Validating a Population Measure of Early Childhood Development in Low- and Middle-Income Countries: The Early Human Capability Index

Alanna Sincovich

School of Public Health

Faculty of Health and Medical Sciences

The University of Adelaide

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

September 2022

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AEDC	Australian Early Development Census
AIM-ECD	Anchor Items for Measuring Early Childhood Development
ARNEC	Asia-Pacific Regional Network for Early Childhood
ASQ	Ages and Stages Questionnaire
AUC	Area under the curve
BL	Baseline
CCDG	Community Child Development Group
CFA	Confirmatory factor analysis
CFI	Comparative Fit Index
CI	Confidence interval
CR	Caregiver report
CREDI	Caregiver Reported Early Development Instrument
DA	Direct assessment
DAG	Directed acyclic graph
DF	Degrees of freedom
EAP-ECDS	East Asia Pacific Early Child Development Scales
EGMA	Early Grade Maths Assessment
EGRA	Early Grade Reading Assessment
ECD	Early childhood development
ECDI	Early Child Development Index
ECE	Early childhood education
ECERS	Early Childhood Environment Rating Scales
EDI	Early Development Instrument
eHCI	early Human Capability Index

EL	Endline
GDP	Gross Domestic Product
GNI	Gross National Income
GSED	Global Scale for Early Development
IDELA	International Development and Early Learning Assessment
IQR	Interquartile range
LL	Lower limit
LF	Long form
LMIC	Low- and middle-income country
MAR	Missing at random
MAT	Multi-Age Teaching
MCAR	Missing completely at random
MDAT	Malawi Developmental Assessment Tool
MDG	Millennium Development Goal
MELE	Measure of Early Learning Environments
MELQO	Measuring Early Learning and Quality Outcomes
MICS	Multiple Indicator Cluster Survey
MODEL	Measure of Development and Early Learning
MNAR	Missing not at random
PDR	People's Democratic Republic
PEARL	Pacific Early Age Readiness and Learning
PRIDI	Regional Project on Child Development Indicators
RCT	Randomised control trial
RMSEA	Root mean square error of approximation
ROC	Receiver Operator Characteristic

SD	Standard deviation
SDG	Sustainable Development Goal
SES	Socioeconomic status
SF	Short form
TFI	Tucker-Lewis Fit Index
TR	Teacher report
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children's Fund
UL	Upper limit
WISC	Wechsler Intelligence Scale for Children
WHO	World Health Organization
WLSMV	Weighted least squares mean and variance adjusted

Abstract

Increased focus on population monitoring of early childhood development has been spurred by the Sustainable Development Agenda, together with burgeoning evidence for the need to support children to reach their developmental potential. Tracking children's development in low- and middle-income countries (LMICs) has been challenged by a lack of appropriate measurement tools and resources to implement measurement. The early Human Capability Index (eHCI) was designed to measure holistic development among children aged 3-5 years, be feasible for large-scale use in low resource settings, and capture locally relevant information. The aim of this thesis was to investigate the reliability, validity, and sensitivity of the eHCI across diverse settings to advance understanding of how the tool can be used to facilitate population measurement of early childhood development.

Four research studies utilised pre-existing data collected using the eHCI from 2013-2020 among children aged 2-6 years in seven LMICs, including Brazil, China, Kiribati, Lao PDR, Samoa, Tonga, and Tuvalu. The first study investigated whether data fit the theoretical structure of the eHCI (nine developmental domains) across seven countries, given the necessary adaptation of the instrument in each country. Confirmatory factor analyses indicated the eHCI maintained the same factor structure across countries, providing evidence for the tool's construct validity. The second study explored the convergent, divergent, discriminant, and concurrent validity of the eHCI, and whether results varied across seven countries. Results provided evidence that the tool captured aspects of early childhood development it was designed to measure. Although the eHCI was intended to measure development among children aged 3-5, results from this study indicated it may be validly applied to children aged 2-6 years.

The first two studies established the eHCI was psychometrically robust using cross-sectional data. The third study used longitudinal data to explore the ability of the eHCI to predict

children's later abilities in Lao PDR, establishing predictive ability of the tool. Specifically, whether scores on the eHCI at 2-5 years predicted cognitive development (literacy, numeracy, executive function) at 6-9 years, four years later. Receiver Operator Characteristic curve analyses demonstrated the summary indicator, eHCI overall development, signalled risk for poor future cognitive development with similar ability to measures of socioeconomic position.

The eHCI was designed to have adequate sensitivity to detect variation in children's development to facilitate program evaluation, which is a limitation of many existing population measures. The final study tested the sensitivity of eHCI scores to inputs promoting children's development, namely quality of early childhood education. Using cross-sectional data in Lao PDR, adjusted linear regressions demonstrated small, positive associations between quality and children's development measured via the eHCI, as was hypothesised.

Together, studies demonstrated that the eHCI, a pragmatic, freely available and locally adapted tool, can be validly applied to children aged 2-6 years across diverse LMICs, for the purposes of locally relevant population monitoring of early childhood development, as well as program evaluation. Ultimately, information collected using the eHCI may be used to inform policy and practice in terms of resourcing and supports to promote children's development.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Alanna Sincovich

02/09/2022

Acknowledgements

I would like to express my gratitude to the children, caregivers, teachers, and schools who participated in the projects for which existing data were used throughout this thesis. I thank the staff and organisations, as well as local communities, who coordinated and supported data collection in Brazil, China, Kiribati, Lao PDR, Samoa, Tonga, and Tuvalu. Without these efforts and willingness to participate, this research would not have been possible.

I also acknowledge the organisations that provided financial and technical support to these projects; including the World Bank, the Global Partnership for Education, the China Development Research Foundation, and the Fundação Maria Cecilia Souto Vidigal.

I acknowledge the funding I received to support my candidature through the Commonwealth Scholarship Program for South Australia. I am thankful also for the funding I received from the Public Health Association of Australia, the University of Adelaide, the Australian Epidemiological Association, and Healthy Development Adelaide, all of which supported dissemination of my research in the form of conference presentations and journal articles.

To my supervisory panel, Professor Sally Brinkman, Dr Tess Gregory, and Professor John Lynch, sincere thanks for your guidance and scientific input throughout my PhD. I have learnt from you all and am grateful for the opportunity to be involved in this research.

Finally, I would like to thank those who formed my support system throughout my PhD, including my colleagues at the Telethon Kids Institute, friends, and family. Special thanks to my mum (Sabine) and sister (Bianca) for your constant encouragement; my colleagues Tess and Mary for being by my side for the journey, and my partner Matt for your unwavering love and support. I am grateful to you all for your contributions throughout my candidature.

Research output

Publications contributing to this thesis

Published

1. Sincovich, A., Gregory, T., Zanon, C., Santos, D.D., Lynch, J., & Brinkman, S.A. (2019). Measuring early childhood development in multiple contexts: The factor structure and reliability of the early Human Capability Index in seven low and middle income countries. *BMC Pediatrics*, 19(471). doi: 10.1186/s12887-019-1852-5.

2. Sincovich, A., Gregory, T., Zanon, C., Santos, D.D., Lynch, J., & Brinkman, S.A. (2020). Measuring early child development in low and middle income countries: Investigating the validity of the early Human Capability Index. *Social Science and Medicine: Population Health*, 11, article 100613. doi 10.1016/j.ssmph.2020. 100613.

3. Rao, N., Mirpuri, S., Sincovich, A., & Brinkman, S. (2020). Overcoming conceptual and methodological challenges in measuring early childhood development across cultures. *The Lancet Child and Adolescent Health*, 4(5), 352-354. doi: 10.1016/S2352-4642(20)30026-2.

Submitted

4. Sincovich, A., Gregory, T., Lynch, J., & Brinkman, S.A., (2022). Child development at ages 2-5 and cognitive outcomes at ages 6-9 in Lao PDR: Measuring children’s development using the early Human Capability Index. *PLOS Global Public Health* – under review.

5. Sincovich, A., Lam, B., Gregory, T., Rodilosso, N., & Brinkman S.A. (2022). Early childhood education quality and child development in Lao PDR. *Early Childhood Research Quarterly* – under review.

Conference proceedings arising from this thesis

1. Sincovich, A. & Brinkman, S. (2021). A census approach to monitoring child development - informing policy and practice beyond survey samples. *International Journal of Epidemiology*, 50 (Issue Supplement 1), online 2 September. doi: 10.1093/ije/dyab168.612.
2. Sincovich, A., Gregory, T., Lynch, J., & Brinkman S. (2021). Early child development at 2–5 years predicts cognitive outcomes at 6–9 years in Lao PDR: A case for population monitoring using the early Human Capability Index in low and middle income countries. *Journal of Epidemiology and Community Health*, 75 (A6-A7). doi: 10.1136/jech-2021-SSMabstracts.13.

Presentations arising from this thesis

1. Sincovich, A. Tracking progress against SDG 4.2: capturing culturally relevant early childhood development. 7th Conference of the International Society for Child Indicators, Tartu, Estonia, August 2019 (oral presentation, awarded best student presentation).
2. Sincovich, A., Gregory, T., Zanon, C., Santos, D.D., Lynch, J., & Brinkman, S. Measuring early childhood development across contexts: The factor structure of the early Human Capability Index (eHCI) across seven low and middle income countries. 7th Conference of the International Society for Child Indicators, Tartu, Estonia, August 2019 (oral presentation).
3. Sincovich, A. & Brinkman, S. Measuring early child development across contexts: the early Human Capability Index. Symposium on Quality of Early Childhood Education, Child Development and Interventions. China Institute of Educational Finance Research, Peking University. Beijing, China (virtual attendance). November 2020 (invited oral presentation).

4. Sincovich, A., & Brinkman, S. Early child development at 2–5 years predicts cognitive outcomes at 6–9 years in Lao PDR: A case for population monitoring using the early Human Capability Index in low- and middle-income countries. Social Science and Medicine Annual Scientific Meeting, Virtual, September 2021 (oral presentation).

5. Sincovich, A., & Brinkman, S. A census approach to monitoring child development: informing policy and practice beyond survey samples. International Epidemiological Association's World Congress of Epidemiology, Virtual, September 2021 (oral presentation).

Thesis overview

This thesis is presented by publication and includes seven chapters. Chapter 1 provides an introduction, outlining the background to and rationale for the research presented in this thesis. This chapter includes a review of existing population measures of early childhood development as well as their limitations when implemented in low- and middle-income countries (LMICs). This summary highlights the challenges of existing approaches to population measurement of early childhood development, and the need for a measurement solution that is better aligned with population measurement ideals, particularly in low resource settings. Here, the early Human Capability Index (eHCI), the measure of focus in this thesis, is described, together with the need to conduct research to establish the psychometric properties of the tool. The overarching aims and four research questions explored throughout this thesis are outlined.

Chapter 2 includes an overview of the data sources and measures utilised, as well as the methodological approaches employed throughout this thesis, including rationale for approaches selected in research studies and how they were undertaken.

Chapters 3 to 6 present the four research studies undertaken to address the research questions posed in this thesis, investigating the reliability, validity, and sensitivity of the eHCI among children aged 2 to 6 years across seven LMICs. These studies are presented as four papers, two of which have been published, and two of which are currently under review. Chapter outlines provide an overview of each research study, its results, and how findings have addressed the relevant research question. Published papers and manuscripts under review, including supplementary materials, are presented in appendices. Additionally, a published commentary paper co-authored by the Candidate is also available in the appendix. This paper presents an argument for the importance of culturally and contextually specific measurement of early childhood development; a theme that is central to both the eHCI and this thesis.

Finally, Chapter 7 includes a discussion of the key contributions the research in this thesis has made to existing evidence regarding population level measurement of early childhood development. Limitations of the thesis are outlined, including how future research may overcome these challenges. The chapter concludes with a discussion of the implications of findings for policy and practice, including how the eHCI may be used to facilitate population measurement of holistic early childhood development in LMICs moving forward.

Chapter 1: Introduction

1.1 Chapter outline

This chapter provides background to and rationale for the research presented in this thesis. It begins by outlining the need to invest in early childhood development, and the critical role population measurement plays in such investment decisions. It summarises existing population measures of early childhood development and their limitations, particularly when applied in low- and middle-income countries. Next, the early Human Capability Index is introduced, including how the tool's design was intended to overcome these challenges. The need for instruments measuring early childhood development to be psychometrically robust is outlined, including descriptions of the aspects of reliability and validity required. The chapter concludes by outlining the aims of this thesis and the four research questions that were explored.

1.2 Background

1.2.1 The importance of investing in early childhood development

Early childhood development, the process of human development from conception to age eight by which children begin to acquire a range of physical, cognitive, language, social, and emotional skills, is considered to be the most important phase in life as it lays the foundation for later learning, behaviour, achievement, health, and wellbeing[1-3]. When children thrive in their early years, fuelled by adequate health care and nutrition, responsive and supportive caregiving, opportunities for early learning, and protection from harm, they are more likely to reach their full developmental potential as adults and successfully participate in economic, social, and civic life.[4, 5] Adversity and negative experiences in early childhood on the other hand, can disrupt brain development and early learning, which can lead to poorer outcomes that persist throughout the life course including lower educational attainment and reduced earning potential, as well as increased risk for poor mental and physical health.[6-8]

Early childhood is a period of great opportunity and vulnerability, and unfortunately it is the case for millions of children worldwide that the window of opportunity for healthy early development is missed. Indeed, an estimated 250 million or 43% of children under the age of five in LMICs are at risk of not reaching their potential for physical, cognitive, or social and emotional development due to a range of adversities which negatively impact their later outcomes, including chronic malnutrition, poverty, and disadvantage.[9] Further, disparities in children's health and development outcomes among world regions are large, with the most recent evidence demonstrating that children living in South Asia and Sub-Saharan Africa have the poorest chance of both surviving and achieving their developmental potential.[10]

Optimal health and development in childhood and throughout the life course is central to the formation of human capital.[10] Therefore, investment in young children will not only benefit the children of today but will have a direct impact on the human capability, stability, and prosperity of nations in the future. Failing to support populations to reach their developmental potential has been equated to the loss of a quarter of the average adult earning potential later in life which not only has harmful long term effects at the individual level, but also constitutes considerable drain on a country's resources and contributes to the intergenerational cycle of poverty, inequality, and social exclusion that affects all countries.[5, 9, 11, 12] The need to invest in early childhood development to minimise such burden and ensure all children are afforded the opportunity to reach their potential is both paramount and irrefutable.

Encouragingly, a large body of research demonstrates that interventions that target the right combination of risk factors can modify children's physical, cognitive, social and emotional development and thus improve outcomes across the life course.[5, 13] Indeed, for almost two decades, economists have argued that investment in early childhood development is the most powerful investment a country can make. For instance, evidence examining the long term benefits of an early childhood program targeting disadvantaged children and families, following participants until their mid-30s, demonstrated a rate of return of 13.7% per annum with a ratio of cost to benefit of 7.3.[14] These benefits manifest in improved health and education outcomes, increased individual earnings of up to 25%, lower crime rates and overall, workforces that are better equipped to face the future challenges of a global, digital economy.[15, 16] Further, investments in children's early years have the largest positive impacts among those who are most disadvantaged.[17, 18] In this way, ensuring adequate supports in the first years of life can help to "even the playing field" and reduce inequalities in children's outcomes. Together with multisectoral supports for all children that continue throughout later childhood and into adolescence, optimal outcomes can be realised.[13]

As a result of decades of research and advocacy, there exists increased awareness amongst governments, service providers, and communities regarding the critical importance of children's early years and the investment required to appropriately support early development. Indeed, in some parts of the world governments are providing their children with adequate health care, nutrition, stimulation, and protection, and as a result are reaping the benefits.[15] However, it remains difficult to convince many governments of the value of investing significantly in early childhood development. The absence of comprehensive child development data highlighting the need for action is a key obstacle to such investment. A more nuanced understanding of the capabilities that are strengthened by early childhood interventions, the mechanisms through which this impact occurs, as well as the dose and quality of supports required – all of which can be acquired via interrogation of child development data – is essential for bolstering confidence in policy makers to invest in the early years.[19] Further, the majority of child development research has been conducted in a small number of high income countries, leaving significant gaps in our knowledge regarding the state of children's development in diverse LMICs.[20] Together, these hurdles pose a significant challenge in promoting adequate investments in early childhood development.

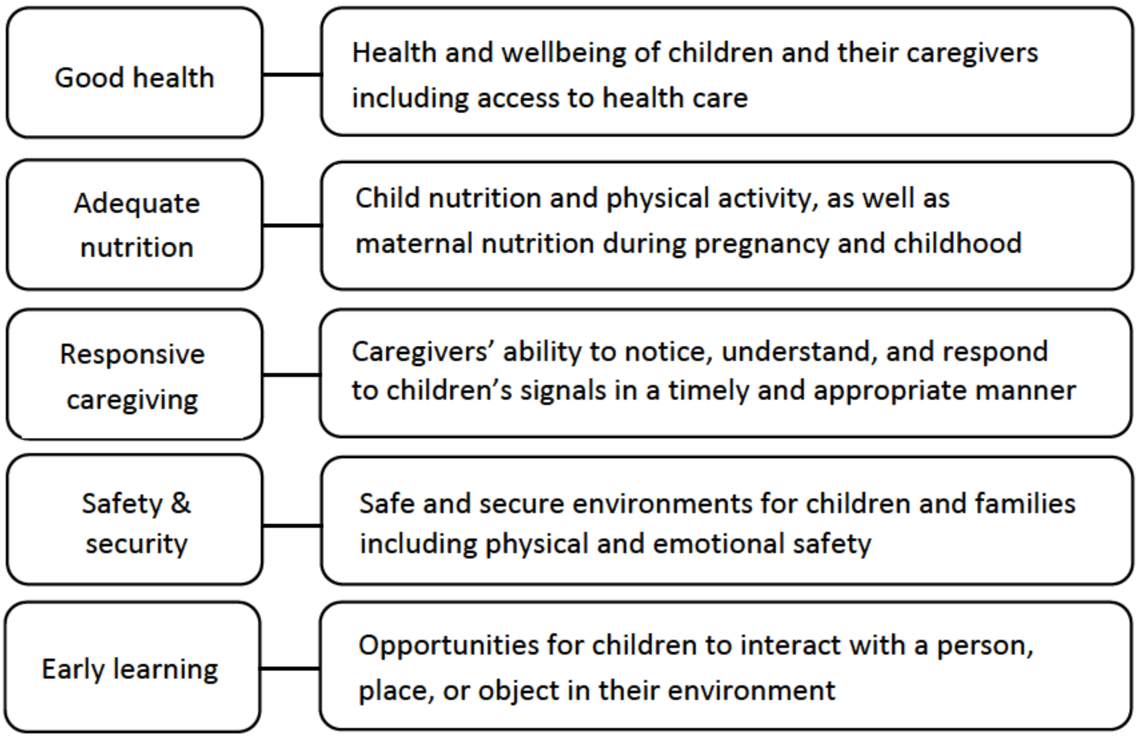
1.2.2 A shift in focus from surviving to thriving

Historically, countries have collected data on indicators including infant and child mortality, breastfeeding, immunisation, and stunting as a means of monitoring children's outcomes.[21] Such monitoring has highlighted great advances in child health and survival outcomes over time. For instance, the global under-five mortality rate shifted from 428 children per 1,000 live births in 1800, to 409 in 1900, 247 in 1950, and then 93 in 1990.[22] The Millennium Development Goals (MDGs), established in 1990 to address disease, poverty, and hunger, among other crises, for the world's poorest, further spurred global action to save and improve the lives of millions of children.[23] The eight goals included a child mortality target which called for a reduction in the under-five mortality rate by two thirds, between 1990 and 2015. While not all countries met this target, these efforts helped to reduce global child mortality rates by greater than half, from 1 in 11 children dying before age five to 1 in 24 children,[24] and demonstrated how collective, global action can have a significant positive impact on society. More recently, in 2020, the global under-five mortality rate was 37.[25] Similar patterns have been observed for other child health indicators, such as exclusive breastfeeding and immunisation rates, albeit improvements are less pronounced.

With these advances in basic child health and survival (noting that not all world regions have experienced such advances equally), the global agenda has in recent decades, shifted from a focus on surviving to a focus on thriving. Indicators of countries' social and economic development have shifted from rates of child mortality to rates of children achieving their developmental potential. As described in the 2016 Lancet Series, Advancing Early Childhood

Development, "children reach developmental potential when they acquire developmental competencies for academic, behavioural, socio-emotional, and economic accomplishments".[9] The Nurturing Care Framework for Early Childhood Development, developed by the World Health Organization (WHO), United Nations Children's Fund (UNICEF), and World Bank and launched in 2018, presents the factors that influence children's acquisition of such competencies and therefore conditions and inputs required to thrive.[5] Specifically, five interrelated components form nurturing care and are needed for children to reach their full developmental potential, including good health, adequate nutrition, safety and security, responsive caregiving, and opportunities for early learning (Figure 1.1).[26]

Figure 1.1. Nurturing Care Framework for Early Childhood Development



Note. Information sourced from World Health Organization[26]

This shift in focus, from surviving to thriving, is reflected in global endorsement of the United Nations Sustainable Development Goals (SDGs), building on the preceding MDGs. Ratified in 2015, the SDGs focus on a multitude of intertwined factors that are needed to promote global equity, specifying targets to end poverty, mitigate inequality, and protect the planet for a better future[27]. Embedded within the goals are targets related to children's malnutrition, mortality, experience of violence, and early learning; targets that constitute a global agenda for early childhood development.[15] SDG 4.2 states that by 2030, countries must ensure that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education. To track progress against this target, countries are required to monitor (i) the proportion of children aged 24-59 months who are developmentally on track in health, learning and psychosocial wellbeing, by sex, and (ii) participation rates in organised learning one year before the official primary entry age, by sex. While many countries monitor national participation in early childhood education enrolment[28], few track the status of children's development. Tracking progress toward SDG 4.2 calls for population level measurement and monitoring of early childhood development. Indeed, early childhood development contributes to several SDG targets, including those related to health, gender equity, and poverty reduction, and thus population monitoring of children's early development is key to supporting progress toward the broader Sustainable Development Agenda.

Spurred by this global agenda together with burgeoning evidence regarding the value of investing in children's early years, leading international organisations in early childhood have highlighted the need for services that promote the development of capabilities in children that will enable them not only to survive, but to thrive.[29] In turn, the use of instruments that measure progress towards such goals is being promoted, with focus shifting toward a holistic approach.

1.2.3 Holistic early childhood development

As described in the comprehensive toolkit for measuring early childhood development in LMICs developed by the World Bank[29], domains of early childhood development typically include cognitive, language, motor, social-emotional, and pre-academic (i.e., early literacy and numeracy) skills. This section provides a brief overview of these developmental domains and what constitutes holistic early childhood development, before moving to discussion of measurement of holistic development. Examples of how skills across these domains are exhibited as children grow from infancy through to preschool-age are provided in Table 1.1.

Table 1.1. Skills captured in domains of early childhood development

Domain	0-2 years	3-5 years
Cognitive	Early understanding of numerical concepts (e.g., meaning of one or two, or big or small) and problem solving (e.g., stacking objects).	More complex understanding of numerical concepts and problem solving (e.g., meaning of more and less, sorting by colour and shape).
Language	Begins with babbling and gesturing, while words and sentences typically emerge in the first two years.	Production and understanding of words, ability to tell stories and identify letters.
Motor	Fine motor skills include picking up and holding objects. Gross motor skills include learning to walk.	Fine motor skills include holding a pencil to write and draw. Gross motor skills include jumping and catching.
Social-emotional	Exploring objects and places, initiating and responding to social interaction, and strategies for dealing with negative emotions.	Getting along with others, self-regulation (e.g., emotional and/or behavioural control), and identification of emotions in self and others.

