy outcomes.

Acknowledgements

We would like to thank the study team in Tanzania (Claudia Gasana, Filmon Samuel, Godfrey Njau, Joyce Kibasa, Joyce Machunda, Modestus Choka, Matilda Kabeho, and Pascal Maibe) for their substantial contributions towards data collection. We would like to thank Katie Alcock and Damaris Ngorosho for their assistance in the early phases of the study. This work was supported by NICHD grant R01HD095277 (PI Buckey).

Conflicts of interest

There are no conflicts of interest.

References

1. Sherr L, Croome N, Castaneda KP, Bradshaw K, Romero RH. **Developmental challenges in HIV infected children—an updated systematic review.** *Child Youth Serv Rev* 2014; **45**:74–89.
2. Mancuso JM. **Health literacy: a concept/dimensional analysis.** *Nurs Health Sci* 2008; **10**:248–255.
3. Gove AK, Wetterberg A, editors. The early grade reading assessment: applications and interventions to improve basic literacy. RTI Press; 2011.
4. Dubeck MM, Gove A. **The early grade reading assessment (EGRA): Its theoretical foundation, purpose, and limitations.** *Int J Educ Dev* 2015; **40**:315–322.
5. Hayiou-Thomas ME. **Genetic and environmental influences on early speech, language and literacy development.** *J Commun Disord* 2008; **41**:397–408.
6. Vernon-Feagans L, Bratsch-Hines M, Reynolds E, Willoughby M. **How early maternal language input varies by race and education and predicts later child language.** *Child Dev* 2020; **91**:1098–1115.
7. Justice LM, Jiang H, Bates R, Koury A. **Language disparities related to maternal education emerge by two years in a low-income sample.** *Matern Child Health J* 2020; **24**:1419–1427.
8. Pace A, Luo R, Hirsh-Pasek K, Golinkoff RM. **Identifying pathways between socioeconomic status and language development.** *Annu Rev Linguist* 2017; **3**:285–308.
9. Hoff E. **How social contexts support and shape language development.** *Dev Rev* 2006; **26**:55–88.
10. Pianta RC, Stuhlman MW. **Teacher-child relationships and children's success in the first years of school.** *School Psychol Rev* 2004; **33**:444–458.
11. Bonacina S, Niemczak C, Lichtenstein JD, Magohe A, Fellows A, Nicol T, *et al*. **Preliteracy assessment in children living with HIV in Tanzania: comparison to results from children living without HIV in Tanzania and the United States.** *AIDS* 2023; **37**:1077–1083.
12. Lichtenstein J, Bowers C, Amato J, Niemczak C, Fellows A, Magohe A, *et al*. **Nonverbal cognitive assessment of children in Tanzania with and without HIV.** *Child Neuropsychol* 2022; **28**:107–119.
13. Lichtenstein J, Niemczak C, Bowers C, Magohe A, Fellows A, Ealer C, *et al*. **Nonverbal neurocognitive assessment during the coronavirus disease of 2019 pandemic: the effect of personal protective equipment.** *Arch Clin Neuropsychol* 2022; **37**:1628–1632.
14. Ealer C, Niemczak CE, Nicol T, Magohe A, Bonacina S, Zhang Z, *et al*. **Auditory neural processing in children living with HIV uncovers underlying central nervous system dysfunction.** *AIDS* 2022:10–97.
15. RTI International. National baseline assessment for the 3Rs (reading, writing, and arithmetic) using EGRA, EGMA, and SSME in Tanzania.
16. RTI International. Assistance to basic education: all children reading (ABE-ACR). findings report, Tanzania National Early Grade Reading Assessment (EGRA).
17. White-Schwoch T, Magohe AK, Fellows AM, Rieke CC, Vilarello B, Nicol T, *et al*. **Auditory neurophysiology reveals central nervous system dysfunction in HIV-infected individuals.** *Clin Neurophysiol* 2020; **131**:1827–1832.
18. White-Schwoch T, Woodruff Carr K, Thompson EC, Anderson S, Nicol T, Bradlow AR, *et al*. **Auditory processing in noise: a preschool biomarker for literacy.** *PLoS Biol* 2015; **13**: e1002196.
19. Hornickel J, Kraus N. **Unstable representation of sound: a biological marker of dyslexia.** *J Neurosci* 2013; **33**:3500–3504.
20. Banai K, Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. **Reading and subcortical auditory function.** *Cereb Cortex* 2009; **19**:2699–2707.
21. Gehsmann KM. **Stages and standards in literacy: teaching developmentally in the age of accountability.** *J Educ* 2012; **192**:5–16.
22. Bialystok E. **The impact of bilingualism on language and literacy development.** *The handbook of bilingualism and multilingualism* 2012:624–648.
23. Rochat T, Netsi E, Redinger S, Stein A. **Parenting and HIV.** *Curr Opin Psychol* 2017; **15**:155–161.
24. Cluver L, Operario D, Gardner F. **Parental illness, caregiving factors and psychological distress among children orphaned by acquired immune deficiency syndrome (AIDS) in South Africa.** *Vulner Child Youth Stud* 2009; **4**:185–198.