Cognitive skills include the processes by which knowledge is acquired including abilities such as problem solving and memory.[30] Language skills include the ability to both understand and express verbal communication.[31] Dependent on culture (e.g., if a culture has a history of written language) vocabulary may or may not be a good indicator of children's language development.[32] This highlights the important influence of culture on children's development, which is discussed in Section 1.2.6. Motor skills include the ability to control and coordinate gross movements of arms and legs and fine movements of the fingers.[33] Social and emotional skills include social competencies and interactions, and the regulation of emotional responses.[33]

In the first years of life, social and emotional development is largely driven by relationships and attachment with caregivers.[34] Finally, pre-academic skills refer to foundational abilities required for development of literacy and numeracy among preschool-aged children, which are important for success in the school environment. This includes letter and number knowledge, counting, reading words and sentences, as well as listening comprehension.[29]

Although definitions can vary, these aspects or domains of development often feature in frameworks and measures of early childhood development internationally. When seeking to measure children's development, capturing holistic development is important as skills across domains overlap and influence mastery of one another.[35] For instance, children's language development and verbal communication can influence other domains spanning literacy and numeracy, as well as social and emotional skills.[36] This list of domains is not exhaustive and there are several other aspects of development that feature in some measurement tools (e.g. executive function, personal/adaptive skills, and approaches to learning). However, the domains summarised above tend to feature most commonly across measures and help to depict what typically constitutes holistic early childhood development.

1.2.4 Measuring early childhood development

Measurement is essential in understanding whether children are developing as expected and to their full potential. Measurement of early childhood development is typically undertaken for three key purposes: screening and diagnostic assessment, program and intervention evaluation, and population monitoring. While the research presented in this thesis is focused on population level measurement, it is important to first highlight differences in measurement approaches.

<u>Screening and diagnostic assessment</u>

Screening tests are used to identify children at risk of developmental difficulties or delay, with the objective of referring children who screen positive for risk to a qualified clinician for further assessment and (potential) diagnosis.[37] Screening tests are short and quick to administer, typically include motor, cognitive, and language domains, and classify children into categories such as delayed, at risk for delay, or developing typically, based on normative data.[29]

The most widely used screening test globally is the Ages and Stages Questionnaire (ASQ).[38] The third and most recent edition, the ASQ-3, is used to measure development among children aged 1-66 months.[39] Domains of development screened by the ASQ include communication, fine and gross motor skills, problem solving (i.e., cognitive development), and personal-social development. The instrument includes several questionnaires with items focused on different stages of development (i.e., a different set of questions is completed dependent on child age). Caregivers complete the short questionnaire in 10-15 minutes, which is then scored by a trained assessor. An example item in the 18-month questionnaire measuring children's fine motor development is "Does your child stack a small block or toy on top of another one?" (response options yes, sometimes, not yet). Based on empirically derived cut-off scores, results on each domain can be categorised as on schedule, for monitoring, or for further assessment.[39] Developed in the United States, the ASQ has been widely validated as a screening instrument in other high-income countries, as well as in various LMICs.[40-44]

Program and intervention evaluation

Program evaluation is designed to evaluate the impact of an intervention on children's outcomes. Measurement is implemented to assess the skills or capabilities that are intended to be impacted by the intervention. Therefore, program evaluation tools may be broad (i.e., holistic) or narrow in terms of domains assessed, dependent upon the intervention.

An example of a developmental assessment tool often utilised for the purposes of program evaluation is the Bayley Scales of Infant and Toddler Development (originally Bayley Scales of Infant Development).[45] The third edition, Bayley-III, measures cognitive, language, social-emotional, motor development, and adaptive behaviour among children aged 1-42 months.[46] Scales are administered via a combination of child behavioural assessments and caregiver questionnaire. Administration ranges from 30-60 minutes and is undertaken by trained individuals with experience in developmental assessment (e.g., psychologist, paediatrician). Items differ with child age, for example, the cognitive scale measures interest and attention among infants, exploration and problem solving among toddlers, and pretend play and object categorisation (e.g., colour, size) among preschool-aged children. While the Bayley-III can be used to identify developmental delay, it differs to typical screening tests in that it collects more comprehensive information through a greater number of items and therefore has greater variability in scores, that can be used to track development and assess impacts of intervention.[47] The Bayley-III was developed in and is primarily used in the United States, due to standardised scoring based on normative data in this setting.[46] Although the tool has been well validated in the United States, psychometric results for its application in LMICs are mixed.[48-50]

Population monitoring

Population monitoring of child development is essential in understanding how children are developing within populations (e.g., community, region, country).[51] It seeks to identify broad patterns in development among children across a whole population. Such monitoring also enables identification of variation or inequalities in development among specific sub-populations of children (e.g., different genders or geographical groups). This information can be collected through either a census or representative sample approach. Information generated through population monitoring can be used to track shifts in development within and across populations over time and across settings.[51] Such measurement aims to inform policies (e.g., preschool provision) and supports (e.g., literacy interventions) to promote community, regional and/or national-level early childhood development outcomes. Subsequently, population measures are designed to be feasible for use at scale.[29]

The Early Development Instrument (EDI) is a Canadian-developed population level measure of development for use among children aged 3 to 6 years.[52, 53] The tool includes approximately 100 items measuring development across five domains that are important for children's readiness for the school environment, including physical health and wellbeing, emotional maturity, social competence, language and cognitive development, and communication skills and general knowledge. An example item in the communication skills and general knowledge domain is "How would you rate this child's ability to tell a story?" (response options poor/very poor, average, very good/good). Teachers take about 15 minutes to complete the EDI for each child in their class (the same items are completed for all children), and aggregated results report the percentage of children who are classified as vulnerable, at risk, or on track in their development at various levels (e.g., school, community). The EDI has been used in more than 20 countries, namely Canada and Australia, but has had limited application in LMICs.[54-57]

In addition to different measurement purposes, measurement of early childhood development is typically conducted via three key collection methods: (i) direct assessment, where children are asked to complete a series of tasks or activities (e.g., naming or counting objects), led by an assessor in a one-on-one setting; (ii) adult report, where respondents who know the child (e.g., caregiver, teacher) complete a series of questions about the child's abilities, based on their knowledge of the child; and (iii) direct observation, where assessors observe, document, and score children's behaviour as it occurs in a natural setting in real time.

The purpose of assessment will determine the type of measurement instrument selected to be implemented, and so too will the age range of target children, as well as any financial and logistical constraints present. Each type of assessment, as well as specific measurement instruments, have a range of advantages and disadvantages that must be considered before selection. The section to follow describes the value of population level measurement, as is the focus of this thesis, and presents a review of currently available population level measures.

1.2.5 Population measurement of early childhood development

Data collected via population level early childhood development measurement create a range of opportunities.[21, 29, 58, 59] Such data provide countries with a baseline regarding the state of child development across the population. As aforementioned, this can be used to identify patterns in children's development across the population, as well as highlight inequalities in outcomes amongst specific population sub-groups (e.g., children from different ethnicities or socioeconomic position). Data can also be used to explore specific areas or domains of development that need strengthening (e.g., children's motor development or early literacy skills). In this way, data can be used to inform policy regarding where further investments are needed to provide all children with the opportunity to achieve their developmental potential.

Population level data can also be used to track changes or population shifts in development among cohorts of children over time. For instance, monitoring the development of children upon entry to primary school (approximately 5 years of age), as is conducted triennially using the EDI in Australia (Australian Early Development Census; AEDC).[60] Such monitoring enables investigation of whether policies and interventions are achieving their goals of improving development for successive cohorts of children. Indeed, only by monitoring progress over time can policy makers, service providers, and communities determine if their efforts are making a difference. Similarly, tracking development among populations over time can enable greater understanding of why some children are doing better than others, through identification of the factors that play an important role in supporting children to reach their developmental potential (for instance, investigation of type, dose, timing, and quality of early childhood education). Importantly, mechanisms driving positive developmental outcomes are likely to vary across different sub-groups of children (e.g., language background, socioeconomic position, disability or additional needs status).[17] Identifying what supports work for whom in this way can help to inform how services and supports should be targeted so that these positive impacts can be maximised across populations.

At a broader level, population level data allow for international monitoring, such as that required for the Sustainable Development Agenda, as well as comparison of children's development within and across countries and world regions. Such comparison investigations can improve understanding of how country-level and region-level policies and service delivery models impact children and their families. Examples might include policies regarding parental leave, antenatal and postnatal supports, or early childhood education and care.

There are a handful of existing population level measures of holistic early childhood development used for a range of purposes in various countries and world regions. As part of this thesis, a review of existing measures (as at August 2022) was conducted with a focus on instruments designed for use among children aged 3 to 5 years.

While early childhood development includes the period from conception to age eight, given the growth in children's development that occurs during this period, a single measure cannot be validly applied across the entire age range. When the Sustainable Development Agenda was ratified in 2015, SDG 4.2.1 incorporated tracking the development of children aged 0-5 years.[61] The indicator was classified as tier three, meaning no internationally established methodology or standards were available.[62] In 2020, led by experts in early childhood development internationally, the indicator was revised to be restricted to children aged 24-59 months, no longer including children aged below two years. This process of reshaping SDG 4.2.1 in a way that enables feasible measurement reflects the challenges in identifying or developing a single population level measure that can address the needs originally set out by the indicator.[62, 63] That is, a measure that captures development among children from birth to five years.

Against this background, the age range of 3 to 5 years was selected as the focus when reviewing holistic population measures of child development in this thesis for several reasons. Measures designed to capture children's development in the first few years of life (i.e., 0 to 2 years) are focused on mastery of developmental milestones and are typically implemented to screen for delay, rather than holistic population monitoring (though there are recently developed exceptions for use in LMICs[64, 65]). Measures designed for use among school-aged children (i.e., 6 years and above) tend to focus on assessment of the skills acquired through education (e.g., literacy and numeracy) and thus typically are not holistic in terms of the domains they capture. Further, an important consideration for population measurement is feasibility (see Section 1.2.6). This includes consideration of how measures can be implemented in a way that allows access to all children in the age range of focus (if using a census approach), or a representative sample of children (if using a sampling approach). For instance, in many countries, preschool (the year before primary school, typically 4 to 5 years dependent on country) is a feasible population capture point as most children attend. Collecting population data for younger children is often more difficult due to lack of a feasible and pragmatic touch point between child and maternal health services in early infancy and preschool at age 4 to 5 years.

Below, instruments identified through the review are presented in chronological order from time of development, including a description of their purpose, target age range, domains measured, administration methods, and utilisation (see Table 1.2 for an overview).

Table 1.2. Overview of existing population measures of early childhood development

Instrument	Age range	Domains measured	Administration	Completion time	Fees	Adaptable
EDI	3-6 years	Physical health and wellbeing, social competence, emotional maturity, language and cognitive development, and communication skills and general knowledge	Teacher report	15 minutes	Yes	Yes (limited)
MICS ECDI	3-4 years	Literacy and numeracy, social-emotional, approaches to learning, and physical development	Caregiver report	<2 minutes	No	No
EAP-ECDS	3-5 years	Cognitive, language, motor, socio-emotional development, health, cultural knowledge, and approaches to learning	Direct assessment	> 60 minutes (LF), 45 minutes (SF)	Yes	No
IDELA	3-6 years	Emergent language and literacy, emergent numeracy, motor development, social-emotional development	Direct assessment	35 minutes	No	Yes (limited)
MELQO MODEL	3-6 years	Literacy, numeracy, socio-emotional, executive function, and fine motor skills	Direct assessment, caregiver report, teacher report	25 minutes (DA), 25 minutes (CR), 15 minutes (TR)	No	Yes
AIM-ECD	4-6 years	Early literacy, early numeracy, executive functioning, and social-emotional competencies	Caregiver report, direct assessment	3 minutes (CR), 30 minutes (DA)	No	No

Note. EDI = Early Development Instrument, MICS = Multiple Cluster Indicator Survey, ECDI = Early Child Development Index, EAP-ECDS = East Asia Pacific Early Child Development Scales, IDELA = International Development and Early Learning Assessment, MELQO = Measuring Early Learning and Quality Outcomes, MODEL = Measure of Development and Early Learning, AIM-ECD = Anchor Items for Measuring Early Childhood Development, DA = direct assessment, CR= caregiver report, TR = teacher report, LF = long form, SF = short form. Completion time of AIM-ECD estimated by authors of the instrument (i.e., actual completion time has not been published).[66]

Early Development Instrument (EDI)

Established in 2000, the Canadian-developed Early Development Instrument (EDI) described earlier represents the first effort to measure children's holistic development at a population level. Designed for use among children aged 3 to 6 years, the instrument includes about 100 items that measure development across five domains including physical health and wellbeing, emotional maturity, social competence, language and cognitive development, and communication skills and general knowledge.[52, 53] Domains include a combination of binary items as well as 3-point response scales (e.g., poor/very poor to good/very good). The EDI has been implemented across individual provinces in Canada via a community-led approach.[67] Kindergarten teachers complete the instrument for each child in their class, which takes about 15 minutes per child. Aggregated results report the percentage of children who are classified to be vulnerable, at risk, or on track in their development, on domains as well as overall, at either the school, community, or provincial level. The EDI is a copyrighted instrument and therefore has costs associated with its use, for licensing and scoring, as well as assistance with adaptation, with all country adaptations requiring maintenance of core items.[68] Several adaptations have been carried out and the EDI has been used across more than 20 countries, including the United States, Indonesia, Vietnam, and, Australia, whereby the EDI is implemented nationally every three years as a census of early childhood development.[60]

Early Child Development Index (ECDI)

The most widely utilised population measure of early childhood development is the Early Child Development Index (ECDI), implemented since 2009 as part of UNICEF's Multiple Indicator Cluster Survey (MICS).[69, 70] The ECDI includes a set of simple indicators designed to generate nationally representative, globally comparable data on early childhood development. For this reason, these indicators are not adapted across countries. The tool includes 10 binary (yes/no) items that capture key developmental milestones for children aged 3-4 years. Caregivers rate their child's behaviour on four domains of development: literacy and numeracy, social-emotional, approaches to learning, and physical development. Children are classified as developmentally on track on each domain, and if on track on three out of the four domains, they are considered developmentally on track overall.[70] The ECDI has been implemented in more than 80 countries, including LMICs, through the MICS as well as other national household surveys throughout the past decade.[70, 71]

With ratification of the SDGs in 2015, the ECDI was proposed as the measure for tracking progress against Indicator 4.2.1 focused on the development of children aged 24 to 59 months in health, learning, and psychosocial wellbeing. However, the tool required revision to adequately capture the domains and age range specified by this indicator, which led to development of the ECDI20130.[72-74] The ECDI2030 includes 20 items that, together, determine the percentage of children aged 24 to 59 months developmentally on track in three domains: health, learning, and psychosocial well-being. Administered using the same method as the ECDI, most items are binary (yes/no), while some of the psychosocial wellbeing items include a 5-point response scale (e.g., how often children exhibit a behaviour, from daily to never). Since recognition of the ECDI2030 as the measure for global monitoring on SDG 4.2.1 in 2019, integration of the tool into national surveys commenced, with the ECDI2030 deemed a standard inclusion in the MICS7 from 2023 onward.[75, 76]

<u>East Asia-Pacific Early Child Development Scales (EAP-ECDS)</u>

The East Asia-Pacific Early Child Development Scales (EAP-ECDS) were developed in 2010 with support from UNICEF, the Asia-Pacific Regional Network for Early Childhood (ARNEC), and the Open Society Foundations. There existed no measure that considered the cultural and contextual diversity within and across the East Asia-Pacific, and so the EAP-ECDS was developed to provide a culturally relevant assessment of children's development in the region.[77] The EAP-ECDS is a direct assessment of holistic development, designed for use among children aged 3 to 5 years at a population level. The scales include 85 items, covering domains including cognitive, language, motor, and socio-emotional development, health, cultural knowledge, and approaches to learning.[78] Assessments are conducted in individual face-to-face sessions by assessors (e.g., teachers, program staff) trained in the use of the tool. Considering the time required to complete the tool (over 60 minutes), a short form was developed in 2016/17. The EAP-ECDS short form is implemented in the same way and captures the same developmental domains as the long form, through administration of 33 items which typically takes 45 minutes to complete.[79, 80] For both long and short forms, scores across domains as well as an overall score can be reported. Use of the scales requires paid training, adaptation, and technical support, as well as data analysis and dissemination support by members of the EAP-ECDS team. The EAP-ECDS short form was first implemented in Myanmar in 2017 and both forms are available for use among nationally representative samples with a focus on informing policy for child development.[79]

International Development and Early Learning Assessment (IDELA)

Developed by Save the Children in 2011, the International Development and Early Learning Assessment (IDELA) is a population level, play-based direct assessment tool designed to measure holistic learning and development among children aged 3 to 6 years.[81] The IDELA was developed to support early years program evaluation and improvement in Save the Children sites (primarily LMIC settings), and provide evidence regarding children's learning and development across countries to promote best practice in early childhood education and care. The tool is comprised of 22 items/tasks in four domains: emergent language and literacy, emergent numeracy, motor, and social-emotional development. Additionally, optional items can be used together with this core set, covering domains such as executive function, learning approaches, and health and hygiene.[81] Assessment is undertaken by trained assessors which takes approximately 35 minutes for each child. Scores for each domain, as well as a total score (combines all core domains) can be reported. The IDELA toolkit, including training and assessment materials, as well as resources to support analysis and dissemination, are available to any organisation for free, given partner registration which requires sharing of data collected, translated and adapted materials and so on. The IDELA is thus a global, open-access tool, designed to be implemented across culturally diverse settings. It is recommended the tool should be adapted (i.e., rephrasing or rewording of items) before use in a new setting to ensure children's understanding of items. The IDELA has been used in almost 80 countries for program evaluation, multi-sectoral studies, and national monitoring.[81, 82]

Measurement of Early Learning and Quality Outcomes (MELQO)

Developed in 2014, the Measuring Early Learning and Outcomes (MELQO) initiative was created through partnership among the Brookings Institution, the World Bank, UNICEF, and United Nations Educational, Scientific and Cultural Organization (UNESCO), with the aim of facilitating measurement of children's early learning and development in LMICs. MELQO includes two modules: (i) the Measure of Development and Early Learning (MODEL; a child-direct assessment and teacher/caregiver interview), and (ii) the Measure of Early Learning Environments (MELE; a classroom observation instrument and head teacher and teacher interviews).[83] Modules were designed for use at the population level to provide data that inform early childhood development policy across countries. Further, MELQO tools were intended to be culturally adapted and aligned with national standards prior to implementation in any new context. In this way, the initiative seeks to produce globally comparable data while also incorporating locally relevant measurement to be used for program evaluation.[83, 84]

Specifically, the MODEL direct assessment was designed for use among children aged 4 to 6 years and captures children's early literacy, early numeracy, social-emotional development, and executive function. Accompanying caregiver and teacher surveys collect information regarding children's behaviour and learning at home and school. While the MODEL is unlicensed and therefore free to use, training of staff to administer the tool/s is required to be led by an accredited individual (i.e., someone who has undertaken master training with the MELQO team). Paid support with adaptation is also available. MODEL results can be reported as domain scores as well as overall scores, or item level reporting according to national standards (e.g. if a child can write their name).[85] Since its development, the MODEL has been used in several LMICs including Kenya, Tanzania, Lao PDR, Ethiopia, and Pakistan.[84, 86, 87]

Anchor Items for Measuring Early Childhood Development (AIM-ECD)

The World Bank's Anchor Items for Measuring Early Childhood Development (AIM-ECD) was most recently developed in 2021. AIM-ECD is a core set of items that measure early childhood development among children aged 4 to 6 years.[66] Core items were identified through a process of analysing data from existing measures of child development from across 12 LMICs to identify items with robust psychometric properties (i.e. reliable and valid) across diverse settings. AIM-ECD includes 20 caregiver report items and 84 direct assessment items spanning domains including early literacy, early numeracy, executive functioning, and social-emotional competencies.[88] The AIM-ECD, focused on the 4-to-6-year-old age range, is intended to be used alongside global measures including the Global Scale for Early Development (GSED; 0 to 2 years)[64] and the ECDI2030[72] (2 to 4 years). While details regarding initial implementations of the AIM-ECD are not yet publicly available, authors of the tool have described that the addition of data and items from other measures to further develop the tool, as well as field testing are among next steps for the measure.[66] It is envisioned that freely available items as well as training, data collection, analysis, dissemination materials (currently under development), will serve as a starting point for linkage across different measurement tools and facilitation of scale up of early childhood measurement globally.[66, 88]

A common feature of all measures described is that they were designed through a process that involved drawing on existing tools. For instance, the IDELA was created using items drawn from the ASQ, Bayley Scales, and the EDI, among other tools.[81] Similarly, both the ECDI and MELQO MODEL were developed through drawing items from the EDI, among other tools.[70, 87] Indeed, creation of many child development measurement tools has taken an approach that incorporates identifying items working well (i.e., psychometrically robust) in a particular setting/s, and then using these (with or without additional items) to achieve a different purpose (e.g., expand measurement to additional domains, or in a different country/region).

In addition to measuring children's development, many of these tools also collect information on aspects such as physical health and nutrition (e.g., anthropometrics) and early learning opportunities (early childhood education, the home learning environment), in line with the Nurturing Care Framework.[5] In doing so, measures seek to capture various inputs for children's development, which can be used to investigate the mechanisms working to promote development, as well as how outcomes could be improved through intervention on such inputs.

1.2.6 Limitations of existing population measurement approaches

There are several population level tools currently used for a range of purposes in various countries and world regions. As Fernald and colleagues[29] discuss, though many early childhood development measures work well in different settings for different purposes, no tool meets all criteria of an ideal measure. Rather, the selection of any assessment requires a compromise among different priorities and measurement ideals, determined by the purpose of measurement (e.g., population monitoring, program evaluation), the age range of target children, and any financial and logistical constraints which are often present in LMICs.

This sentiment is applicable to the existing population measures described. That is, various characteristics of instruments currently in use, including their costs, the training required prior to administration, the time they take to administer, how they are administered (e.g., direct assessment adult report), the aspects or domains of development they capture, their sensitivity to capturing variation in children's abilities, and their applicability across diverse settings, all constitute significant barriers to their utilisation in low resource settings. These limitations are described below, highlighting some of the key challenges in using existing instruments.

Feasible for use at scale

A population measure of early childhood development needs to be cost-effective for use at scale, therefore fees to use the tool (including licensing, training, data analysis, and dissemination of results), the level of administrator training required, and administration time must all be minimal.[29] For example, licensing costs of the EDI payable to instrument creators or costs associated with having EAP-ECDS team members travel to a country to provide training and implementation support, present a significant financial investment and does not deem these measures feasible for use at a large scale in low resource settings. Administration method is also an important consideration. For instance, while direct assessment measures are typically argued to produce scores with less bias than those through adult report measures[89], they require highly trained staff to administer measurement, are more time consuming to conduct, more costly to implement, and thus are not feasible for use across large populations.[90]

Sensitive to variation in children's development

Although adult report measures are generally more cost-effective as they are quick and simple to administer and do not require developmental expertise[91], often such tools are based on developmental milestones with a pass or fail outcome, lacking the sensitivity required to adequately detect variations in children's development, as well as changes in development over time.[92] Population measures should cover a range of domains to capture children's holistic development, as well as levels of ability, from low to high, as is appropriate to the target age range. This intends to ensure scores on a measure have the sensitivity to detect both variation in development between children, as well as shifts in patterns of capabilities among cohorts of children over time, which is an essential aspect of population monitoring for informing policy and targeting supports.[92, 93] For instance, to ensure feasibility for use at scale, the original MICS ECDI relies on a short set of 10 items designed to identify broad trends in children's development across four domains (literacy and numeracy, social-emotional, approaches to learning, and physical development). As a result, the instrument captures limited information within these domains of development which hinders its ability to generate meaningful information required to inform directing of investments in promoting children's outcomes.[29]

Culturally adaptable and locally relevant

The measurement of holistic child development is complex as it is influenced by culture, language, and theory, and thus the concept of what good child development looks like, when it emerges, and how it is valued, will vary across contexts.[94, 95] As a result, there are few internationally accepted measures of early childhood development. This is most often due to concerns that tools developed for use in high income countries are not appropriate for use in LMICs due to differences in culture and context and thus early learning frameworks, as well as a lack of consensus regarding the constructs of development to be measured.[78] This has been an ongoing challenge in the child development field, as the majority of research has been conducted in high income countries, leaving significant gaps in knowledge regarding the state of children's development, as well as measurement approaches, in diverse LMICs.[20, 93]

Although tools that produce internationally comparable results allow for population monitoring across countries, such instruments are often not well aligned with local culture, context, or early learning and development frameworks, and thus the information they produce tends to have limited utility locally.[96] For instance, the MICS ECDI does not enable countries to adapt or align the global set of items to local expectations for children's development, which reduces relevance to local policy and practice.[29] A measure of early childhood development should reflect the capabilities considered to be important in any given cultural or national framework, as well as be culturally appropriate in terms of administration method and materials.[93] Thus, tools should be adaptable across diverse cultures and contexts so that they not only accurately reflect children's abilities, but also capture locally-relevant and culturally-influenced aspects of development to inform local policy and practice.[93, 97]

These ideal criteria for population measures are in addition to an instrument being psychometrically robust (i.e., reliable and valid).[98] Evidently, there is need for a solution to measuring early childhood development that is better aligned with population level measurement ideals, particularly in diverse, low resource settings. It was against this background that the early Human Capability Index (eHCI) was developed.

1.2.7 The early Human Capability Index

Initiated by the World Bank in 2014, the Pacific Early Age Readiness and Learning (PEARL) Program aimed to pilot interventions to promote children's school readiness and early literacy, sharing learnings to support children and their families across Pacific Island Countries.[99, 100] Impact evaluation of the PEARL Program in Tonga required an early childhood development measurement instrument that captured holistic development, was sensitive to changes in children's development over time to determine effects of the program, was free, quick and easy to implement at a population level, and was applicable to the Tongan culture and context.[99] No existing tool met these requirements, and so the eHCI was developed to facilitate program evaluation via population monitoring of early childhood development.

Designed to capture holistic development among children aged 3 to 5 years, the eHCI was developed with the vision of being feasible for large-scale use in low resource settings, while having the sensitivity to detect variation in children's development and capture change in development over time.[99] The tool was intended to be adapted to local cultures and contexts for purposes including population monitoring, evaluation of early years policies and programs, and longitudinal studies seeking to predict children's future capabilities.[99]

The eHCI and its supporting materials are available for anyone to use free of charge, without licensing, training, or data analysis fees.[101] As described by creators on the eHCI website, the tool requires minimal administrator training and can be completed by an adult familiar with the child (e.g., caregiver, preschool teacher, health professional) in approximately 10 minutes. Thus, the eHCI offers a pragmatic, efficient, and affordable solution to measuring early childhood development across large populations. Designed through a comprehensive process of local expert consultation together with existing theoretical conceptualisations, the instrument measures nine domains of development: Physical Health, Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, Approaches to Learning, Numeracy, Reading, and Writing.[99] Figure 1.2 provides an example of items in each domain, and the full instrument is presented in Paper 1 supplementary materials (Chapter 3). The eHCI captures both positive and negative aspects of how a child is developing, allowing for a holistic approach rather than a focus on pathology or developmental delay, as is the case with a screening or diagnostic assessment tool. In this way, the eHCI places children on a developmental continuum, from poor to optimal development, which seeks to maximise sensitivity to detect changes in development over time and/or through intervention.

Overall, the eHCI was created with the goal of overcoming challenges of other available measures of children's early development. Not all instruments included in the review presented above were available at the time the eHCI was developed. Figure 1.3 presents a timeline of the development of tools, including the eHCI. The tool has since been adapted for use in diverse contexts and implemented to support various early childhood health, development, and education projects across several LMICs (see Figure 1.4).[99, 102-107]

Figure 1.2. eHCI domains and example items

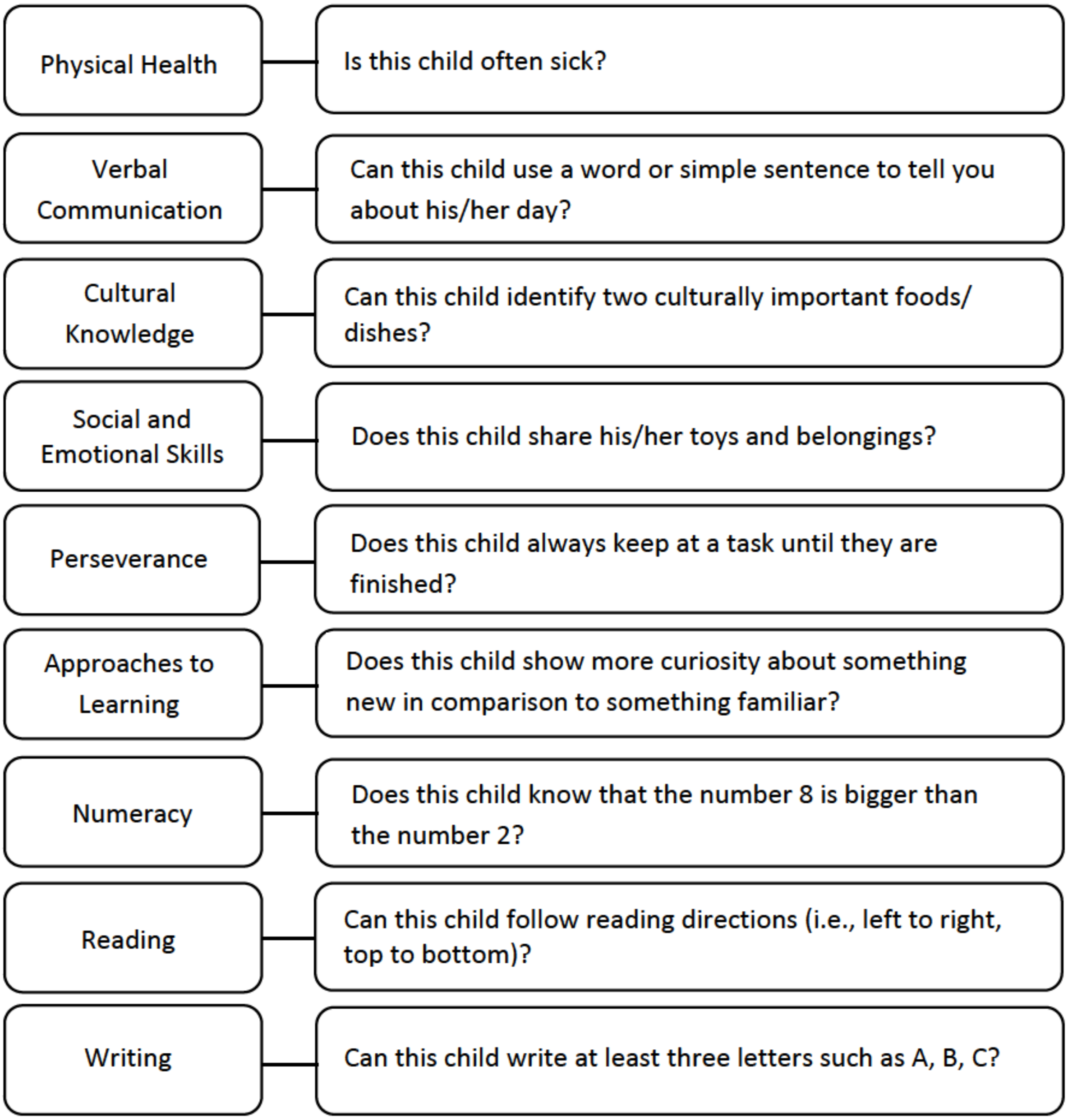


Figure 1.3. Timeline of development of population measures of early childhood development

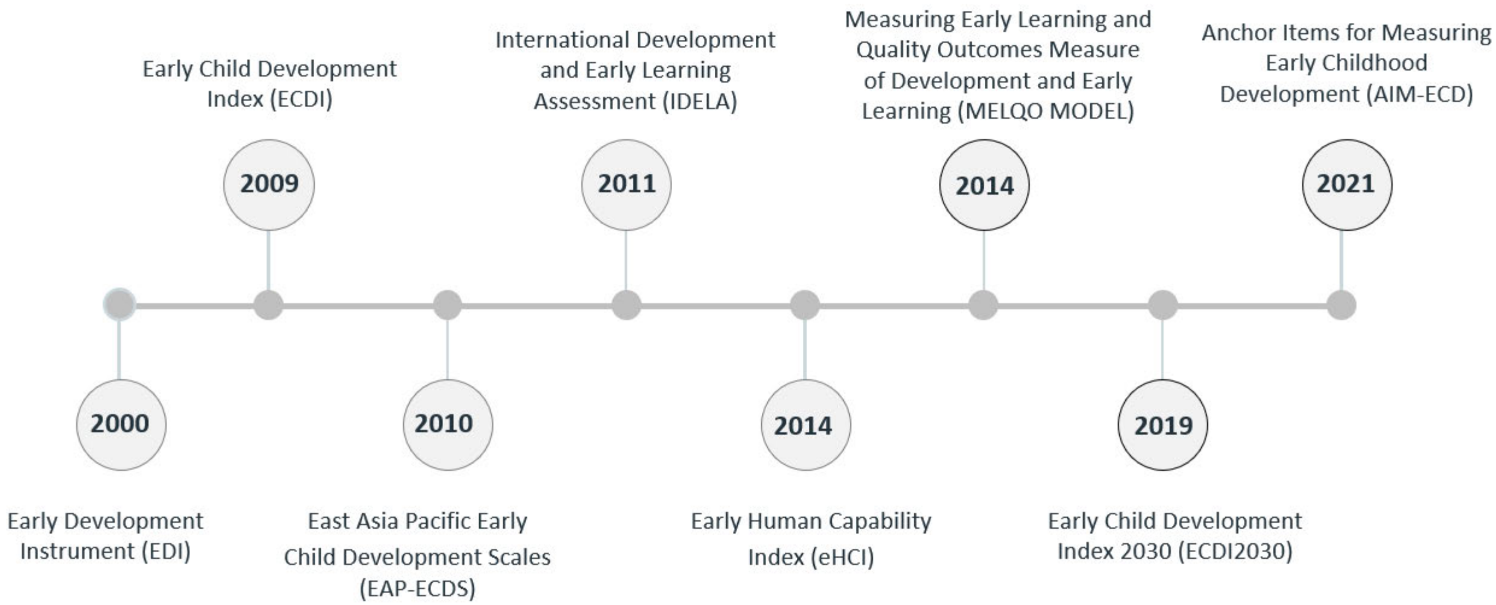


Figure 1.4. Map of countries in which the eHCI has been implemented

Note. Information sourced from the eHCI website[101] as well as the instrument's creator.

For instance, after the eHCI was developed for use in Tonga, it was adapted to the local context in several additional Pacific Island Countries including Kiribati, Samoa, and Tuvalu. The eHCI was then implemented as a national census of children's development in these countries, as part of the World Bank PEARL Program, establishing a baseline of early childhood development with information used to guide early years policies and programs.[108] Another example includes adaptation of the eHCI to Lao People's Democratic Republic (PDR). This adaptation sought to measure impact of the World Bank Early Childhood Education Project; a large-scale randomised control trial that aimed to increase access to quality early childhood education and improve the development of children living in disadvantaged villages.[102] The eHCI was implemented across the study population at three time points over a five-year period, to capture change in children's development as a result of project interventions.

With the eHCI adapted for use in multiple countries and data collected at several time points across diverse settings, the next step in the tool's ongoing process of development is the investigation of its psychometric properties. A measure must be reliable and valid; reliable in that the instrument produces consistent scores across similar measurement conditions, and valid in that the scores the tool produces reflect the constructs of development that it was designed to measure. An instrument measuring early childhood development that lacks in reliability and validity could produce biased scores and lead to ill-informed policy and program decisions. Thus, evaluating the psychometric properties of a measurement tool is fundamental for its future utilisation and effectiveness.[98] There are a range of aspects of reliability and validity that need to be established in order to constitute a psychometrically robust measure (Section 1.2.7 presents a detailed description). Adaptation of the eHCI in each new context was informed by a combination of both theoretical conceptualisation as well as local expert consultation. Therefore, while content and face validity were established in each setting, comprehensive investigation of the reliability and validity of the eHCI was yet to be explored.

The research presented in this thesis sought to undertake this process of investigation with the aim of producing evidence to establish the reliability and validity of the tool. An instrument with the properties of the eHCI that is psychometrically robust across diverse settings could better enable population level measurement and monitoring of children's development, as is required for tracking against SDG 4.2, particularly in low resource settings.

1.2.8 Reliability and validity

This chapter has outlined the opportunities generated by population level child development data for informing policy and supports to ensure all children thrive. However, these opportunities can be realised only when data collected are reliable and valid, and thus accurately capture children's abilities. This section describes aspects of reliability and validity required to ensure a measure of early childhood development is psychometrically robust.

Exploring the psychometric properties of a tool designed for use at a population level bring together disciplines of psychology and public health. Psychometrics, a research discipline concerned with measurement of latent constructs that cannot be directly observed (e.g., intelligence, personality etc), has included development of several measurement theories including classical test theory and item response theory (a type of modern test theory).[109] The purpose of this section is not to comprehensively describe such theories, their assumptions, and key concepts. Rather, this section seeks to provide a description of the most common forms of reliability and validity that need to be established for a population measure of children's early development, based on existing literature in the field. Descriptions below include examples, as relevant to this thesis, to provide background to the research questions to follow.

<u>Reliability</u>

Reliability refers to a measure's consistency in producing scores across time (test-retest reliability), items (internal reliability/consistency), and respondents (inter-rater reliability).[110,

[111] While reliability alone is not sufficient in ensuring a measure produces scores that accurately reflect true ability, it is a necessary component of a psychometrically robust measure.

Test-retest reliability is an indicator of how consistent a measurement score is over time (given the construct measured is considered to be stable over this period).[112] For instance, if the same developmental test is administered to a child several times within the period of a month, a measure with test-retest reliability would produce highly correlated scores on each testing occasion. *Internal reliability/consistency* refers to consistency of an individual's responses across items in a multiple-item measure.[113] Typically, all items in a measure or scale seek to capture the same underlying construct, so an individual's scores across items should be related. For example, a child's responses to five items in a scale measuring literacy should be highly correlated with one another, which would indicate the items capture the same construct and thus the scale demonstrates internal consistency. *Inter-rater reliability* tests the extent to which an individual's scores on a tool are consistent across different administrators or respondents (e.g., teacher, caregiver).[114] For instance, if two different respondents complete a direct assessment with the same child, a measure with good inter-rater reliability would produce highly correlated scores for the child across both administrators.

Validity

Validity refers to the extent to which the scores produced by a measure accurately reflect the attributes or constructs the tool was designed to measure.[114, 115] While definitions of different aspects of validity can vary across disciplines, validity can typically be described in four categories including face validity, content validity, construct validity, and criterion validity.

Face validity is the extent to which a measure appears, at face value, to measure the aspects the tool was designed to measure.[116] For instance, a measure designed to assess children's early numeracy skills would be expected to include items about numerical concepts and simple mathematical problems. Similarly, *content validity* refers to when a measure comprehensively captures all aspects that form the construct of interest.[117] For example, a measure designed to assess children's language development would be considered to have good content validity if it captured both receptive and expressive language abilities. Unlike other types of validity (as well as aspects of reliability summarised), face and content validity are not typically tested statistically. Rather, they are generally explored through consultation with content experts as well as alignment with the theoretical conceptualisation of the construct being measured.

Construct validity is evident when scores on a measure are consistent with theoretical expectations of the construct it was designed to measure.[117] These expectations can be established in several ways, including through exploring the internal factor structure (i.e., developmental domains) of the measure in observed data and how this aligns with theoretical conceptualisations of the tool, relationships between responses to different items/scales on the tool, and relationships between scores on the tool and measures of other constructs. Construct validity is often described as an overarching concept, with other forms of validity (including convergent, divergent, and discriminant validity) providing evidence for construct validity.[115]

Convergent validity refers to how closely related a measure (or scale) is to other measures/scales of similar constructs.[118] It posits that a measure accurately reflects a construct (e.g. social and emotional skills) if it correlates highly with other measures of similar constructs. In contrast, *divergent validity* is evident when a measure/scale is not highly correlated to other measures/scales of different or conceptually unrelated constructs.[119] Both convergent and divergent validity can be established by examining correlations among scales within the measure of focus, or that between the measure of focus and an additional, separate measure. For example, a holistic measure of children's development with convergent validity would show stronger correlations between aspects of development that are more conceptually related (e.g., literacy and numeracy scales), and demonstrate divergent validity through weaker correlations between aspects of development that are less conceptually related according to theory (e.g., social emotional skills and motor development).

Divergent validity, as per the definition provided above, is often also referred to as discriminant validity. However, in this thesis the term *discriminant validity* is used to refer to a measure's ability to produce scores that differ according to participant characteristics, such as age and sex, as theoretically expected. For instance, children's age has a strong association with development, therefore a measure of child development would be expected to discriminate between children of different ages, with younger children scoring lower on average, and older children scoring higher. This provides additional evidence of construct validity in that it examines if an instrument is capturing what it was designed to measure.

Finally, *criterion validity* refers to the extent to which an individual's score on a measure is related with other measures or variables (i.e., criteria) as theoretically expected.[120] Typically, criterion measures are expected to have established reliability and validity and therefore intend to serve as a "gold standard" measurement of the construct. When the criterion is measured at the same time as the instrument of focus, criterion validity is referred to as *concurrent validity*. For example, a measure of children's motor development with concurrent validity would have scores that are closely related to results from a paediatric screen of the child's developmental milestones such as crawling and picking up objects. When the criterion is measured at a time point after the construct has been measured, criterion validity is referred to as *predictive validity*. For instance, a measure of children's early literacy skills with predictive validity would have scores that predict later outcomes on reading ability in primary school.

1.3 Research aims

The overarching aim of this thesis was to investigate the psychometric properties of the eHCI within various LMICs. To do so, pre-existing data collected using the eHCI among children aged 2 to 6 years in seven countries throughout 2013 to 2020, including Brazil, China, Kiribati, Lao PDR, Samoa, Tonga, and Tuvalu, was utilised. Various steps in the validation process were conducted through exploring three research questions:

1. Does the eHCI measure the same underlying constructs (i.e., developmental domains) in seven LMICs, given the necessary local adaptation of the instrument in each country?
2. Does the eHCI have adequate validity, and does this vary across seven LMICs?
3. Does the eHCI measured at 2 to 5 years have predictive validity for children’s cognitive development (i.e., literacy, numeracy, and executive function) at 6 to 9 years?

Additionally, this thesis sought to explore whether the eHCI was able to achieve the measurement ideal of having the sensitivity required to adequately detect variations in children’s development; a limitation of many existing population measures of early childhood development. A fourth research question was focused on this aim, as follows:

4. Are scores on the eHCI sensitive to the quality of early education children attend?

Investigation of these research questions seek to advance understanding of how the eHCI can be used to facilitate population measurement of early childhood development in future.

1.4 Chapter conclusion

This chapter summarised the background against which the eHCI was developed, highlighting the need for a more pragmatic solution to population level measurement of early childhood development in LMICs. Against this background, the overarching aims and specific research questions of this thesis were outlined. The following chapter summarises the methodology that was applied to explore the research questions this thesis set out to investigate.

Chapter 2: Methodology

2.1 Chapter outline

This chapter describes the methodological approaches employed throughout this thesis. Given it is a thesis by publication, there is some overlap in the details presented in this chapter with those in the methods sections of papers in Chapters 3 to 6. This chapter first provides an overview of data sources across countries and measures utilised. It then focuses on methodological approaches employed, describing rationale for approaches selected in research studies and how they were undertaken. This chapter also includes discussion of how methods conducted address research questions and contribute to the overarching aims of this thesis.

2.2 Data sources

Pre-existing data, collected in seven LMICs throughout 2013-2020 were utilised. Research settings included seven countries, predominantly across the East Asia and Pacific region. Data were collected using different sampling techniques and collection methods, as summarised in Table 2.1. Datasets used were primarily cross-sectional, with one longitudinal dataset from Lao PDR. Children in the majority of country samples ranged in age from 2 to 6 years[a], with the longitudinal dataset including children up to 9-years-old at follow up. All datasets included the eHCI as an indicator of early childhood development, together with other child, caregiver, household, and village level variables which differed across countries (see Section 2.3).

Data were originally collected as part of four independent projects funded for program evaluation and population monitoring endeavours by local, national, and international agencies

[a] While the eHCI was designed for use among children aged 3 to 5 years, existing projects that utilised the instrument often included a broader age range of children to achieve project objectives. For this reason, eHCI data among children aged 2 to 6 years was available and utilised throughout this thesis.

to support children's health, development, and education. The following sections provide brief contextual information regarding each country and original projects for which data were collected, including their aims and implementation of the eHCI, to provide this background information prior to describing the specific measures and methods used throughout this thesis.

2.2.1 Brazil

Brazil is the world's fifth most populous country, occupying a large area on the eastern coast of South America. Classified as an upper middle-income country, in 2020, Brazil had a population of approximately 213 million and Gross National Income (GNI) per capita was USD7850.[121] Early childhood education is a constitutional right for children in Brazil, with municipalities responsible for service provision. Children aged 0-3 years attend day care centres and children aged 4 to 6 attend preschool. While not mandatory, both are recognised as educational institutions.[122] Government spending on education has increased in recent years (approximately 6% Gross Domestic Product (GDP) in 2018). As a result, almost all children accessed pre-primary education in 2020, with a gross enrolment ratio of 95%.[121]

Led by researchers at the University of Sao Paulo, a project was undertaken with the aim of enabling school-based population monitoring of early childhood development.[123] It was intended that such monitoring would provide information regarding children who were not developmentally on track across populations (i.e., classes, schools), as well as evaluate impacts of early years health and education programs. In 2015, the eHCI underwent local translation and adaptation and was piloted among children attending public schools in a city in Southern Brazil (n = 1,810). Data were collected from kindergarten (i.e., preschool) teachers to pilot the tool and test reliability and validity in the Brazilian context, but also to explore feasibility of the eHCI for large-scale implementation within a school setting. Resulting data were used to develop a short form of the instrument and undertake further psychometric testing.[107]

2.2.2 China

Situated in East Asia, China is the world's most populous country, accounting for approximately an eighth of the world's population under the age of five. Classified as an upper middle-income country, in 2020, China's population was approximately 1.4 billion and the GNI was USD10550 per capita.[121] There are three key forms of early childhood education in China, including nursery, kindergarten, and pre-primary classes. In 2018, government expenditure on education represented about 4% of GDP, and the pre-primary gross enrolment ratio was 90% in 2020.[121] Despite this, there are great inequalities in access to early education among children living in rural versus urban areas, with access as high as 98% in some urban areas and as low as 30% in some poor, rural areas.[124]

In 2015, the eHCI was translated and adapted for use in rural China. Led by the China Development Research Foundation, data collection was undertaken in two of China's Northern provinces[105], which were selected on the basis of pre-existing preschool programs in these areas. Data collection sought to explore inequalities in children's development across different population groups and establish a baseline of children's development in these provinces, prior to implementation of further early years interventions including the OneSky program. Respondents included a combination of children's caregivers and teachers. These initial data collection efforts (n = 11,421), which constitute the data that have been utilised in this thesis, demonstrated feasibility of use of the tool at scale in China. The eHCI has since been employed to collect information on more than 200,000 children across both rural and urban settings in China for various population monitoring and program evaluation endeavours.[125]

2.2.3 Lao PDR

A landlocked country in Southeast Asia, in 2020 the population of Lao PDR was approximately 7.3 million and the GNI was USD2520 per capita.[121] As one of the fastest growing economies in East Asia and the Pacific, Lao PDR recently moved to lower middle-income classification. However, there exist significant health, education, and economic disparities between different ethnicities and geographies.[126] Prominent early childhood education programs provided by the government include kindergarten and pre-primary targeting children aged 3-5, prior to compulsory primary education. In 2014, government expenditure on education was approximately 3% and in 2020, the pre-primary gross enrolment ratio was about 49%.[121]

The World Bank Early Childhood Education Project (herein ECE Project) aimed to increase access to quality early childhood education (ECE) services among children aged 3 to 5 years in Lao People's Democratic Republic (PDR).[127] The project was implemented by the Government of Lao PDR in almost 400 villages across five northern provinces. Project interventions included three components designed to increase demand for, coverage, and quality of ECE, as well as strengthening monitoring and evaluation in the early childhood sector. Three clustered randomised control trials were employed to evaluate the impact of interventions on children's development over a four-year period.[102] Data were collected at three time points throughout 2015 to 2020, two of which were included in the studies presented in this thesis; baseline in 2015/6 (n = 7,493) and endline in 2020 (n = 13,896). Endline data collection included follow up of children from baseline, as well as data collection among a new cohort of children aged 2 to 6 years.[128] Data collection utilised various measurement tools and respondents including village head, household head, and caregiver questionnaires (including the eHCI), a child direct assessment, and an ECE classroom observation measure.

2.2.4 Pacific Island Countries

Kiribati

Comprised of 33 coral atolls in the Central Pacific, the population in Kiribati was approximately 119,000 in 2020[121]. Classified as a lower middle-income country, the GNI in

2020 was USD2960 per capita.[121] Kiribati is one of the world's nations most vulnerable to rising sea levels as a result of climate change. Early childhood education in Kiribati was recently formalised in 2017, with preschool services provided by the non-government sector including churches and community groups, predominantly for children aged five years. In 2020, the pre-primary gross enrolment ratio was 89%.[121] Government expenditure on education represented 12% of the country's GDP in 2019[121], and this is reflected in that the first nine years of education in Kiribati are free and compulsory.[129]

Samoa

Samoa is comprised of two main islands in the South Pacific. In 2020, Samoa's population was approximately 198,000 and the GNI was USD4050 per capita.[121] Classified as a lower middle-income country, Samoa has experienced less frequent natural disasters relative to other Pacific Island Countries, though has been devastated by significant economic and social shocks resulting from previous weather events. In 2020, government expenditure on education represented 5% of Samoa's GDP.[121] Early childhood education is provided primarily for children aged 3 to 4 years (with 2-year-old children also eligible to attend) by non-government organisations including churches and other organisations.[130] In 2019, the pre-primary gross enrolment ratio in Samoa was 41%.[121]

Tonga

Tonga is an archipelago of more than 170 South Pacific islands, 36 of which are inhabited. In 2020, the population in Tonga was 106,000, with 70% residing on the main island (Tongatapu). Tonga is an upper middle-income country, and in 2020, had a GNI of USD5190 per capita. Like many Pacific Island Countries, emigration out of Tonga is high with remittances accounting for almost 40% of GDP.[131] In 2020, the Education Act was revised to include early childhood education (primarily provided by non-government organisations including churches and community groups) for children aged 4-5 in the compulsory education system.[132]

Government expenditure on education in 2019 was 8% of GDP, and in 2020, the pre-primary gross enrolment ratio in Tonga was 48%.

Tuvalu

Situated in the South Pacific, Tuvalu comprises 9 small islands totalling just 26 square kilometres. In 2020 the population of Tuvalu was approximately 12,000[121], with a third of the population under 15 years of age. While Tuvalu is classified as an upper middle-income country, it is considered one of the most economically vulnerable countries globally[133] due to limited resources as well as risk of adverse impacts of climate change and extreme weather events. Primary and secondary education is free and compulsory, with preschool provided by church, community, and private organisations for children under 6 years of age. The education sector in Tuvalu receives the largest proportion of the national budget (18% in 2019)[b]. In 2020, the pre-primary gross enrolment ratio was 79%.[121]

Initiated by the World Bank in 2014, the Pacific Early Age Readiness and Learning (PEARL) Program supported capacity building to design, implement, and monitor evidence-based policies and programs that prepare children and families for school.[99, 100] Measurement of child development, among other aspects, and the design of interventions to improve these outcomes were supported in Kiribati, Samoa, Tonga, and Tuvalu. Across countries, a census approach to measuring children's development was employed, using the eHCI to collect information for every child aged 3 to 5 years nationally. This was enabled through a combination of teacher and caregiver reported information. In Tonga, data collection was designed to establish a baseline of children's development and evaluate the impact of project interventions. A census of children's development was conducted in 2013/14 across Tonga's 54 inhabited islands (n =

[b] Recent government expenditure (since 1997) on education as a proportion of country's GDP was not available in Tuvalu. Proportion of the national budget allocated to the education sector has been described in place of this.

6,214), from which data are utilised in this thesis, with a repeat census in 2017 indicating positive impacts on school readiness. In Kiribati (n = 8,339), Samoa (n =12,191), and Tuvalu (n = 549), data collection sought to understand the current status of children's health and development. For instance, in 2015, Tuvalu implemented the eHCI census across all nine islands. Results informed health and development supports, with repeat data collection in future intended to enable monitoring and evaluation.[104] Samples from these additional three collections were also utilised throughout this thesis.

2.2.5 Data access and ethics approval

Access to data sources was existing through the roles of researchers Sincovich and Brinkman in the adaptation of the eHCI in the seven countries described. Approval to use data was granted by data custodians within each country, including the University of Sao Paulo (Brazil); the China Development Research Foundation (China); and the World Bank (Lao PDR and Pacific Island Countries including Kiribati, Tonga, Tuvalu, and Tonga). Approval was granted to use data for the purposes of exploring the reliability and validity of the eHCI within and across countries, with results to be published within this thesis as well as academic journals.

Research activities undertaken for this thesis included secondary analyses of pre-existing, de-identified data and thus were deemed exempt from requiring ethical review by the University of Adelaide Human Research Ethics Committee (see Appendix A).

Table 2.1. Overview of data sources utilised

Country	Year	Research design	Sampling technique	Respondents	Age range	Sample
Brazil	2015	Cross-sectional	Sample; children from all 37 public schools one city in Southwest Brazil	Preschool teachers	2-5 years	1,810
China	2015/16	Cross-sectional	Sample; children from two provinces across Northern China	Preschool teachers & caregivers	2-6 years	11,421
Lao PDR	2015/16 (BL); 2020 (EL)	Longitudinal	Sample; children from 376 villages in 5 provinces across Northern Lao PDR	Caregivers	2-6 years (BL); 2-9 years (EL)	7,493 (BL); 13,896 (EL)
Kiribati	2017	Cross-sectional	Census; aimed to collect data for all children aged 3-5 years nationally	Preschool teachers & caregivers	2-6 years	8,339
Samoa	2016	Cross-sectional	Census; aimed to collect data for all children aged 3-5 years nationally	Preschool teachers & caregivers	2-5 years	12,191
Tonga	2013/14	Cross-sectional	Census; aimed to collect data for all children aged 3-5 years nationally	Preschool teachers & caregivers	2-6 years	6,214
Tuvalu	2015	Cross-sectional	Census; aimed to collect data for all children aged 3-5 years nationally	Preschool teachers & caregivers	2-6 years	549

Note. Respondents refers to those who completed the eHCI in each country. BL = baseline, EL = endline. While in some countries the objective was to collect data among children within a specified age range (e.g., 3-5 years), the resulting sample often had a wider age range (2-6 years) due to small numbers of children falling outside of the intended age range

2.3 Measures

In this section, the measures and variables collected previously as part of original projects summarised and used in the four studies that form this thesis, are described. Table 2.2 provides an overview of the information for this thesis used within each country. While measures used across countries and papers differed (see Section 2.4), the eHCI was collected in all countries and used across all four research studies in this thesis. Additional child, caregiver, household, and village level information, collected alongside the eHCI in each country, were also used.

2.3.1 Early childhood development – early Human Capability Index (eHCI)

The eHCI includes approximately 60 items (dependent upon country adaptation, ranging from 56 in Lao PDR to 66 in Tuvalu) measuring children's development across nine domains: Physical Health, Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, Approaches to Learning, Numeracy, Reading, and Writing. The eHCI underwent a local adaptation process in each country to ensure the tool's content and face validity, as described in Chapter 1. Therefore, although there are many similarities across adaptations of the tool, some items and domains differ across countries. For example, the Perseverance domain is measured by the same four items across all adaptations. The Physical Health domain, however, varies from 2 items in Brazil to 5 items in Kiribati and Tuvalu, while the Laotian version of the eHCI does not capture Physical Health. Each countries' version of the adapted instrument is presented in Paper 1 supplementary materials (Chapter 3).

Table 2.2. Overview of variables across countries utilised

Measurement level	Variable	Brazil	China	Lao PDR		Kiribati	Samoa	Tonga	Tuvalu
				BL	EL				
Child level	Early childhood development (eHCI)	✓	✓	✓	✓	✓	✓	✓	✓
	Sex	✓	✓	✓	✓	✓	✓	✓	✓
	Age	✓	✓	✓	✓	✓	✓	✓	✓
	Ethnicity			✓	✓				
	Preschool attendance		✓	✓	✓	✓	✓	✓	✓
	Stunting		✓	✓	✓	✓	✓	✓	✓
	Health status			✓	✓				
	Home learning environment			✓		✓	✓	✓	✓
	Cognitive development (direct assessment)			✓	✓				
Caregiver level	Education		✓	✓		✓	✓	✓	✓
	Literacy			✓					
Household level	Socioeconomic status				✓				
Village level	Geographical remoteness				✓				
	Early childhood education quality				✓				

Note. BL = baseline, EL = endline.

The eHCI is completed by an adult who knows the child. Respondents included children's teacher or primary caregiver in the studies presented throughout this thesis. Response options for each item are binary (yes/no, able/unable, can do already/cannot do yet). All items are completed for all children (i.e., no skip functions are employed, items do not differ according to children's age). The majority of items are positively worded so that the yes/able/can do already responses are scored as 1, and the no/unable/cannot do yet responses are scored as 0. A small number of items (ranging from 4 in Kiribati to 6 in Tonga) are negatively worded and were reverse scored. Derivation of domain scores and overall development scores were completed in accordance with guidelines provided by instrument creators.[101] Individual item scores were averaged so that children receive a score for each domain ranging from 0 to 1, with higher scores indicating better development. Where ≥ 20% of items in a domain had a missing response, children were allocated a missing value for that domain. An overall development score was derived by taking the average of domain scores, also ranging from 0 to 1, with higher scores indicating better development. Children with missing scores on ≥ 20% domains were allocated a missing value for the overall development score.

2.3.2 Demographic and contextual characteristics

Various demographic and contextual information were collected alongside the eHCI in each country, at the child, caregiver, household, and village level, with some variation in information available across countries. This information is summarised below.

<u>Child level</u>

Child level information included children's age in years (calculated using child's date of birth and date of data collection), sex (male, female), and preschool attendance (yes, no). Additionally, child ethnicity (Lao Tai, Khmu, Hmong, Other), as well as caregiver-reported child health status (very healthy, normal, unhealthy/often sick, or don't know based on the item "What do you think is the current health status of the child?") were used in Lao PDR.

Caregiver level

Caregiver level information included children's mother's highest level of education (collected in all countries except Brazil), with response categories varied across countries. In Lao PDR, caregiver literacy was also utilised, measured through asking the caregiver to read a short sentence (cannot read at all, can read a little, can read well, vision problem).

Household level

In Lao PDR, several household level variables indicative of socioeconomic status (e.g., assets, number of household members, house construction materials etc), collected through household head interview, were used to derive an indicator of socioeconomic status, with five quintiles from least to most advantaged (see Section 2.4.4 for further details).

Village level

Also in Lao PDR, information collected via a village head questionnaire included an indicator of remoteness based on the item "Can cars access the village during rainy seasons?" (yes, no).

2.3.3 Stunting

Children's height and weight, measured in centimetres and kilograms by trained administrators, were recorded at the time of data collection in each country (except Brazil). Anthropometric measures were converted into WHO Child Growth Standards height-for-age z-scores, and stunting was defined as a height-for-age z-score < -2.[134] For example, 95.0cm is the average height for a girl aged 36 months (standard deviation = 3.8). A z-score of -2 corresponds to a height of 87.4cm (i.e., $z = (87.4 - 95.0)/3.8 = -2.0$) Therefore, a 3-year-old girl equal to or less than 87.4cm in height would be classified at stunted.

2.3.4 Home learning environment

Information about caregiver-child engagement in six types of learning activities in the home was also collected alongside the eHCI in each country (except in China and Brazil). These binary questions (yes, no) were based on items from the MICS questionnaire for children under five[135] and asked if, in the last 3 days (or in the last 7 days in Lao PDR and Tonga), a member of the household aged 15 years or above had: read books or looked at picture books with the child; told stories to the child; sang songs or danced with the child; played with the child; took the child outside of the home; and named, counted or drew with the child. In Kiribati, Samoa, and Tonga, respondents were also asked if there were children's reading materials (e.g., picture books) in the home (also based on the MICS with a yes/no response).

2.3.5 Cognitive development

Concurrent to administration of the eHCI in Lao PDR, children's literacy, numeracy, and executive function were measured via 92 direct assessment items. Assessment was administered by a trained administrator who asked children to complete a series of activities in a one-on-one setting. Items formed an early working version (2014/5) of the Measurement of Development Early Learning (MODEL; part of the MELQO initiative)[83], prior to validation studies.[84, 87] Literacy and numeracy assessments were based on items from the Early Grade Reading Assessment and the Early Grade Maths Assessment and adapted to the context in Lao PDR.[102] Both assessments have been adapted for use in a number of countries, with established reliability and validity.[136] Executive function assessments were based on items from two sub-tests from the Wechsler Intelligence Scale for Children (WISC-IV; digit span forward and digit span backward) and the Head-Toes-Knees-Shoulders task.[137, 138] Overall, children were assessed on 6 sub-domains of literacy (37 items e.g., print familiarity, letter knowledge), 9 aspects of numeracy (39 items e.g., number identification, spatial vocabulary), and 3 aspects of executive function (16 items e.g., backward digit span) (a list of sub-domains is included in

Paper 3; Chapter 5). Correct item responses were scored as 1 and incorrect responses scored 0. Skip functions were employed in sub-domains that included a large number of items. For instance, the number identification sub-domain involved children reading a sheet of 20 numbers, with administrators instructed to skip remaining items if children respond incorrectly to five numbers consecutively. For sub-domains measured via multiple items, individual item scores were averaged so that children received a score for each sub-domain, ranging from 0 to 1. Sub-domain scores were then averaged to provide a total score for literacy, numeracy, and executive function, ranging from 0 to 1, with 1 indicating better development.

2.3.6 Early childhood education quality

The Measure of Early Learning Environments (MELE) captures constructs related to the quality of children's early learning environments.[83] The MELE does not intend to evaluate teachers or schools. Rather, it was designed to enable tracking of the quality of children's learning environments at a population level and identify areas in which training and supports might be needed to promote early learning outcomes. Several adaptations of the MELE have been developed and field-tested, with validation established in selected LMICs.[84, 86, 139]

The MELE classroom observation tool was adapted and implemented to capture early childhood education quality as part of the ECE Project in Lao PDR[140]. The tool measures four key domains of quality: Learning Activities (8 items); Classroom Interactions and Approaches to Learning (10 items); Classroom Arrangement, Space, and Materials (14 items); and Facilities and Safety (6 items). A description of domains and items is included in Paper 4, Chapter 6. With the exception of the Classroom Arrangement domain, items were scored from 1 to 4, with 4 representing high quality. Items in the Classroom Arrangement domain were either binary (yes/no) or multiple choice to indicate presence of materials (no materials present/materials present but children do not use/materials present and children use). Domain scores were calculated by summing scores on all items. Domain scores were transformed to range from 0 to 10, so that all domains were on the same scale to aid in interpretability, with 10 indicating high quality. An overall quality score was derived by taking the average of domain scores, also ranging from 0 to 10, with higher scores indicating better quality.

2.4 Research methodology

The overarching objective of this thesis was to investigate the psychometric properties of the eHCI using data collected across several LMICs. Additionally, it aimed to explore whether the tool was able to achieve the measurement ideal of having adequate sensitivity to detect variations in children's development according to the quality of ECE children attend. Together, this research sought to advance understanding of how the eHCI can be used to facilitate population measurement of early childhood development in future.

Four research questions were investigated to achieve this aim:

1. Does the eHCI measure the same underlying constructs (i.e., developmental domains) in seven LMICs, given the necessary local adaptation of the instrument in each country?
2. Does the eHCI have adequate validity, and does this vary across seven LMICs?
3. Does the eHCI measured at 2 to 5 years have predictive validity for children's cognitive development (i.e., literacy, numeracy, and executive function) at 6 to 9 years?
4. Are scores on the eHCI sensitive to the quality of early education children attend?

Each research question was explored separately in four studies, which align with the four chapters presented in this thesis to follow. Table 2.3 provides an overview of the aspects of reliability, validity and sensitivity that were explored by each research study, as described in Chapter 1. The first two studies focused on investigating various aspects of the reliability and validity of the eHCI using data sources in all seven LMICs. The third and fourth studies focused on exploring additional properties of the tool in one country, Lao PDR. This included whether scores on the eHCI predicted children's later outcomes and were sensitive to variation in children's development based on the quality of early education children attended, both of which are desirable characteristics of a population measure of early childhood development.

Findings from the four studies are integrated in Chapter 7 to address the overarching research aims. The research questions that were able to be addressed in this thesis were conducted within the context of the available data at the time the project was conceived. Thus, Chapter 7 also includes discussion of additional aspects of reliability and validity that were not able to be explored here but are important steps in the process of instrument validation.

All four research questions were investigated using quantitative methods. These methods, as well as descriptions of how data sources were used and analyses were conducted to address research questions, are described separately for each study in sections to follow. This chapter then concludes with information regarding how missing data were accounted for in analyses.

Table 2.3. Overview of types of reliability, validity, and measurement ideals addressed by research studies

		Description	Study 1	Study 2	Study 3	Study 4
Reliability	Internal reliability	Evident when a set of items that form a scale or measure are highly correlated.	✓			
Construct validity	Internal factor structure	Evident when scores on domains demonstrate alignment to the measure's theoretical structure.	✓			
	Convergent validity	Evident when measures of theoretically related constructs are highly correlated.		✓		
	Divergent validity	Evident when measures of theoretically less related constructs are not highly correlated.		✓		
	Discriminant validity	Evident when measures produce scores that differ according to participant characteristics, such as age and sex, as theoretically expected.		✓		
Criterion validity	Concurrent validity	Evident when measures of constructs are highly correlated with established measures of the same or similar constructs, measured at the same time.		✓		
	Predictive validity	Evident when measures of constructs are highly correlated with established measures of the same or similar constructs, measured at a later time.			✓	
Additional ideals	Sensitivity	Evident when scores on a measure reflect differences based on inputs as theoretically expected (e.g., intervention/treatment quality or dose).				✓

2.4.1 Research question 1

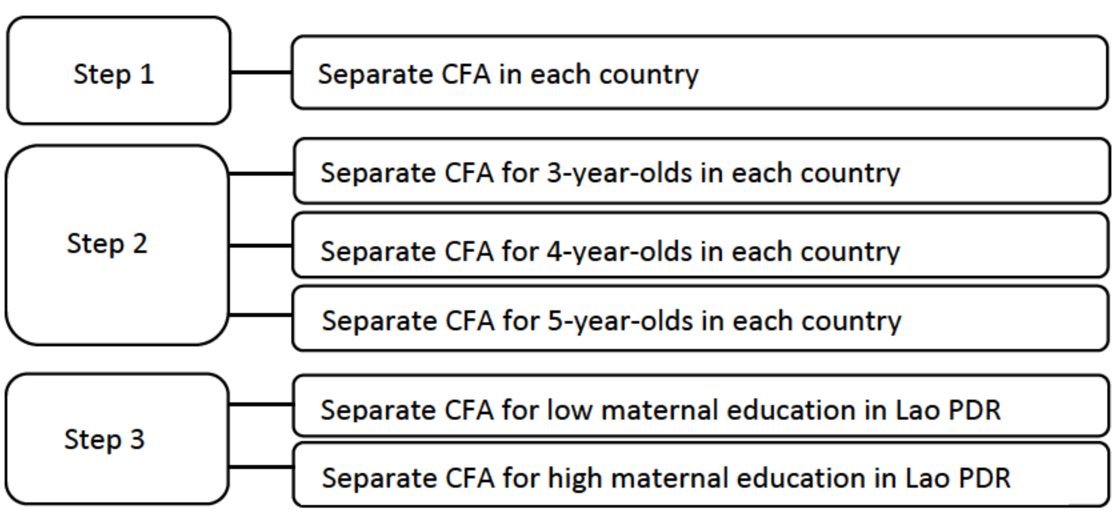
Does the eHCI measure the same underlying constructs (i.e., developmental domains) in seven LMICs, given the necessary local adaptation of the instrument in each country?

Development of the eHCI, and each country adaptation of the instrument, was informed by a combination of both theoretical conceptualisation as well as local expert consultation regarding the key aspects of children's development that are predictive of future capabilities.[99] This process was intended to produce an instrument that measures children's development relative to what is considered a well-developed child in a particular culture or context, producing locally relevant and meaningful information. While items differ across adaptations of the eHCI, the same nine developmental domains were maintained across countries. One exception was in Lao PDR, whereby the adaptation of the eHCI did not include the Physical Health domain.

Although the eHCI was designed to measure a specific set of developmental domains, this theoretical factor structure had not been explored using factor analysis methods. Confirmatory factor analysis (CFA) is an analytical approach used to verify the structure of items/variables that form a measure, testing the hypothesised relationship between items and underlying constructs. This study used CFAs to explore whether eHCI data across seven countries fit the theoretical structure of the instrument (i.e., 9 developmental domains, or 8 in Lao PDR), given the necessary adaptation of the instrument within each country. That is, whether the tool captured nine distinct underlying constructs (or domains) across countries, and the degree to which items designed to measure each of these constructs loaded onto the relevant factor (e.g., whether all items in the numeracy domain captured the same construct). Overall, this sought to provide evidence for the tool's construct validity across countries, indicating it is measuring what it was designed to measure. This enabled confirmation of the theoretical structure of the tool before subsequent psychometric analyses were conducted based on this structure. Internal reliability of eHCI domains was also examined in each country, using Cronbach's alpha as well as Ordinal alpha.

This study utilised cross-sectional data from seven countries (baseline data in Lao PDR; Table 2.1). A series of CFAs were conducted in three steps, summarised in Figure 2.1. First, 9-factor CFAs (or an 8-factor CFA in Lao PDR) were conducted separately for each country, addressing the primary research question of this study by exploring the fit of data in each country to the theoretical structure of the eHCI (i.e., 9 developmental domains, or 8 in Lao PDR). Two additional steps explored any variation in the fit of eHCI data according to children's age or respondent's education. Specifically, 9-factor CFAs (or an 8-factor CFA in Lao PDR) were conducted separately for children aged 3, 4, and 5 years in each country (as this was the age range that was consistent in datasets across countries). Lastly, using Lao PDR data, the sample was split into two groups (low and high maternal education) and 8-factor CFAs were conducted separately for each group. This analysis was conducted among the Lao PDR sample only, as it was the only sample for which data were available on the education of all respondents (caregivers).

Figure 2.1. Confirmatory factor analyses conducted



CFAs were conducted using MPlus 7.31[141], employing polychoric correlation matrices and weighted least square mean and variance adjusted (WLSMV) estimation, both of which are most appropriate for use with binary data[142-144], as well as oblique (goemin) factor rotation which assumes correlations among factors.[144] Various goodness-of-fit indices, together with standardised factor loadings were used to evaluate model fit to the eHCI data in each country. Specifically, fit indices included $\chi 2$ ($p > 0.05$ indicates good model fit)[144], root mean square error of approximation (RMSEA; values ≤ 0.06 indicate good model fit, values between 0.06-0.08 indicate satisfactory fit)[145, 146], and the Comparative Fit Index and the Tucker-Lewis Fit Index (CFI and TLI, respectively; values ≥ 0.95 indicate good fit, values between 0.90-0.95 indicate satisfactory fit)[145-147]. Factor loadings ≥ 0.40 deemed items a good indicator of the underlying latent construct.[148]

The internal reliability of eHCI domains was also examined in each country. This is often conducted in conjunction with factor analysis to assess how interrelated a set of items are and thus how well they, collectively, measure the underlying construct.[149] For instance, the relatedness of items in the eHCI numeracy domain will indicate whether items are capturing similar skills. Many psychometric research studies assess the internal reliability of scales/domains using Cronbach's alpha. Increasingly, ordinal reliability coefficients, specifically ordinal alpha, have been deemed more appropriate to evaluate the internal reliability of scales comprised of items with binary or ordinal response options, such as the eHCI.[150] Thus, ordinal alpha coefficients were calculated for each eHCI domain separately for each country, along with Cronbach's alpha to enable comparison with previous research. Coefficients ≥ 0.70 are deemed acceptable for both internal reliability indicators[113, 150] and were calculated using package 'psych' in R Statistical Software.[151]

2.4.2 Research question 2

Does the eHCI have adequate validity, and does this vary across seven LMICs?

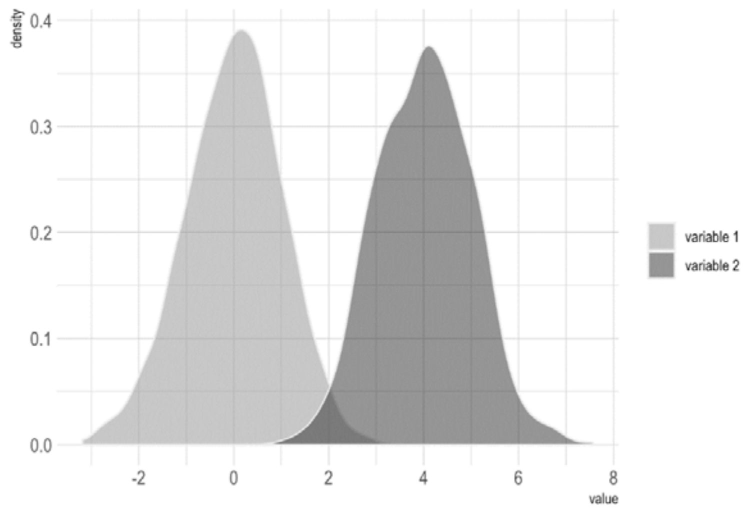
With internal factor structure and reliability of the eHCI investigated, this research study focused on exploring the validity of the tool. As described in Table 2.3, this study evaluated convergent and divergent validity, discriminant validity, and concurrent validity. This sought to provide further evidence that the eHCI was measuring what it was designed to measure across diverse settings (i.e., construct and criterion validity). Evaluation of these aspects of validity required investigation of how eHCI domain scores were associated with one another, as well as other measures. Using cross-sectional data from seven countries (baseline data in Lao PDR; Table 2.1), descriptive statistics, correlations, and linear regressions were calculated together with density plots in each country, to determine psychometric robustness of the tool. Results in each country were examined to identify any variations in validity across countries. Analyses were conducted using IBM SPSS Statistics 25[152], with density plots generated using R Statistical Software 3.6.0.[153]

Convergent and divergent validity were tested by examining correlations amongst eHCI domains within each country. Two-tailed Spearman's rho correlations, appropriate for use with ordinal scales, were used to indicate associations between domains, where $r = 1$ indicates a perfect positive correlation and $r = -1$ indicates a perfect negative correlation.[154] It was predicted that the strongest associations would be observed between eHCI domains measuring children's literacy (i.e., reading and writing) and numeracy across countries, with smaller correlations between remaining domains (e.g., physical health and social and emotional development). This hypothesis was based on the intertwined nature of children's literacy and numeracy development, relative to other domains.

Discriminant validity of eHCI domains was tested using various child and family characteristics in each country, including child age, sex, stunting status, preschool attendance, maternal education, and home learning environment. Whether eHCI scores could discriminate amongst children's development by these characteristics was examined by first calculating the mean and standard deviation of eHCI domain scores separately across countries, stratified by response values for each of the child and family variables. This enabled comparison of scores for males versus females, children who attended and did not attend preschool and so on. It was predicted that higher domain scores would be observed among females, older children, those not stunted, children who attended preschool, children of more educated mothers, and children with learning opportunities at home. It was also hypothesised that, if measuring children's development as intended, differences in scores by child and family characteristics would vary across eHCI domains. For instance, it was predicted there would be a larger difference in mean literacy and numeracy domain scores between children who were and were not stunted, compared to the social and emotional domain, because of the well-established adverse impacts stunting has on children's cognitive development.[155]

Kernel density plots were generated to provide a visual representation of the distribution of eHCI scores for different demographic groups within each country. Density plots, a smoothed variation of the histogram, represent the distribution of one or multiple numerical variables. An example is provided in Figure 2.2 below. In this instance, density plots help to highlight variations in the discriminant ability of eHCI domains for different demographic variables across countries, which may not be as apparent from examining domain mean and standard deviation scores. For example, plots may help to demonstrate the difference in age gradient in children's scores on Literacy versus Social and Emotional Skills domains, and the extent to which this varies across countries.

Figure 2.2. Example of density plot utilised to investigate discriminant validity of eHCI domains



Unadjusted linear regressions were used to test associations between child and family characteristics (independent variables) and eHCI domains (dependent variables) in each country. Regressions, rather than correlations, enabled exploration of the increase or decrease in eHCI domain scores associated with various characteristics explored. Unstandardised beta coefficients were reported with 95% confidence intervals. Coefficient size together with confidence intervals that did not overlap zero were interpreted as meaningful associations.[156]

Concurrent validity was investigated using baseline data in Lao PDR, as in this study, children's development was measured via direct assessment in addition to the eHCI (see Section 2.3.5). This analysis sought to demonstrate criterion validity of the eHCI (though noting lack of a gold standard measure discussed in Chapter 1) and was tested through exploring the association between children's scores across both measures. Two-tailed Spearman's rho correlations among eHCI domains and direct assessment literacy and numeracy domains were calculated. It was predicted that the strongest associations would be observed between direct assessment scores and eHCI domains measuring literacy and numeracy, as these domains measure the same constructs, with smaller correlations between direct assessment scores and remaining eHCI domains.

2.4.3 Research question 3

Does the eHCI measured at 2 to 5 years have predictive validity for children's cognitive development (i.e., literacy, numeracy, executive functioning) at 6 to 9 years?

The first two studies explored whether the eHCI was psychometrically robust across diverse contexts using cross-sectional data. This research study was conducted with the aim of exploring the ability of the eHCI to predict children's later abilities, establishing the predictive validity of the tool. Longitudinal data in Lao PDR enabled this investigation, and Receiver Operator Characteristic curve analyses were conducted to explore this research question.

The ECE Project in Lao PDR, for which data were utilised for this study, included three waves of data collection; baseline (November 2015 to March 2016), midline (November to December 2017), and endline (February to March 2020). The eHCI and direct assessments of children's literacy, numeracy, and executive function (see Section 2.3.5) were employed at all waves. Table 2.4 depicts child age ranges at each data collection time point[c]. Children for whom data were collected at both baseline and endline (as opposed to baseline and midline, or midline and endline) were selected as the sample for this study. This enabled investigation of the ability of eHCI scores at ages 2 to 5 years predicting cognitive development outcomes at 6 to 9 years, four years later. A longer time frame between measurement of the eHCI and the criterion (direct assessments) enabled the most rigorous investigation of the predictive validity of the eHCI possible with available data.

[c] Age ranges indicate ages of the majority of children. Each data collection was undertaken over a period (from two to five months) and child age differs based on the child's birth date and data collection date across time points. For instance, some children who were aged 2 at baseline will have been 4 years, rather than 3 years, at midline.

Table 2.4. Child age across data collection points for ECE Project in Lao PDR

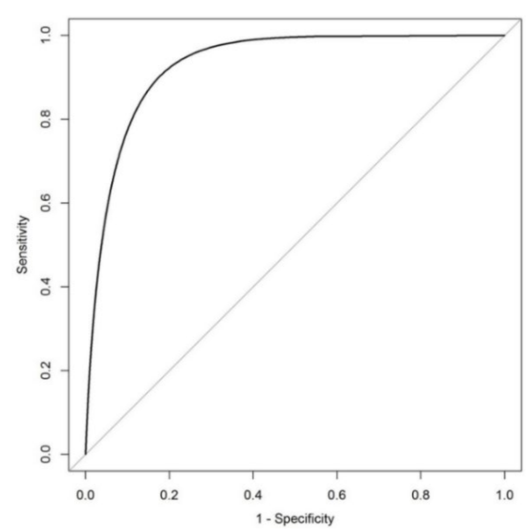
Collection	Year	Child age (years)			
Baseline	2015/6	2	3	4	5
Midline	2017	3	4	5	6
	2018	4	5	6	7
	2019	5	6	7	8
Endline	2020	6	7	8	9

Limited early childhood development measures have established predictive validity. Predictive validity has typically been undertaken using methods, such as correlations and regression analyses, that demonstrate an association between scores on the measure and a later outcome. For instance, predictive ability of scores on the EDI at age 5 for literacy and numeracy skills at ages 8-12 was established using correlations and regressions.[157] Similar methods were used to explore the ability of the IDELA among children in pre-primary school to predict literacy and numeracy skills 2.5 years later.[158] While an association between scores is important, it does not demonstrate the ability of a measure to discriminate between groups of children with different outcomes in future (e.g., developing well versus poorly). This can also be described as predicting a particular outcome.

Although it is not the intention of a population measure of child development, such as the eHCI, to provide an individual diagnosis for poor development, it is intended that data collected be used to inform investment in supports to promote children's outcomes. Thus, establishing that a poor score on the eHCI discriminates between children doing well versus poorly on later indicators of ability, is important to demonstrate results are meaningful and bolster confidence in use of the tool. In this study, methods designed to investigate the ability of eHCI domains to predict scores on direct assessment measures (i.e., literacy, numeracy, and executive function) were employed.

Receiver Operator Characteristic (ROC) curve analyses have commonly been used to evaluate diagnostic accuracy of a clinical assessment for a dichotomous outcome.[159, 160] In this study, ROC curves were used to explore the predictive validity of the eHCI and are presented alongside the area under the ROC curve (AUC), or C-statistic, to determine the ability of measures at ages 2 to 5 to predict (discriminate between) outcomes at ages 6 to 9. An AUC of 0.5 indicates the measure is no better than chance at discriminating between later outcomes. The ROC curve graphs sensitivity (percentage of true positive outcomes) as a function of 1-specificity (percentage of true negative outcomes) with the diagonal reference line equal to chance (see Figure 2.3 for an example). The AUC or C-statistic represents the probability that a randomly selected case who experienced an outcome will have a higher predicted probability of experiencing the outcome, compared to a randomly selected case who did not experience the outcome. Therefore, an AUC of 1.0 indicates the predictor perfectly discriminates between outcomes (i.e., 100% sensitivity and 100% specificity). Typically, for clinical diagnostic purposes, an AUC between 0.5 to 0.7 is considered low, 0.7 to 0.9 moderate, and greater than 0.9 indicates high predictive ability for the measure to discriminate between two outcome groups.[161] However, these guidelines should not be applied as strict "rules" and are dependent on context and purpose of assessment.

Figure 2.3. Example of Receiver Operator Characteristic curve utilised to investigate predictive validity of the eHCI



ROC curve analyses were conducted in three steps. First, ROC curves were calculated to explore the ability of eHCI overall development and direct assessment literacy, numeracy, and executive function domain scores at baseline to predict poor development on the same direct assessments at follow up. Poor development in outcome measures was defined as scores in the bottom 10th percentile (for each age in years stratum). Second, a sensitivity analysis was conducted by varying the definition of poor development to scores in the bottom 20th percentile. Third, additional ROC curve analyses explored the ability of the eight individual eHCI domain scores at baseline to predict poor development scores at follow up, as defined as scores in the bottom 10th percentile on the direct assessment literacy, numeracy, and executive function. To help interpret the size of effects in all ROC curve analyses, the relationship between two measures of socioeconomic status (caregiver education and literacy) at baseline and poor development at follow up was also explored, as these factors would also be expected to predict children's developmental outcomes. Analyses for this study were conducted using IBM SPSS Statistics 27.[162]

All analyses were stratified by child age in years at baseline. This step was critical because of the strong positive association between age and children's development. As analyses explored ability to predict scores in the bottom 10th percentile on direct assessment measures, use of a single cut-off score for poor development among children of all ages would have ignored this age gradient. Classification of poor development would have been unequally applied across ages, resulting in an over-representation of the youngest children in the bottom 10% of scores and erroneous conclusions regarding the ability of the eHCI to predict ability at later ages.

2.4.4 Research question 4

Are scores on the eHCI sensitive to the quality of early education children attend?

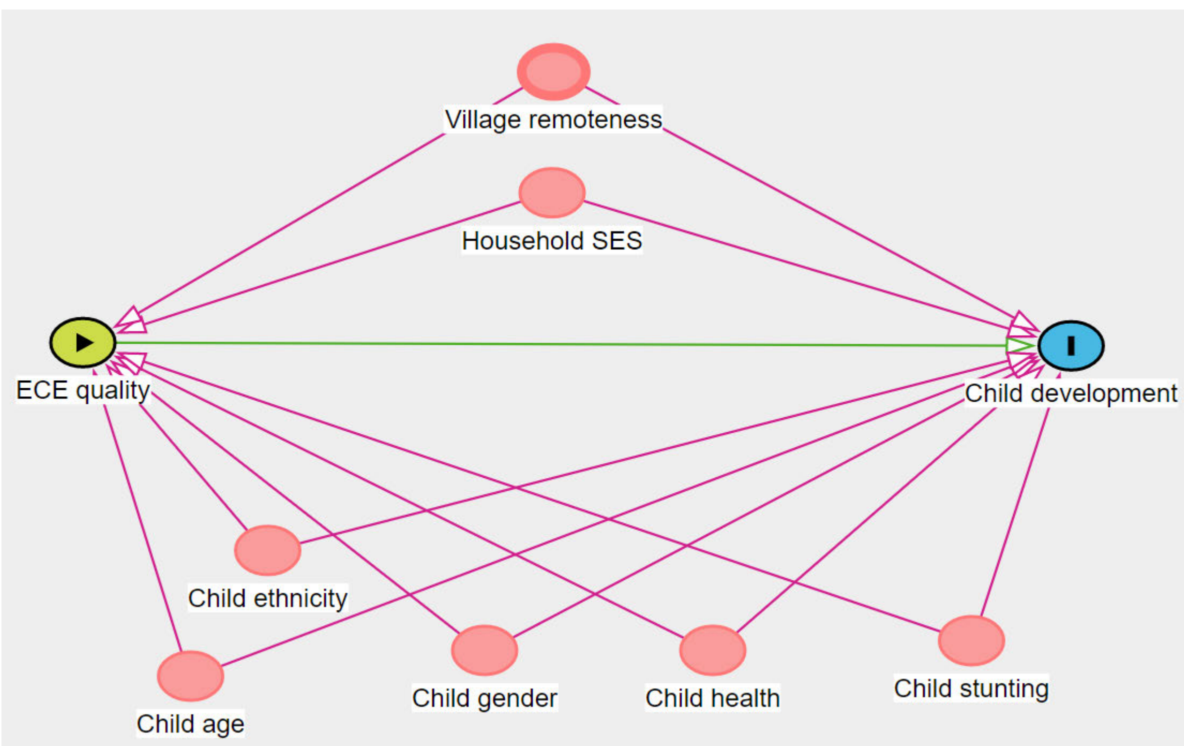
The first three research studies explored the psychometric properties of the eHCI across diverse contexts, including whether scores on the tool were able to predict children's later outcomes. The eHCI was designed to be sensitive to variation in children's development to facilitate program evaluation. Indeed, this is why the tool was originally developed in Tonga. This final study was designed to investigate the sensitivity of eHCI scores to inputs or supports designed to promote children's development, namely quality of ECE. This sought to further explore whether the eHCI has the sensitivity to detect these variations in children's development, which is a limitation of many existing population measures of child development, as discussed in Chapter 1.

Higher quality ECE typically results in improved child development outcomes.[163] Quality of ECE was measured in the final wave of data collection for the ECE Project in Lao PDR, alongside both the eHCI and direct assessment of children's development (Table 2.2). It was hypothesised that, on average, children attending better quality ECE would have higher development scores on the eHCI and direct assessment of children's development, after adjusting for child, household, and village level confounders. As described in Section 2.3.6, the measure of ECE quality utilised was the MELE. Being a relatively new measure of ECE quality, the MELE had not previously been validated in Lao PDR, and research had demonstrated mixed evidence regarding the measure's relationship with child development across selected LMICs. Therefore, this study had the unique opportunity to test application of both the eHCI and MELE in Lao PDR. Inclusion of direct assessments (i.e., literacy, numeracy, executive function) as an outcome measure, alongside the eHCI, enabled comparison of the sensitivity of both measures to ECE quality.

Descriptive statistics for MELE, eHCI, and direct assessment measures were assessed to investigate distribution of scales (i.e., identify any floor or ceiling effects). Two-tailed Spearman's rho correlations, appropriate for use with ordinal scales, were used to explore associations between exposure (MELE) and outcome (eHCI, MODEL) variables.

A directed acyclic graph (DAG) was used to depict the hypothesised effect of ECE quality on early childhood development (see Figure 2.4 below). DAGs are commonly used in epidemiologic research to describe the causal relationships between variables, and help to identify confounding variables (i.e., factors that influence both ECE quality and children development).[164]

Figure 2.4. Directed acyclic graph of the association between early childhood education quality and child development



Note. ECE = early childhood education, SES = socioeconomic status

A household SES index was derived using Principal Components Analysis and included the following items: number of household members, assets (e.g., electronics, vehicles, animals), house construction materials, caregiver literacy, access to electricity, receipt of financial aid, and purchasing ability (e.g., school supplies, food.). A continuous factor score was created using the first component for each household which was then divided into quintiles to create five levels of SES.[165] Standard errors were clustered on village as children living in the same village were likely to have more similar levels of SES compared to children living in different villages.

A series of linear regressions were then conducted to explore the association between ECE quality (MELE overall quality and individual domain scores) and both adult reported and directly assessed child development (eHCI and MODEL overall development and domain scores). Models were adjusted for confounding variables including child age, sex, ethnicity, stunting, household SES, and village remoteness. Unstandardised beta coefficients were reported with 95% confidence intervals, and coefficient size together with confidence intervals that did not overlap zero were interpreted as meaningful associations.[156] Analyses were conducted using Stata/SE 17.[166]

2.5 Missing data

Missing data is typical in epidemiological research due to non-response to items or questionnaires, or participant attrition in longitudinal studies. Missingness can be classified as missing completely at random (no systematic differences between observed and missing data; MCAR), missing at random (systematic differences between observed and missing data can be explained by associations with the observed data; MAR), and missing not at random (associations with the observed data is not able to explain all systematic differences between the observed and missing data; MNAR).[167] Failure to account for missingness in analyses can lead to bias in results.[168] Over the past few decades, there has been considerable development in the statistical methods for analysis of data with missing values, including full maximum likelihood estimation, multiple imputation, and weighting adjustment methods.[168] Recent advances demonstrate that there is no best, single method for accounting for missing data in all studies. Rather, choice of method is dependent on the mechanisms of missingness as well as analytical methods employed.

However, as argued by Hughes and colleagues[167] among others[169, 170], there are instances in which a complete case analysis (i.e., one that is restricted to participants with complete information available on variables included in the main analysis) is appropriate, and multiple imputation or alternative methods are not required. While a detailed exploration of mechanisms for missingness in eHCI data across countries was not within scope for this thesis, complete case analysis was employed throughout analyses presented. For instance, in study four (Chapter 6), analysis was limited to children with development (eHCI and MODEL; outcome) and ECE quality (MELE; exposure) scores, as well as the child, family, and village level characteristics included as confounders. Missing data in each research study in this thesis are described in respective chapters.

2.6 Chapter conclusion

This chapter described the methodological approaches employed to address the four research questions explored throughout this thesis, as well as how each study contributes to advancing understanding of how the eHCI may be used to facilitate population measurement of early childhood development. Next, Chapters 3 to 6 present the four research studies described; each in the form of a journal article (published or under review) with supplementary materials appended.

Chapter 3: Does the eHCI measure the same underlying constructs (i.e., developmental domains) in seven LMICs, given the necessary local adaptation of the instrument in each country?

3.1 Chapter outline

This chapter presents the first of four papers contributing to this thesis. The article was published in *BMC Pediatrics* in December 2019 and is presented in Appendix B.[171]

As described in Chapter 1, to generate locally relevant data and meaningfully inform supports to promote children's outcomes, a measure of early childhood development should reflect the skills and capabilities considered to be important in any given cultural or contextual setting. The eHCI was designed to be adaptable across diverse settings and therefore underwent a process of local adaptation in each country. While some items and domains differ across countries, adaptations of the tool maintained the same theoretical structure as was originally developed in Tonga (with the exception of Lao PDR). Although the eHCI was designed to measure a specific set of developmental domains, this theoretical factor structure had not been explored using factor analysis methods. Demonstrating a consistent internal factor structure is one important aspect of the comprehensive evaluation of a measurement instrument's validity and reliability. This study explored whether eHCI data across seven LMICs fit the theoretical structure of the instrument (i.e., nine developmental domains, or eight domains in Lao PDR), providing evidence for the tool's construct validity and internal reliability.

Confirmatory factor analyses and internal consistency coefficients demonstrated that after local adaptation, translation, and different implementation methods across countries, the eHCI maintained the same factor structure of nine theoretically based developmental domains in each countries: Physical Health, Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, Approaches to Learning, Numeracy, Reading, and Writing (with the exception of Lao PDR, for which the local adaptation did not include the Physical Health domain). While some domains were found to be operating consistently across countries (e.g., Numeracy, Reading), others tended to demonstrate greater variation across country samples (e.g., Verbal Communication, Social and Emotional Skills) highlighting the important influence of culture on the measurement of children's development. Further, additional analyses stratified by children's age (i.e., 3 to 5 years) and respondent education provided insights into how these characteristics may play a role in the functioning of the eHCI.

Finally, most domains demonstrated adequate internal reliability within all countries. However, two domains (Physical Health, Perseverance) exhibited less than satisfactory internal reliability in multiple countries. These findings may be due to a large proportion of reverse-scored items in these domains, which could be measuring constructs separate to those intended. Nevertheless, local adaptation of the tool in each country deemed all items important in capturing the skills and capabilities of a well-developing child in their respective contexts, and thus it was not the intention of this research to exclude items on the basis of psychometric results.

Overall, findings from this study lend support to the aims of the eHCI in being adaptable and applicable for use within a range of LMICs to facilitate population measurement of children's early development. Subsequent psychometric analyses, presented in chapters to follow, were conducted based on the confirmed structure (domains) of the instrument.

3.2 Statement of authorship

Title of Paper	Measuring early childhood development in multiple contexts: the internal factor structure and reliability of the early Human Capability Index in seven low and middle income countries
Publication Status	Published
Publication Details	Sincovich, A., Gregory, T., Zanon, C., Santos, D.D., Lynch, J., & Brinkman, S.A. (2019). Measuring early childhood development in multiple contexts: The factor structure and reliability of the early Human Capability Index in seven low and middle income countries. *BMC Pediatrics*, 19(471). doi: 10.1186/s12887-019-1852-5.

Principal Author

Name of Principal Author (Candidate)	Alanna Sincovich		
Contribution to the Paper	Contributed to the acquisition of data. Led research design, data analysis, interpretation of results, and drafting of the manuscript. Revised and finalised manuscript.		
Overall percentage (%)	85		
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	14/07/2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that: (i) the candidate's stated contribution to the publication is accurate (as detailed above); (ii) permission is granted for the candidate in include the publication in the thesis; and (iii) the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Tess Gregory		
Contribution to Paper	Supported study design, analysis plan, and interpretation of results. Reviewed the manuscript.		
Signature		Date	07/07/2022

Name of Co-Author	Cristian Zanon		
Contribution to Paper	Contributed to the acquisition of data. Supported interpretation of results. Reviewed the manuscript.		
Signature		Date	03/08/2022

Name of Co-Author	Daniel Santos		
Contribution to Paper	Contributed to the acquisition of data. Supported interpretation of results. Reviewed the manuscript.		
Signature		Date	14/07/2022

Name of Co-Author	John Lynch		
Contribution to Paper	Supported interpretation of results. Reviewed the manuscript.		
Signature		Date	04/07/2022

Name of Co-Author	Sally Brinkman		
Contribution to Paper	Contributed to the acquisition of data. Supported study design, analysis plan, and interpretation of results. Reviewed the manuscript.		
Signature		Date	13/07/2022

3.3 Measuring early childhood development in multiple contexts: the internal factor structure and reliability of the early Human Capability Index in seven low- and middle-income countries

3.3.1 Abstract

The fourth year of the Sustainable Development Agenda era calls for countries to continue to invest not only in interventions and policies that will promote global equity and sustainability, but also in the monitoring systems required to track progress against these targets. A more pragmatic solution to measuring children's early development in low- and middle-income countries in particular, is required. This study explores the psychometric properties of the early Human Capability Index (eHCI), a population measure of holistic development for children aged 3–5 years, designed with the vision of being flexible and feasible for use in low resource and capacity settings. Utilising data from seven low- and middle-income countries: Brazil (n = 1,810), China (n = 11,421), Kiribati (n = 8,339), Lao People's Democratic Republic (n = 7,493), Samoa (n = 12,191), Tonga (n = 6,214), and Tuvalu (n = 549), analyses explored the internal factor structure and reliability of scores produced by the tool within each country. Confirmatory factor analyses and internal consistency coefficients demonstrated that after local adaptation, translation, and different implementation methods across countries, the eHCI maintained the same factor structure of nine theoretically based developmental domains: Physical Health, Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, Approaches to Learning, Numeracy, Reading, and Writing. Findings support the aims of the eHCI in being adaptable and applicable for use within a range of low- and middle-income countries to facilitate measurement and monitoring of children's early development, as is required for the tracking of progress towards the Sustainable Development Agenda.

3.3.2 Introduction

Global endorsement of the United Nations Sustainable Development Goals (SDGs), an agenda for which the healthy development of children is central[172], together with burgeoning evidence regarding the value of investing in children's early years[5], have highlighted the need for services and supports that provide children with the opportunity not only to survive, but to developmentally thrive.[29] In turn, the creation and utilisation of instruments that measure such development in children has gained momentum. The early Human Capability Index (eHCI), a population measure designed to capture the holistic development of children aged 3–5 years, represents one such effort. This paper presents preliminary evidence of the psychometric properties of the eHCI and highlights how the tool could make an important contribution to the task of evaluating early childhood policies and programs as well as monitoring children's development in the early years.

Tracking progress towards healthy child development

Ratified by United Nations member countries in 2015, the Sustainable Development Agenda specifies 17 goals and 169 targets to end poverty, mitigate inequality, and protect the planet for a better future[27]. The fourth year of the SDG era calls for countries to continue to invest not only in interventions and policies that will promote global equity and sustainability, but also in the monitoring systems required to track progress against these targets and thus identify those at risk of falling behind. Of particular relevance to early childhood development, SDG 4.2 states that by 2030, countries must ensure that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education. To track progress against this target, countries are required to monitor (i) the percentage of children under 5 years of age who are developmentally on track in health, learning and psychosocial wellbeing, and (ii) national participation rates in early childhood education.

While many countries monitor national enrolment rates in early childhood education[28], few track the status of children's early development. Measuring progress toward SDG 4.2 calls for population monitoring of children's early health and development outcomes. Indeed, child development contributes to a number of SDG targets, including those related to health, gender equity, and poverty reduction, and thus global monitoring of children's early development is key to supporting progress toward the broader Sustainable Development Agenda.[15] Faced with limitations in terms of the measurement instruments available as well as the resources and capacity required to implement such monitoring systems, tracking children's health and development in low- and middle-income countries will be a challenge. In this, however, lies an important opportunity to promote and address the current obstacles associated with measuring children's early development.

A call for new measurement solutions

Measurement of children's development is influenced by culture, language, and theory. What are considered important aspects of and appropriate goals for children's development, as well as what are deemed suitable assessment techniques to capture this information, can vary considerably across cultures.[78, 94, 96] Consequently, although tools need to capture aspects of child development that are important to outcomes throughout the life course, they should also be aligned with local culture and early learning and development frameworks, so that they not only accurately reflect children's capabilities, but also produce information relevant to local policy and practice.[96, 172]

A number of measurement initiatives are currently underway to monitor children's early development at national, regional, and global levels. Some examples include the Early Development Instrument (EDI), the Early Childhood Development Index (ECDI), the Caregiver Reported Early Development Instrument (CREDI), the International Development and Early

Learning Assessment (IDELA), the East Asia-Pacific Early Chid Development Scales (EAP-ECDS), the Malawi Developmental Assessment Tool (MDAT), the Measurement of Development and Early Learning (MODEL), and the Regional Project on Child Development Indicators (PRIDI).[29] Various characteristics of the instruments in use however, including the cost of licensing fees, the level of enumerator training required prior to administration, the time they take to administer, how they are administered, and their applicability and adaptability within different contexts, constitute considerable barriers to their utilisation. This is especially the case in contexts where resources and capacity are limited. To overcome these challenges, international leaders in early childhood have called for a more pragmatic and reliable solution to measuring children's early development in low- and middle-income countries in particular. It was against this background that the eHCI was developed.

Early Human Capability Index

Designed to capture key aspects of holistic development in children aged 3–5 years, the eHCI was developed with the vision of being feasible for use in low resource and capacity settings while having the ability to capture change in children's development over time.[99] The tool includes approximately 60 items (dependent upon country adaptation) spanning nine developmental domains (Physical Health, Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, Approaches to Learning, Numeracy, Reading, and Writing) and can be completed via adult report (e.g., by children's caregivers, teachers, or early childhood practitioners) in less than 10 min. The eHCI requires minimal resources to be implemented; the tool is available for anyone to use free of charge, little enumerator training is required, and it can be completed quickly and easily by any adult who knows the child. Further, the tool was designed so that it can be easily adapted and utilised within diverse contexts for a range of purposes, including population monitoring, the evaluation of effects of early childhood policies and programs, as well as longitudinal studies seeking to predict children's future capabilities.

Development of the eHCI

The eHCI was originally developed in 2013 for the purposes of evaluating a program designed to support children and families to be better prepared for school in Tonga.[100] Consultations were undertaken to understand locally, what good child development at school entry looks like, first establishing broad areas (i.e., domains) of development and then identifying specific items within these areas. On the basis of consultations and the child development literature, a draft instrument was developed and independently reviewed by child development experts. Once translated into Tongan, stakeholders originally consulted reviewed the instrument to ensure content and face validity. Piloting was then conducted to determine respondent understanding of items, efficiency of data collection methods, if scale distributions were discriminating between children as theoretically expected, as well as any floor or ceiling effects (i.e., if items were too hard or too easy) for the targeted range of children aged 3–5 years. After revisions based on findings from the pilot as well as a final review by local stakeholders, the eHCI was implemented nationally. Exploration of the psychometric properties of scores produced by the eHCI census in Tonga demonstrated adequate discriminant validity (comparison of mean scores across children grouped by demographic characteristics met theoretical expectations, i.e., older children scored higher than younger children, girls received slightly higher scores than boys, children of more educated mothers received higher scores than those of less educated mothers) and internal scale reliability (through Rasch analysis).[99]

Utilisation of the eHCI

The eHCI has since been adapted and utilised to support a range of early childhood development projects in several low- and middle-income countries, predominantly across the Asia-Pacific region.[102-107] Similar to a number of tools designed to measure children's development[53, 82, 173], development and adaptations of the eHCI in each new context were informed by a combination of both theoretical conceptualisation as well as local expert consultation regarding the key aspects of children's development that are predictive of future capabilities. Through these consultative processes, content and face validity of the instrument were established and adaptations and translations were ensured to be capturing the true intent of each item.[174] The internal factor structure of the instrument, however, is yet to be explored within multiple countries. Evaluating the psychometric properties of scores produced by a measurement tool is fundamental for its future utilisation and effectiveness.[98] An instrument measuring children's development that lacks in reliability and validity could produce biased scores that lead to ill-informed decisions. With eHCI data now available in multiple countries, work is needed to explore the tool's validity and reliability. An instrument with the properties of the eHCI that produces scores that are psychometrically robust and appropriate for use within diverse settings has potential for global applicability. Such a tool could better enable population monitoring of children's development, as is required for SDG 4.2, particularly in low- and middle-income countries, with the ultimate goal of shaping services and policy to promote global equity of children's health and development.

The current study

This research is a first step in working to establish the psychometric properties of scores produced by the eHCI within different cultures and contexts. Utilising data previously collected from seven low- and middle-income countries, Brazil, China, Kiribati, Lao People's Democratic Republic

(PDR), Samoa, Tonga, and Tuvalu, analyses sought to explore the internal factor structure and reliability of scores produced by the tool within each country. Findings will be used to guide recommendations regarding the reporting of eHCI results moving forward.

3.3.3 Method

Measures

The early Human Capability Index

Completed by an adult who knows the child, the eHCI includes approximately 60 items (dependent upon adaptation) measuring children's development across nine domains: Physical Health, Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, Approaches to Learning, Numeracy, Reading, and Writing. Response options to each item are binary (yes/no or able/unable). The majority of items are positively worded so that the yes/able responses are scored as 1, and the no/unable responses are scored as 0. A small number of items (ranging from 4 in Kiribati and Lao PDR to 6 in Tonga) are negatively worded and thus are reverse scored. Individual item scores in each domain are averaged so that children receive a score for each developmental domain ranging from 0 to 1, with higher scores indicative of better development.

The eHCI underwent a local adaptation process to ensure the tool's content and face validity in each country. Thus, although many items are similar across different adaptations of the eHCI, some items and domains differ across countries. To illustrate, the Perseverance domain is measured by the same 4 items across all adaptations of the eHCI. In contrast, the Physical Health domain varies from 2 items in Brazil, 3 items in Samoa and China, 4 items in Tonga, and 5 items in Tuvalu and Kiribati, while the Laotian version of the eHCI does not capture physical health as a result of local expert consultation. Table 3.1 presents the Tongan eHCI and the percentage of children for whom respondents reported yes/able for each item, while Supplementary Tables 3.1-

3.6 (Appendix C) present the same information for remaining countries, highlighting similarities and differences between adapted versions of the instrument.

Demographic characteristics

In addition to measuring children's development, the eHCI collects information about children's demographic characteristics as well as relevant contextual information. Specifically, respondents provide information about children's age, gender, and special needs status, and then dependent upon country adaptation, they also provide information about children's height and weight, their mother's highest level of education, whether the child has attended preschool, if there are reading materials (i.e., books) in the child's home, and caregivers' engagement in six different types of stimulating activities with their children in the home (e.g., reading a book, playing, counting etc). Variables presented in this manuscript include children's age, gender, preschool attendance, as well as maternal education.

Data collection procedures

Data included in this manuscript were collected from seven countries between 2013 and 17, utilising different sampling techniques and data collection methods. Contextual information regarding data collection procedures in each country are summarised in Table 3.2.

Participants

Characteristics of each country sample are presented in Table 3.3. Samples ranged in size from 549 children in Tuvalu to 12,191 in Samoa, with children ranging in age from 2 to 6 years. Though the eHCI was designed to capture the development of children aged 3–5 years, the tool has also been used to collect data on children who fall slightly outside of this age range. This is a result of varied data collection purposes across countries. For instance, in Lao PDR, 2-year-olds were included in data collection as this dataset serves as the baseline measure for a randomised control trial; younger children needed to be included at baseline to ensure they also fall into midline and endline data collections in years to come. Each country sample had a relatively even split of males and females; maternal education ranged from the majority of children with mothers who had never attended school (30.2%), started (27.3%), or finished primary school in Lao PDR (29.1%), to the majority of children with mothers who had completed secondary school (42.2%) or tertiary studies in Tonga (17.8%); while the percentage of children who had attended preschool ranged from 23.2% in Lao PDR to 100.0% in Brazil.

Statistical analysis

First, confirmatory factor analyses (CFAs) were conducted separately for each country to determine the fit of eHCI data to the theoretical structure of the instrument (i.e., nine developmental domains, or eight domains in the case of Lao PDR). Next, CFAs were conducted separately for children aged 3, 4, and 5 years old in each country (as this was the age range consistent across all countries) to explore any variation in fit based on children's age. Children with missing age data were excluded from this analysis (Brazil n = 2, China n = 56, Kiribati n = 884, Lao PDR n = 852, Tonga n = 53, Tuvalu n = 3). Additional CFAs were conducted for the Lao PDR sample stratified by maternal education, to explore if fit of data to the theoretical structure of the eHCI varied by respondent's level of education. Specifically, the sample was split into two groups: low maternal education (i.e., no school, started primary, finished primary) and high maternal education (i.e., finished secondary, tertiary) and CFAs were conducted separately for each group. Children with missing maternal education data were excluded from this analysis (n = 5). This analysis was conducted for the Lao PDR sample only, as it was the sole sample for which data were available on the education level of all respondents.

Table 3.1. Tongan eHCI items and n (%) children for whom their caregiver/teacher reported yes/able

Domain	Item	Yes/Able	Missing
Physical Health	Is this child sickly or looked after poorly?*	826 (13.3)	5 (0.1)
	Does this child have good hygiene (i.e., always wash their hands after toileting)?	4,929 (79.3)	13 (0.2)
	Does this child have positive habits?	4,073 (65.5)	16 (0.3)
	Does this child know good foods from bad foods?	3,710 (59.7)	25 (0.4)
Verbal Communication	Can this child use a group of words in talking?	5,705 (91.8)	8 (0.1)
	Can this child converse with others?	5,988 (96.3)	8 (0.1)
	Can this child talk about something that he/she has done?	5,619 (90.4)	10 (0.2)
	Can this child give detail with good Tongan words?	3,844 (61.9)	10 (0.2)
	Can this child hold an adult like conversation (for example talkative, always questioning)?	5,078 (81.7)	16 (0.3)
Cultural Knowledge	Shows compassion, understanding and tolerance of others?	5,229 (84.1)	15 (0.2)
	Can this child identify two culturally important foods/dishes?	5,043 (81.1)	11 (0.2)
	Can this child identify two local plants that provide food/fruits?	4,741 (76.3)	14 (0.2)
	Does this child show the Tongan cultural values of humility?	2,892 (46.5)	21 (0.3)
	Does this child show the Tongan cultural values of devotion/commitment/obligation/responsibility?	2,745 (44.2)	16 (0.3)
	Does this child show the Tongan cultural values of reciprocity in relationships?	2,849 (45.8)	14 (0.2)
	Does this child participate in cultural routines?	4,764 (76.7)	11 (0.2)
	Is this child able to say a short prayer?	5,503 (88.5)	16 (0.3)
Social and Emotional	Is this child happy to share their toys and belongings?	5,296 (85.2)	9 (0.1)
	Does this child take care of their own things?	4,579 (73.7)	4 (0.1)
	Does this child demonstrate respect for adults?	3,870 (62.3)	11 (0.2)
	Does this child demonstrate respect for other children?	3,945 (63.5)	11 (0.2)
	Does this child accept responsibility for their actions?	4,264 (68.6)	9 (0.1)
	Is this child considerate of other people's feelings?	4,274 (68.8)	6 (0.1)
	Does this child repeatedly do something wrong even though he/she has been told to stop?*	3,612 (58.1)	9 (0.1)
	Is this child always helpful?	5,445 (87.6)	6 (0.1)
	Is this child friendly to other children?	5,644 (90.8)	9 (0.1)
	Does this child kick, bite or hit adults or other children?*	2,294 (36.9)	12 (0.2)
	Is this child impatient?*	3,868 (62.2)	11 (0.2)
	Does this child always understand the difference between acceptable and unacceptable behaviour?	4,522 (72.8)	8 (0.1)
	Does this child follow simple directions on how to do something?	5,210 (83.8)	6 (0.1)

Perseverance	Does this child always perform tasks independently?	4,583 (73.8)	8 (0.1)
	Does this child always keep at a task until they are finished?	3,348 (53.9)	9 (0.1)
	Does this child need constant reminding to finish something off?*	4,468 (71.9)	10 (0.2)
	Does this child get easily distracted from a task?*	4,664 (75.0)	14 (0.2)
Approaches to Learning	Does this child show more curiosity about something new in comparison to something familiar?	4,976 (80.1)	10 (0.2)
	Does this child investigate/explore the function of a new toy/game/puzzle or object?	4,900 (78.8)	9 (0.1)
	Is this child always wanting to learn new things?	4,850 (78.0)	12 (0.2)
	When in an unfamiliar environment with a familiar person, does this child feel free to explore?	3,611 (58.1)	15 (0.2)
	Is this child diligent in their approach to a new job or task?	3,840 (61.8)	19 (0.3)
Numeracy	Can this child recognise geometric shapes (e.g., triangle, circle, square)?	3,322 (53.5)	13 (0.2)
	Can this child name and identify at least 3 colours?	4,278 (68.8)	8 (0.1)
	Can this child sort and classify objects by common characteristics (e.g., shape, colour, size)?	3,218 (51.8)	12 (0.2)
	Can this child name and recognise the symbol of all numbers from 1 to 10?	3,020 (48.6)	15 (0.2)
	Can this child count to 10?	5,053 (81.3)	7 (0.1)
	Can this child count to 20?	2,073 (33.4)	9 (0.1)
	Can this child count to 100?	430 (6.9)	13 (0.2)
	Does this child know that a horse is taller than a dog?	4,464 (71.8)	11 (0.2)
	Does this child know the order of the day (e.g., morning, then afternoon and then evening)?	2,411 (38.8)	17 (0.3)
	Does this child understand the concepts of yesterday, today and tomorrow?	1,812 (29.2)	14 (0.2)
	Does this child know that a vehicle weighs more than a cup?	4,197 (67.5)	11 (0.2)
	Does this child know that the number 8 is bigger than the number 2?	2,258 (36.3)	16 (0.3)
Reading	Does this child know the sounds of three letters of the alphabet? (phonics)	3,932 (63.5)	16 (0.3)
	Can this child identify at least 3 letters of the alphabet?	3,263 (52.5)	13 (0.2)
	Can this child identify at least 10 letters of the alphabet?	2,232 (35.9)	14 (0.2)
	Can this child hold a book and turn the pages in the right way?	3,811 (61.3)	10 (0.2)
	Can this child follow reading directions (i.e., left to right, top to bottom)?	1,799 (28.9)	9 (0.1)
	Can this child read at least 4 popular words?	2,263 (36.4)	13 (0.2)
Writing	Can this child draw something identifiable? (e.g., a stick person)	3,531 (56.8)	9 (0.1)
	Can this child copy (trace) the shape of a letter (e.g., A, E, F)?	3,392 (54.6)	5 (0.1)
	Can this child write their own name?	2,428 (39.1)	9 (0.1)
	Can this child write short and simple words?	1,452 (23.4)	9 (0.1)
	Can this child write short and simple sentences?	778 (12.5)	9 (0.1)

Note: * = reverse scored items

Table 3.2. Data collection contexts and procedures

	Country context	Year/s	Respondent/s	Method	Data collection purpose	Sample
Brazil	Occupies a large area on the eastern coast of South America. In 2017, population was approx. 209000 and GNI was USD8600 per capita.	2015	Preschool teachers	Pen and paper	Adaptation of the eHCI for the Brazilian context[107]	Children from one city in Southwest Brazil
China	In East Asia, the world's most populous country. In 2017, population was approx. 1.4 billion and GNI was USD8690 per capita.	2015/16	Preschool teachers and caregivers	Pen and paper	Evaluate inequality in children's outcomes across population groups[105]	Children from two provinces across Northern China
Kiribati	Comprised of 33 coral atolls and isles in the Central Pacific. In 2017, population was approx. 116000 and GNI was USD3010 per capita.	2017	Preschool teachers and caregivers	Tablet	National baseline of child development to guide policy and programs[106]	Population; aimed to collect data for all children aged 3-5 years nationally
Lao PDR	A landlocked country in the Southeast of Asia. In 2017, population was approx. 6.9 million and GNI was USD2270 per capita.	2015/16	Caregivers	Tablet	Baseline data for an RCT designed to support children's early development[102]	Children from five provinces across Northern Lao PDR
Samoa	Comprised of 2 main islands in the South Pacific. In 2017, population was approx. 196000 and GNI was USD4090 per capita.	2016	Preschool teachers and caregivers	Tablet	National baseline of child development to guide policy and programs[103]	Population; aimed to collect data for all children aged 3-5 years nationally
Tonga	An archipelago of more than 170 South Pacific islands (36 inhabited). In 2017, population was approx. 108000 and GNI was USD4010 per capita.	2013/14	Preschool teachers and caregivers	Pen and paper	Baseline data for an RCT designed to support children's school readiness[99]	Population; aimed to collect data for all children aged 3-5 years nationally
Tuvalu	An island country in the South Pacific, comprising 9 small islands. In 2017, population was approx. 11000 and GNI was USD4970 per capita.	2015	Preschool teachers and caregivers	Pen and paper	National baseline of child development to guide policy and programs[104]	Population; aimed to collect data for all children aged 3-5 years nationally

Note. GNI = Gross National Income. RCT = Randomised Control Trial. Population and Gross National Income figures sourced from World Bank[121]

Table 3.3. Sample descriptive characteristics

		Brazil n (%)	China n (%)	Kiribati n (%)	Lao PDR n (%)	Samoa n (%)	Tonga n (%)	Tuvalu n (%)
Child gender								
	Male	853 (47.1)	5,587 (48.9)	4,269 (51.2)	3,824 (51.0)	6,293 (51.6)	3,247 (52.3)	272 (49.5)
	Female	855 (47.2)	5,338 (46.7)	3,915 (46.9)	3,669 (49.0)	5,898 (48.4)	2,967 (47.7)	277 (50.5)
	Missing	102 (5.6)	496 (4.3)	155 (1.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Child age								
	2 years	57 (3.1)	54 (0.5)	948 (11.4)	1,410 (18.8)	1,159 (9.5)	13 (0.2)	1 (0.2)
	3 years	571 (31.5)	3,396 (29.7)	2,185 (26.2)	1,749 (23.3)	4,377 (35.9)	1,609 (25.9)	163 (29.7)
	4 years	760 (42.0)	3,329 (29.1)	2,136 (25.6)	1,867 (24.9)	4,616 (37.9)	2,058 (33.1)	180 (32.8)
	5 years	420 (23.2)	3,360 (29.4)	2,013 (24.1)	1,599 (21.3)	2,039 (16.7)	2,038 (32.8)	195 (35.5)
	6 years	0 (0.0)	1,226 (10.7)	173 (2.1)	16 (0.2)	0 (0.0)	443 (7.1)	7 (1.3)
	Missing	2 (0.1)	56 (0.5)	884 (10.6)	852 (11.4)	0 (0.0)	53 (0.9)	3 (0.5)
Child preschool attendance								
	Yes	1,810 (100.0)	9,159 (80.2)	7,665 (91.9)	1,738 (23.2)	4,657 (38.2)	2,701 (43.5)	498 (90.7)
	No	0 (0.0)	2,176 (19.1)	674 (8.1)	5,755 (76.8)	7,534 (61.8)	3,513 (56.5)	51 (9.3)
	Missing	0 (0.0)	86 (0.8)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Maternal Education								
	No school	-	307 (2.7)	-	2,265 (30.2)	-	-	-
	Started Primary	-	1,242 (10.9)	714 (8.6)	2,045 (27.3)	-	-	-
	Finished Primary	-	3,372 (29.5)	1,438 (17.2)	2,182 (29.1)	222 (1.8)	81 (1.3)	47 (8.6)
	Started Secondary	-	3,563 (31.2)	2,319 (27.8)	-	603 (4.9)	2,399 (38.6)	89 (16.2)
	Finished Secondary	-	1,358 (11.9)	2,710 (32.5)	754 (10.1)	10,037 (82.3)	2,621 (42.2)	107 (19.5)
	Tertiary	-	751 (6.6)	785 (9.4)	242 (3.2)	1,329 (10.9)	1,107 (17.8)	117 (21.3)
	Missing	-	828 (7.2)	373 (4.5)	5 (0.1)	0 (0.0)	0 (0.0)	189 (34.4)
Total n		1,810	11,421	8,339	7,493	12,191	6,214	549

Note. Maternal education data were not collected in Brazil. In Samoa, Tonga, and Tuvalu, when responding to the maternal education item, respondents could select only one response option pertaining to primary school and so the proportions represented against the 'finished primary school' category for these countries may include a combination of children for whom their mother either started or finished primary school.

Goodness-of-fit indices including $\chi 2$ ($p > 0.05$ indicates good fit[144]), root mean square error of approximation (RMSEA; values ≤ 0.06 indicate good model fit while those between 0.06 and 0.08 indicate satisfactory fit[145, 146]), the Comparative Fit Index and the Tucker-Lewis Fit Index (CFI and TLI, respectively; values ≥ 0.95 indicate good fit while values between 0.90 and 0.95 indicate satisfactory fit[145-147]), as well as standardised factor loadings (≥ 0.4 considered high and thus deemed a good indicator of the underlying construct[148]) were used to evaluate model fit. CFAs were conducted in MPlus 7.31[141] utilising polychoric correlation matrices and the weighted least squares mean and variance adjusted (WLSMV) estimation method, both of which are deemed most appropriate for use with binary-type data such as that of the eHCI[142-144], as well as oblique (geomin) factor rotation which assumes correlations amongst factors.[144]

The internal reliability of eHCI domains was also examined for each country, which is often conducted in conjunction with factor analysis to measure how interrelated a set of items are and thus how well they, collectively, measure the underlying construct of focus.[149] Although the majority of similar research assesses internal reliability using Cronbach's alpha[82, 173], increasingly, ordinal reliability coefficients, specifically ordinal alpha, are deemed to be more appropriate in the case of evaluating the internal reliability of scales including items with binary response options in particular.[150] As such, ordinal alpha coefficients were calculated as well as Cronbach's alpha (≥ 0.70 deemed acceptable for both coefficients[113, 150]) to allow for comparison with previous research. These analyses were conducted using the package 'psych' in R-Studio.[153]

3.3.4 Results

Model fit indices yielded from a CFA for each country are presented in Table 3.4. Although a statistically significant $\chi 2$ indicates poor model fit, this test is sensitive to sample size and thus large samples are likely to yield significant $\chi 2$ values.[145, 175] Considering country samples were large, we instead rely more heavily on other fit indices. While χ2 values were statistically significant in all countries, RMSEA values indicated good model fit (i.e., ≤ 0.06) in each country with the exception of Samoa, for which the RMSEA value (0.07) indicated satisfactory fit. CFI and TLI values were ≥ 0.90 in all countries, again indicating satisfactory fit of eHCI data to the theoretical structure of the instrument (i.e., 9 developmental domains, or 8 domains in Lao PDR). Table 3.4 also presents model fit indices from CFAs conducted separately for 3, 4, and 5-year-old children, demonstrating relatively inconsistent results across countries. For instance, although model fit indices indicated better fit of data to the theoretical structure of the eHCI for 5-year-old children in Brazil and Samoa, slightly better fit was observed for 3-year-olds in Tonga and Tuvalu, relative to that of older children.

Table 3.4. Model fit indices from confirmatory factor analyses in each country

		X^2 (df)	RMSEA (90% CI)	CFI	TLI
Brazil	Full sample	4595.48 (1503)	0.034 (0.033-0.035)	0.94	0.94
	3-year-olds	3772.55 (1394)	0.055 (0.053-0.057)	0.75	0.74
	4-year-olds	4413.833 (1503)	0.050 (0.049-0.052)	0.82	0.80
	5-year-olds	2157.71 (1448)	0.034 (0.031-0.037)	0.92	0.92
China	Full sample	35272.80 (1616)	0.043 (0.042-0.043)	0.92	0.92
	3-year-olds	10356.63 (1616)	0.040 (0.039-0.041)	0.94	0.93
	4-year-olds	12039.11 (1616)	0.044 (0.043-0.045)	0.87	0.86
	5-year-olds	9892.98 (1616)	0.039 (0.038-0.040)	0.84	0.83
Kiribati	Full sample	28191.32 (1979)	0.040 (0.039-0.040)	0.93	0.93
	3-year-olds	8200.71 (1979)	0.038 (0.037-0.039)	0.92	0.92
	4-year-olds	7057.17 (1682)	0.039 (0.038-0.040)	0.91	0.91
	5-year-olds	6245.19 (1979)	0.033 (0.032-0.034)	0.91	0.91
Lao PDR	Full sample	22116.62 (1456)	0.044 (0.043-0.044)	0.90	0.90
	3-year-olds	4744.53 (1456)	0.036 (0.035-0.037)	0.87	0.86
	4-year-olds	5528.15 (1456)	0.039 (0.038-0.040)	0.86	0.85
	5-year-olds	5379.22 (1456)	0.041 (0.040-0.042)	0.89	0.89
Samoa	Full sample	89517.11 (1674)	0.066 (0.065-0.066)	0.94	0.93
	3-year-olds	31146.12 (1616)	0.065 (0.064-0.065)	0.93	0.93
	4-year-olds	34712.58 (1674)	0.065 (0.065-0.066)	0.94	0.93

	5-year-olds	14613.05 (1674)	0.062 (0.061-0.062)	0.94	0.94
Tonga	Full sample	23936.63 (1793)	0.045 (0.044-0.045)	0.93	0.93
	3-year-olds	6172.07 (1793)	0.039 (0.038-0.040)	0.91	0.91
	4-year-olds	7724.07 (1793)	0.040 (0.039-0.041)	0.92	0.92
	5-year-olds	7839.75 (1793)	0.041 (0.040-0.042)	0.90	0.90
Tuvalu	Full sample	4144.74 (2043)	0.043 (0.041-0.045)	0.95	0.94
	3-year-olds	2540.33 (2043)	0.039 (0.033-0.043)	0.95	0.95
	4-year-olds	2483.74 (1854)	0.043 (0.039-0.048)	0.89	0.88
	5-year-olds	2600.80 (2043)	0.037 (0.033-0.042)	0.89	0.89

Note. $p < .001$ for all models, df = degrees of freedom, CI = confidence interval. Differing df across CFAs within a country highlight instances where items were dropped from the model to enable estimation of model parameters. In Brazil, read6 and write6 were dropped from the 3yo CFA, and write1 was dropped from the 5yo CFA. In Kiribati, all phys items were dropped from the 4yo CFA. In Samoa, phys2 was dropped from the 3yo CFA. In Tuvalu, phys3, phys4, and phys5 were dropped from the 4yo CFA.

Standardised factor loadings yielded from full sample CFAs are presented in Table 3.5 for Tonga, and Supplementary Tables 3.7-3.12 (Appendix C) for remaining countries. Items had factor loadings ≥ 0.40 across domains and countries with few exceptions. Items that form Numeracy, Reading, and Writing domains in particular had high factor loadings (≥ 0.80 on average) consistently across countries with few exceptions, while the strength of factor loadings for remaining developmental domains tended to be more varied across countries. For example, factor loadings for items in the Verbal Communication domain in China ranged from 0.52–0.90, while those in Lao PDR ranged from 0.73–0.94. Similarly, factor loadings for items in the Cultural Knowledge domain in Lao PDR ranged from 0.58–0.83, while those in Brazil ranged from 0.76–0.99. Reverse scored items in Physical Health, Social and Emotional Skills, and Perseverance domains had weak factor loadings in all countries but Brazil. In contrast, Brazil was the only country in which some non-reverse-scored items had weak factor loadings. Specifically, Reading item 6 for which just under 2% of children were reported to be able to read complex sentences had a factor loading of 0.36, and Writing item 1 for which all but just under 2% of children were reported to be able to scribble on paper had a factor loading of 0.21.

Table 3.5. Factor loadings from confirmatory factor analysis in Tonga

	F1	F2	F3	F4	F5	F6	F7	F8	F9
Phys 1	0.10								
Phys 2	-0.73								
Phys 3	-0.69								
Phys 4	-0.82								
Comm 1		0.81							
Comm 2		0.91							
Comm 3		0.96							
Comm 4		0.85							
Comm 5		0.65							
Cult 1			0.81						
Cult 2			0.88						
Cult 3			0.88						
Cult 4			0.88						
Cult 5			0.88						
Cult 6			0.88						
Cult 7			0.66						
Cult 8			0.73						
Soc 1				0.65					
Soc 2				0.70					
Soc 3				0.88					
Soc 4				0.88					
Soc 5				0.61					
Soc 6				-0.10					
Soc 7				-0.73					
Soc 8				0.69					
Soc 9				0.65					
Soc 10				0.05					
Soc 11				-0.14					
Soc 12				0.81					
Soc 13				0.82					
Persev 1					0.76				
Persev 2					0.77				
Persev 3					-0.11				
Persev 4					-0.46				
Appr 1						0.82			
Appr 2						0.90			
Appr 3						0.90			

Appr 4						0.68			
Appr 5						0.93			
Num 1							0.86		
Num 2							0.85		
Num 3							0.86		
Num 4							0.86		
Num 5							0.79		
Num 6							0.79		
Num 7							0.76		
Num 8							0.84		
Num 9							0.85		
Num 10							0.85		
Num 11							0.84		
Num 12							0.87		
Read 1								0.77	
Read 2								0.93	
Read 3								0.92	
Read 4								0.78	
Read 5								0.89	
Read 6								0.75	
Writ 1									0.86
Writ 2									0.94
Writ 3									0.93
Writ 4									0.96
Writ 5									0.94

Note. Phys = Physical Health, Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

Table 3.6 presents model fit indices for CFAs for low versus high maternal education in Lao PDR, with standardised factor loadings presented in Supplementary Tables 3.13-3.14 (Appendix C). RMSEA, CFI and TLI values indicated better fit of eHCI data to the theoretical structure of the instrument (i.e., 8 domains in Lao PDR) when respondents had a higher level of education. Factor loadings for reverse-scored items in Social and Emotional Skills and Perseverance domains, however, were weak across both education groups.

Table 3.6. Model fit indices from confirmatory factor analyses in Lao PDR grouped by maternal education

	X^2 (df)	RMSEA (90% CI)	CFI	TLI
Low maternal education	19188.69 (1456)	0.043 (0.043-0.044)	0.90	0.89
High maternal education	3222.03 (1456)	0.035 (0.033-0.037)	0.94	0.93

Note. $p < .001$ for all models, df = degrees of freedom, CI = confidence interval.

Finally, Table 3.7 presents internal consistency coefficients for eHCI domains in each country, demonstrating varied results across domains. The Numeracy domain had consistently high internal reliability across countries, with ordinal $\alpha \geq .91$, which is considered high.[150] Verbal Communication (ordinal $\alpha \geq 0.87$), Cultural Knowledge (ordinal $\alpha \geq 0.79$), Social and Emotional Skills (ordinal $\alpha \geq 0.70$), Approaches to Learning (ordinal $\alpha \geq 0.86$), Reading (ordinal $\alpha \geq 0.87$), and Writing (ordinal $\alpha \geq 0.78$) domains also yielded internal consistency coefficients deemed to be acceptable across countries. In contrast, the remaining two domains, Physical Health and Perseverance, demonstrated less than satisfactory internal reliability with ordinal $\alpha < 0.70$ in all countries with the exception of Tuvalu and Kiribati on the Physical Health domain (ordinal α = 0.77 and 0.76, respectively) and Tuvalu (ordinal $\alpha = 0.75$) and Brazil (ordinal $\alpha = 0.75$) on the Perseverance domain.

3.3.5 Discussion

The current study presents the psychometric properties of scores produced by the eHCI in seven low- and middle-income countries. Results demonstrated adequate fit of eHCI data to the theoretical structure of the instrument measuring children's development across 9 domains (or 8 domains in the case of Lao PDR). Overall, findings lend support to the aims of the eHCI in being adaptable and applicable for use within a range of low- and middle-income countries to facilitate measurement of children's early development.[99]

Table 3.7. Internal reliability of eHCI domains

	Brazil		China		Kiribati		Lao PDR		Samoa		Tonga		Tuvalu	
	Cα	Oα	Cα	Oα	Cα	Oα	Cα	Oα	Cα	Oα	Cα	Oα	Cα	Oα
Physical Health	.20	.52	.25	.43	.61	.76	-	-	.34	.48	.47	.56	.61	.77
Verbal Communication	.67	.93	.67	.87	.78	.91	.70	.88	.78	.90	.69	.92	.73	.93
Cultural Knowledge	.52	.89	.80	.90	.84	.93	.53	.79	.89	.95	.79	.89	.82	.92
Social and Emotional	.85	.95	.71	.84	.83	.91	.73	.84	.85	.90	.50	.70	.80	.91
Perseverance	.75	.89	.48	.62	.26	.33	.15	.17	.38	.45	.23	.31	.60	.75
Approaches to Learning	.71	.91	.70	.86	.67	.86	.74	.88	.85	.94	.80	.91	.73	.90
Numeracy	.84	.94	.87	.95	.86	.94	.80	.91	.92	.97	.89	.96	.90	.97
Reading	.61	.87	.79	.90	.79	.91	.72	.91	.87	.95	.82	.92	.82	.93
Writing	.78	.87	.59	.78	.83	.94	.65	.93	.89	.96	.83	.96	.83	.94

Note. Cα = Cronbach's alpha, Oα = Ordinal alpha

Psychometric findings

Samples utilised in this research differed considerably across countries in terms of children's demographic backgrounds, data collection methods and purposes, as well as sampling techniques and sizes. Although it might be argued that such differences present a challenge to exploring and comparing the validity and reliability of the eHCI within multiple contexts, this is the pragmatic nature in which the instrument was intended to be used; for a range of purposes and across varied contexts. As Yapa and Bärnighausen discuss, the resource constraints that come with research in low- and middle-income countries are often the driving force behind creative solutions.[176] As such, we argue that there is strength in that the eHCI was found to demonstrate a common underlying factor structure within the varied contexts in which the instrument has been implemented.

Numeracy, Reading, and Writing domains in particular were found to be working consistently across countries. Items that form these domains had high factor loadings and these scales had high internal reliability across countries. Similar results have been reported for other measures of children's development, for instance, factor analyses of domains that constitute the EDI, a teacher-completed checklist measuring children's holistic development in their first year of school, have demonstrated the Language and Cognitive Development domain (which captures children's literacy and numeracy skills) to have the best fit across multiple countries.[54] Examination of items that form Numeracy, Reading, and Writing domains in the eHCI highlight that little adaptation of these items was required across countries. For instance, the Numeracy domain covers the same concepts of shape, colour, and number recognition, counting ability, and knowledge of numerical concepts such as time and weight, across countries. This might suggest these domains to be the more universal aspects of children's development, indeed such skills have been demonstrated to be important predictors of outcomes throughout the life course[157, 177], and thus results are consistently strong across countries. Such skills are also arguably more easily observable (i.e. it is likely that a caregiver or teacher knows if a child can read or count, as opposed to whether they know if a child is always wanting to learn new things as measured in the Approaches to Learning domain, or if a child knows good from bad foods as captured in the Physical Health domain), which may also have had an influence on results.

In contrast, results across Physical Health, Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, and Approaches to Learning domains tended to demonstrate more variation across country samples. This is unsurprising considering the nature of the skills measured by these domains, including health and hygiene practices, verbal communication abilities, knowledge of culture and culturally acceptable behaviours, social interactions and emotional regulation, as well as how tasks are approached and the ability to complete them, which are aspects of development that would be considered to be more contextually and culturally specific. To illustrate, although Social and Emotional Skills was demonstrated to be one of the distinct domains that the eHCI captures within each country, variation in the strength of factor loadings of items that form this domain between Brazil and Lao PDR might be, in part, explained by cultural differences in social interactions and the expression of emotions between the two countries. Item factor loadings were lower in Lao PDR (on average around 0.65–0.70), an individualistic culture in which emotion is perceived to be experienced internally within an individual, whereas the same item loadings were higher in Brazil (on average around 0.80–0.85), a collectivist culture whereby emotions are thought to occur between people and thus are expressed openly.[178] These results could also be attributed to variation in methodological bias across countries.[179] For instance, acquiescence, the tendency to agree with statements, has been demonstrated to be more common in collectivist cultures.[180] Variation of results across countries in this way highlights the important influence of culture on the measurement of children's development, and the need for tools to capture not only the aspects of children's development that are important predictors of later outcomes, but to also be aligned with local culture in order to produce information that both accurately reflects children's abilities and is relevant to local policy and practice.[96, 172]

Reverse-scored items in the Physical Health, Social and Emotional, and Perseverance domains had weak factor loadings in all countries but Brazil, indicating that, compared to other items, they are poorer measures of the underlying constructs being measured by these domains. These results reflect initial findings from analyses conducted in Tonga following the development of the tool, with Rasch analyses indicating that a number of reverse-scored items did not fit the model well relative to other items.[96] It is possible that enumerators and/or respondents had difficulty in understanding and/or responding to these negatively worded items. Previous research has shown that reverse-scored items tend to load onto a factor separate to the construct they are intended to measure, that instead reflects aspects of item method.[181, 182] This was not observed in Brazil however, and this could be due to sample differences in this context relative to other countries. Specifically, Brazil was the only sample for which only preschool teachers completed the eHCI, as opposed to a combination of children's caregivers and teachers, or caregivers only as in other samples. It might be that children's caregivers and teachers do not respond to the eHCI in the same way, or that a minimum level of education or literacy is required to understand and respond to items. Indeed, results demonstrated better fit of data to the theoretical structure of the eHCI in Lao PDR amongst more educated caregivers who responded to the tool, compared to those less educated. However, weak factor loadings for reverse-scored items were maintained when analyses were run separately for caregivers who had low versus high education. Together, results raise important questions regarding respondent reliability that need to be explored by future research.

When considering children's age in determining how the instrument is operating, results were inconsistent across countries (i.e., in some countries best fit was observed for 3-year-olds and in others for 5-yearolds), and with the exception of Brazil, model fit indices did not vary in magnitude greatly, indicating that he eHCI appears to work relatively consistently across the age range of 3–5 years. Although these results provide some insight into how effective the eHCI is in measuring development across children of different ages, analyses focused on the discriminant validity of the tool, including if items capture a continuum of development for children of different ages, are needed to further explore this question.

Finally, internal reliability results for Physical Health and Perseverance domains indicated that items that form these domains, collectively, are not a good measure of these underlying constructs. It is possible that this is a result of reverse-scored items (which could be measuring constructs separate to that intended) making up a large proportion of these domains (i.e. one in two items in the Perseverance domain are reverse-scored across countries, while for the Physical Health domain this is the case for one in two items in Brazil, one in three items in Samoa and China, one in four items in Tonga, and one in five items in Tuvalu and Kiribati). Nevertheless, local adaptation of the tool in each country deemed all items important to children's early development in their contexts, and thus it was not the intention of the current research to exclude items on the basis of psychometric results. An example is the first item in the Physical Health domain regarding children being frequently sick. Although we would not naturally assume that this item measures a child's skills or capabilities (and subsequently it does not work well in the model), in the contexts of countries of focus whereby illness is common, it was deemed important for the eHCI to provide information regarding children's experience of illness as one aspect of their holistic development.

Study limitations

When interpreting results of the study, it is important to be cognisant of three limitations. First, although the majority of countries studied utilized a census approach to data collection and thus are considered nationally representative, sampling strategies employed in Brazil, China, and

Lao PDR (see Table 3.2) posit that results may differ if eHCI data were to be collected on nationally representative samples in these countries. Next, results indicate that reverse-scored items may not be operating as intended. Beyond analyses presented in this study, however, the information required to be able to explore this further (for example, insight into respondents' understanding of reverse-scored items) is not currently available. Finally, it is important to reiterate that demonstrating consistent internal factor structure and reliability, as has been done in this study, is not complete evidence of a valid tool and must be considered together with results from additional psychometric analyses.

Study implications

Relative to other measures of early childhood development currently utilised, the eHCI requires minimal resources to be implemented. Initial psychometric results suggest that this has not come at the cost of the validity and reliability of the instrument. Demonstrating a consistent internal factor structure and reliability is one important aspect of the comprehensive evaluation of an instrument's validity and reliability. Although not within the scope of the current study, additional work is underway to explore the extent to which eHCI domains can discriminate amongst children's abilities by a range of demographic and contextual variables, are associated with scores on other measures of child development, show reliability amongst respondents, and are able to predict children's future outcomes. A low-burden instrument that is both easily adaptable and psychometrically robust within multiple contexts in this way has potential utility internationally. Indeed, such a tool might better enable population monitoring of children's development, as is required for the tracking of progress towards the Sustainable Development Agenda, particularly in low- and middle-income countries.

In terms of the reporting of and utilisation of data produced by the eHCI, results suggest that eHCI data should continue to be reported across the instrument's 9 theoretically based developmental domains, or 8 domains in the case of Lao PDR. Reporting of children's development across different areas of development in this way enables the identification of areas of both strength and need, and as a result can help to shape more targeted approaches to intervention or policy development. SDG 4.2 however, calls for the monitoring of children who are "developmentally on track", a concept that, as with children's development more broadly, is likely to vary across contexts. As such, if the eHCI is to be recommended to track progress against SDG 4.2 in future, research needs to not only work to further validate the instrument, but also determine how "developmentally on track" might be classified utilising eHCI data.

Conclusion

Initial psychometric results demonstrate that scores produced by the eHCI, after processes of local adaptation, translation and implementation, maintained a similar factor structure of 9 theoretically based developmental domains (or 8 domains in the case of Lao PDR) within a range of low- and middle-income countries. Future research is needed to build on these results and help to determine if the eHCI is able to fulfil its purpose of being a reliable, valid, and feasible tool which can help to facilitate the evaluation of early childhood policies and programs as well as measurement and monitoring of children's development in the early years, particularly in low- and middle-income countries.

Chapter 4: Does the eHCI have adequate validity, and does this vary across seven LMICs?

4.1 Chapter outline

With the internal factor structure and reliability of the eHCI established, the next study focused on investigating the validity of the instrument. This chapter presents the second paper contributing to this thesis. The article was published in *Social Science and Medicine – Population Health* in June 2020 and is presented in Appendix D.[183]

Relative to other population measures of early childhood development available at the time the tool was developed, the eHCI requires minimal resources to be implemented. Specifically, it is available for anyone to use free of charge, little training is required, and it can be completed by an adult who knows the child (e.g., caregiver, teacher) in approximately 10 minutes. It was important to investigate whether such characteristics, designed to maximise feasibility for use at scale, compromised the validity of scores produced by the tool. This study used data from across seven LMICs to explore the convergent, divergent, and discriminant validity (i.e., aspects of construct validity), as well as concurrent validity (i.e., an aspect of criterion validity) of the eHCI, through correlations and linear regressions, seeking to provide further evidence that the eHCI was capturing what it was designed to measure.

The strongest associations across eHCI domains in all countries were observed among those measuring literacy and numeracy, demonstrating convergent and divergent validity. Findings demonstrated that eHCI domain scores discriminated between the development of children as theoretically expected, across countries. For instance, development scores increased with child age and on average, girls scored slightly higher than boys across developmental domains. Being stunted was associated with poorer eHCI scores, while attending preschool was associated with better development. Further, differences between scores among older and younger children, stunted and non-stunted children, and children who did and did not attend preschool, were largest on domains capturing literacy and numeracy skills, with smaller differences on remaining domains. This was aligned with theoretical expectations considering the relationships between children's age, preschool attendance, and stunting with cognitive development. Overall, results provided evidence for the discriminant validity of the tool across countries. Lastly, concurrent validity was evident through associations with direct assessment of children's literacy and numeracy skills in Lao PDR, as hypothesised.

Although associations observed among eHCI domains and with child and family characteristics were relatively consistent overall, some variation in results across countries reflected the context-specific nature of early childhood development measurement, lending support to the need for measures to generate information that reflects local settings. For example, variation in the strength of the association between caregiver-child interactions and eHCI scores may reflect cultural and/or contextual factors, including those related to caregiving practices and early years service provision. Importantly, findings from this study also suggested that although the eHCI was designed to measure development among children aged 3 to 5 years, the instrument may be validly applied to children aged 2 to 6 years. This expands potential application of the tool as discussed in Chapter 7. Additionally, results indicated that both caregivers and teachers provided responses to the eHCI that were valid. This is aligned with the instrument's intended pragmatic, feasible nature in that information can be collected through methods (i.e., respondents) most convenient in any particular setting.

Together, the first two papers in this thesis demonstrated that the eHCI was able to provide valid measurement of early childhood development in a range of countries, among children of a wide age range, with various respondents. With these aspects of reliability and validity established, the next step undertaken in the process of exploring the psychometric properties of the tool was assessment of how well the eHCI predicts children's later outcomes.

4.2 Statement of authorship

Title of Paper	Measuring early child development in low and middle income countries: Investigating the validity of the early Human Capability Index
Publication Status	Published
Publication Details	Sincovich, A., Gregory, T., Zanon, C., Santos, D.D., Lynch, J., & Brinkman, S.A. (2020). Measuring early child development in low and middle income countries: Investigating the validity of the early Human Capability Index. *Social Science and Medicine: Population Health*, 11, article 100613. doi 10.1016/j.ssmph.2020. 100613.

Principal Author

Name of Principal Author (Candidate)	Alanna Sincovich		
Contribution to the Paper	Contributed to the acquisition of data. Led study design, data analysis, interpretation of results, and drafting of the manuscript. Revised and finalised manuscript.		
Overall percentage (%)	85		
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	14/07/2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that: (i) the candidate's stated contribution to the publication is accurate (as detailed above); (ii) permission is granted for the candidate in include the publication in the thesis; and (iii) the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Tess Gregory		
Contribution to Paper	Supported study design, analysis plan, and interpretation of results. Reviewed the manuscript.		
Signature		Date	07/07/2022

Name of Co-Author	Cristian Zanon		
Contribution to Paper	Contributed to the acquisition of data. Supported interpretation of results. Reviewed the manuscript.		
Signature		Date	03/08/2022

Name of Co-Author	Daniel Santos		
Contribution to Paper	Contributed to the acquisition of data. Supported interpretation of results. Reviewed the manuscript.		
Signature		Date	14/07/2022

Name of Co-Author	John Lynch		
Contribution to Paper	Supported study design, analysis plan, and interpretation of results. Reviewed the manuscript.		
Signature		Date	04/07/2022

Name of Co-Author	Sally Brinkman		
Contribution to Paper	Contributed to the acquisition of data. Supported study design, analysis plan, and interpretation of results. Reviewed the manuscript.		
Signature		Date	13/07/2022

4.3 Measuring early child development in low- and middle-income countries: Investigating the validity of the early Human Capability Index

4.3.1 Abstract

Inclusion of early child development in the United Nations Sustainable Development Agenda raises issues of how this goal should be monitored, particularly in low resource settings. The aim of this paper was to explore the validity of the early Human Capability Index (eHCI); a population measure designed to capture the holistic development of children aged 3-5 years. Convergent, divergent, discriminant and concurrent validity were examined by exploring the associations between eHCI domains and child (sex, age, stunting status, preschool attendance) and family (maternal education, home learning environment) characteristics. Analyses were repeated using data from seven low- and middle-income countries: Brazil (n = 1,810), China (n = 11,421), Kiribati (n = 8,339), Lao PDR (n = 7,493), Samoa (n = 12,191), Tonga (n = 6,214), and Tuvalu (n = 549). Correlations and linear regressions provide evidence that within these country samples, the tool is capturing the aspects of early child development that it was designed to measure. Although the tool was intended to measure development of children aged 3-5 years, results suggest it can be validly applied to children aged 2-6 years. The eHCI is free, requires minimal implementation resources, captures development across domains and abilities, and is designed to allow cultural and contextual concepts to be included. The eHCI appears psychometrically robust in diverse country contexts and could enable evaluation of early years policies and programs, as well as monitoring of children's development to track progress towards the Sustainable Development Agenda.

4.3.2 Introduction

Monitoring children's outcomes is key to improving understanding of the early determinants of health and development because it helps identify the supports required to enable children to reach their developmental potential.[51] Tracking child health and development in low- and middle-income countries (LMICs) is a challenge due to lack of appropriate tools and capacity to implement measurement. The early Human Capability Index (eHCI) was designed to measure holistic development in children aged 3-5 years, be feasible for use in low resource settings, and capture locally-relevant early child development (ECD).[184] This paper explores the convergent, divergent, discriminant, and concurrent validity (Figure 4.1) of the eHCI in several LMICs, highlighting how the tool could enable ECD measurement in these contexts.

Figure 4.1. Types of validity explored in the current study

Convergent validity is evident when measures of theoretically related constructs are highly correlated. Divergent validity is evident when measures of theoretically less related constructs are not highly correlated. Discriminant validity is evident when measures produce scores that differ according to participant characteristics, such as sex and age, as theoretically expected. Concurrent validity is evident when measures of constructs are highly correlated with previously established measures of the same construct.

The SDG challenge: characteristics of a useful measure of ECD in LMICs

Sustainable Development Goal (SDG) 4.2 is focused on ECD. To track progress against this target, countries are seeking population measures to better understand their children's early health and development. In addition to being psychometrically robust, such a measure needs to be cost-effective, therefore fees to use the tool, enumerator training required, and administration time need to be minimal.[29] It should cover a range of development as well as levels of ability, and importantly, be sensitive enough to detect changes in children's capabilities.[59] Further, such a tool should be adaptable across diverse cultures and contexts so that it not only accurately reflects children's abilities, but also captures locally-relevant and culturally-influenced aspects of development to inform local policy and practice.[97]

The selection of a tool for the purposes of measuring ECD requires compromise among different priorities and measurement ideals determined by the aims of assessment, the age range of children, and any financial and logistical constraints. For instance, although the direct assessment of ECD[82] is often argued to produce scores with less bias than those through a measure of adult report[185], this method of assessment requires highly trained enumerators, is more time consuming to conduct, more costly to implement and is infeasible for whole-of-population implementation.[90] Although adult report measures are generally more cost-effective as they are quick and simple to administer and do not rely on developmental expertise to be delivered [91], often such tools are based on developmental milestones with a pass or fail outcome, lacking the sensitivity required to detect changes in development over time.[59] Although tools that produce globally comparable results [186] allow for monitoring of ECD across countries, such instruments may not be aligned with local culture or early learning and development frameworks, and thus the information they produce may have limited utility locally.[96] Local 'ownership' of results is crucial for local action to invest in ECD.

Measuring ECD at a population-level

Existing population-level measures of ECD include the Early Development Instrument (EDI) and UNICEF'S Multiple Indicator Cluster Survey-Early Childhood Development Index (MICS-ECDI). The EDI was the first population-level measure of ECD to be implemented in multiple countries including Canada, the United States, Jamaica, and Australia.[187] The 100-item checklist is completed by teachers of children in the first year of full-time school. The EDI formed the basis of the Australian Early Development Census (AEDC), a triennial national census of ECD which has captured over 1.2 million children across Australia to date.[60] The most widely utilised population-level measure of ECD is the MICS-ECDI. Consisting of 10 caregiver-reported items for children aged 3-4 years, the ECDI has been embedded in the MICS (i.e. a household survey) to collect globally comparable ECD information in more than 80 countries.[70] Characteristics of both tools pose challenges to their use in diverse, low resource settings. For instance, the EDI licensing requirements and the specific intention of the ECDI for international comparison limit adaptation to local culture and context.

Early Human Capability Index

The eHCI was developed to facilitate program evaluation in Tonga by monitoring population-level child development.[99, 100] The tool has been adapted to support a range of early child education and development projects in several LMICs.[102-107] Utilising these data, the internal structure of the eHCI was explored, with results demonstrating a similar factor structure of nine theoretically-based developmental domains across countries.[184] The current study examines the tool's convergent, divergent, discriminant, and concurrent validity using data from seven LMICs: Brazil, China, Kiribati, Lao People's Democratic Republic (PDR), Samoa, Tonga, and Tuvalu. We examine associations between eHCI domain scores for evidence of convergent and divergent validity; the ability of eHCI domain scores to discriminate among children of different age, sex, stunting status, preschool attendance, maternal education, and home learning environments (discriminant validity); and associations between scores on eHCI domains and direct assessment of literacy and numeracy for evidence of concurrent validity.

4.3.3 Method

Participants

Data from seven country samples were utilised (Table 4.1). Sample sizes ranged from 549 in Tuvalu to 12,191 in Samoa, with children aged between 2-6 years. Data were collected from 2013-17 through studies funded for program evaluation and population monitoring purposes by local, national, and international agencies that utilised different sampling techniques and data collection methods (Supplementary Table 4.1; Appendix E). In Brazil, data were collected from teachers for children attending all 37 public schools in a city in Southern Brazil, to identify areas of intervention and enable evaluation of ECD programs.[107] In China, data were collected for children in two Northern provinces selected by the China Development Research Foundation on the basis of pre-existing programs in these areas.[105] Respondents were a combination of children's caregivers and teachers, with data collected to explore development across different population groups, as well as to form a baseline of ECD before further intervention. In Lao PDR, data were collected from children's caregivers to form the baseline for a randomised control trial. Five Northern provinces were selected by the Government of Lao PDR based on high levels of poverty, and 14 districts within these provinces were selected on the basis of presence of a district level education office. All villages within selected districts in which at least 20 children resided were sampled, and random sampling methods were used to select 20 households in each village.[102] In Kiribati, Samoa, Tonga, and Tuvalu, a census approach was employed with data collection seeking to capture information from a combination of caregivers and teachers for every child aged 3-5 years nationally.[99, 102-104]

Measures

Early Human Capability Index

The eHCI is unlicensed and free to use, requires minimal enumerator training, and can be completed quickly by an adult familiar with the child. Thus, the eHCI can be implemented feasibly across large populations in low resource settings. The tool captures both positive and negative aspects of how a child is developing, rather than developmental delay only. The eHCI places children on a developmental continuum which improves ability to detect changes in development over time and/or through intervention. Further, the eHCI was designed to be adapted to local culture and context for a range of purposes, including population monitoring, evaluation of early years policies and programs, and longitudinal studies seeking to predict children's future capabilities.

Table 4.1. Sample characteristics for data collected in 7 LMICs from 2013-2017

		Brazil n (%)	China n (%)	Kiribati n (%)	Lao PDR n (%)	Samoa n (%)	Tonga n (%)	Tuvalu n (%)
Sex								
	Female	855 (47.2)	5,338 (46.7)	3,915 (46.9)	3,669 (49.0)	5,898 (48.4)	2,967 (47.7)	277 (50.5)
	Missing	102 (5.6)	496 (4.3)	155 (1.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Age								
	2 years	57 (3.1)	54 (0.5)	948 (11.4)	1,410 (18.8)	1,159 (9.5)	13 (0.2)	1 (0.2)
	3 years	571 (31.5)	3,396 (29.7)	2,185 (26.2)	1,749 (23.3)	4,377 (35.9)	1,609 (25.9)	163 (29.7)
	4 years	760 (42.0)	3,329 (29.1)	2,136 (25.6)	1,867 (24.9)	4,616 (37.9)	2,058 (33.1)	180 (32.8)
	5 years	420 (23.2)	3,360 (29.4)	2,013 (24.1)	1,599 (21.3)	2,039 (16.7)	2,038 (32.8)	195 (35.5)
	6 years	0 (0.0)	1,226 (10.7)	173 (2.1)	16 (0.2)	0 (0.0)	443 (7.1)	7 (1.3)
	Missing	2 (0.1)	56 (0.5)	884 (10.6)	852 (11.4)	0 (0.0)	53 (0.9)	3 (0.5)
Stunted								
	Yes	-	1,602 (14.0)	1,935 (23.2)	3,652 (48.7)	2,332 (19.1)	374 (6.0)	132 (24.0)
	Missing	-	1,342 (11.8)	1,401 (16.8)	852 (11.4)	0 (0.0)	91 (1.5)	98 (17.9)
Preschool								
	Attended	1,810 (100.0)	9,159 (80.2)	7,665 (91.9)	1,738 (23.2)	4,657 (38.2)	2701 (43.5)	498 (90.7)
	Missing	0 (0.0)	86 (0.8)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Maternal Education								
	No school	-	307 (2.7)	-	2,265 (30.2)	-	-	-
	Started Primary	-	1,242 (10.9)	714 (8.6)	2,045 (27.3)	-	-	-
	Finished Primary	-	3,372 (29.5)	1,438 (17.2)	2,182 (29.1)	222 (1.8)	81 (1.3)	47 (8.6)
	Started Secondary	-	3,563 (31.2)	2,319 (27.8)	-	603 (4.9)	2,399 (38.6)	89 (16.2)
	Finished Secondary	-	1,358 (11.9)	2,710 (32.5)	754 (10.1)	10,037 (82.3)	2,621 (42.2)	107 (19.5)
	Tertiary	-	751 (6.6)	785 (9.4)	242 (3.2)	1,329 (10.9)	1,107 (17.8)	117 (21.3)
	Missing	-	828 (7.2)	373 (4.5)	5 (0.1)	0 (0.0)	0 (0.0)	189 (34.4)

Home learning environment							
Read to	-	-	3,930 (47.1)	2,149 (28.7)	2,424 (19.9)	3,566 (57.4)	297 (54.1)
Missing	-	-	6 (0.1)	0 (0.0)	3,470 (28.5)	704 (11.3)	164 (29.9)
Tell story	-	-	5,886 (70.6)	1,903 (25.4)	2,856 (23.4)	3,836 (61.7)	337 (61.4)
Missing	-	-	6 (0.1)	0 (0.0)	3,470 (28.5)	702 (11.3)	165 (30.1)
Sing/dance	-	-	4,964 (59.5)	2,215 (29.6)	5,043 (41.4)	5,025 (80.9)	353 (64.3)
Missing	-	-	6 (0.1)	0 (0.0)	3,470 (28.5)	700 (11.3)	165 (30.1)
Play	-	-	4,774 (57.2)	1341 (17.9)	4800 (39.4)	5180 (83.4)	359 (65.4)
Missing	-	-	6 (0.1)	0 (0.0)	3470 (28.5)	699 (11.2)	165 (30.1)
Go outside	-	-	3,455 (41.4)	1775 (23.7)	4722 (38.7)	5174 (83.3)	338 (61.6)
Missing	-	-	6 (0.1)	0 (0.0)	3470 (28.5)	700 (11.3)	165 (30.1)
Name, count, draw	-	-	3,380 (40.5)	2280 (30.4)	2720 (22.3)	3152 (50.7)	309 (56.3)
Missing	-	-	6 (0.1)	0 (0.0)	3470 (28.5)	701 (11.3)	166 (30.2)
Books in home	-	-	3,948 (47.3)	-	6253 (51.3)	4129 (66.4)	-
Missing	-	-	267 (3.2)	-	0 (0.0)	11 (0.2)	-
Total n	1,810	11,421	8,339	7,493	12,191	6,214	549

Note. Children's height and weight, maternal education, and home learning environment data were not collected in Brazil. Home learning environment data were not collected in China. Information regarding books in the home was not collected in Lao PDR and Tuvalu. In Samoa, Tonga, and Tuvalu, when responding to the maternal education item, respondents could select only one response option pertaining to primary school and so the proportions represented against the 'finished primary school' category for these countries may include a combination of children for whom their mother either started or finished primary school.

The eHCI includes approximately 60 items (ranging from 56 in Lao PDR to 66 in Tuvalu) designed to measure ECD across nine domains: Physical Health, Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, Approaches to Learning, Numeracy, Reading, and Writing. The eHCI underwent a local adaptation process to ensure the tool's content and face validity in each country. Thus, although many items are similar across different adaptations of the eHCI, some items and domains differ across countries. For instance, the Perseverance domain is measured by the same four items across all adaptations of the instrument, while the Physical Health domain varies from 2 items in Brazil to 5 items in Kiribati and Tuvalu, while the Laotian version of the eHCI does not capture physical health.[171] In each country, all items are applied to all children (i.e., rather than a sub-group of items for different age groups). Response options for each item are binary (yes/no, able/unable, can do already/cannot do yet). Most items are positively worded so that the yes/able/can do already responses were scored as 1, and the no/unable/cannot do yet responses were scored as 0. A small number of items (ranging from 4 in Kiribati and Lao PDR to 6 in Tonga) are negatively worded and were reverse scored. Individual item scores were averaged so that children received a score for each domain ranging from 0 to 1, with higher scores indicative of better development.

Demographic characteristics

Children's demographic characteristics were collected alongside the eHCI in each country, including child age, sex, previous or current preschool attendance, and their mother's highest level of education (except in Brazil).

Stunting

Children's height and weight were also recorded at the time of data collection in each country (except in Brazil). Anthropometric measures were converted into World Health Organization

Child Growth Standards height-for-age z-scores, and stunting was defined as a height-for-age z-score < -2.[134]

Home learning environment

Information about caregiver-child engagement in six types of learning activities at home was also collected alongside the eHCI in each country (except in China and Brazil). These binary questions (yes/no) were based on items from the Multiple Indicator Cluster Surveys (MICS) questionnaire for children under five[135] and asked if, in the last 3 days (or in the last 7 days in Lao PDR and Tonga), a member of the household aged 15 years or above had: read books or looked at picture books with the child; told stories to the child; sang songs or danced with the child; played with the child; took the child outside of the home; and named, counted or drew with the child. In Kiribati, Samoa, and Tonga, respondents were also asked if there were reading materials in the home (also based on the MICS with a yes/no response option).

Literacy and numeracy

Concurrent to administration of the eHCI in Lao PDR, children's literacy and numeracy were measured via direct assessment. These assessments were based on items from the Early Grade Reading Assessment (EGRA) and the Early Grade Maths Assessment (EGMA), and adapted to the context in Lao PDR.[102] The EGRA and EGMA have been adapted for use in a number of countries, with the reliability and validity of the EGRA in particular, well established.[136] In Lao PDR, children were assessed on six aspects of early literacy and eight aspects of early numeracy (Supplementary Table 4.2; Appendix E). Correct item responses were scored as 1, and incorrect responses were scored as 0. For domains measured via multiple items, individual item scores were averaged so that children received a score for each domain, ranging from 0 to 1, with higher scores indicative of better literacy and numeracy. All literacy domain scores were averaged to provide a total literacy score, and all numeracy domain scores were averaged to provide a total numeracy score.

Statistical analysis

Convergent and divergent validity were tested by examining patterns of correlation (Spearman's rho) amongst eHCI domains in each country. We predicted that the strongest associations would be observed between eHCI domains measuring children's literacy and numeracy, with smaller correlations between remaining domains. The ability of eHCI domains to discriminate amongst children's development by a range of child and family characteristics was tested in each country. Discriminant ability according to child age (2-6 years), child sex (female, male), child stunting status (yes, no), child preschool attendance (yes, no), maternal education (no school, started primary school, finished primary school, started secondary school, finished secondary school, tertiary education), and home learning environment items (yes, no) were examined using linear regressions. Children with missing data were excluded from relevant analyses. Density plots for eHCI domains were generated to further explore the distribution of domain scores by child and family characteristics in each country, and unstandardised regression coefficients were graphed to examine associations amongst eHCI domains and child and family characteristics in each country. We expected that higher domain scores would be observed among females, older children, those not stunted, children who attended preschool, children of more educated mothers, and children with learning opportunities at home. Concurrent validity was tested by exploring correlations (Spearman's rho) among eHCI domains and literacy and numeracy direct assessment in Lao PDR. We hypothesised that the strongest associations would be observed between direct assessment scores and eHCI domains measuring literacy and numeracy, with smaller correlations between direct assessment scores and remaining eHCI domains.

4.3.4 Results

Correlations amongst eHCI domains are shown in Table 4.2. The largest correlations were observed amongst Numeracy, Reading, and Writing domains in all countries (ranging from r_s

= .54 in Lao PDR to r_s = .85 in Tuvalu between Numeracy and Reading). Smaller correlations were observed between Physical Health and other domains in Brazil and China (r_s = .16 with Writing in China), and between Perseverance and other domains in Kiribati, Lao PDR, Samoa, and Tonga (r_s = .08 with Approaches to Learning in Samoa).

Table 4.2. Convergent and divergent validity: correlations among scores on eHCI domains in 7 LMICs

		Phys	Comm	Cult	Soc	Persev	Appr	Num	Read
Brazil									
	Comm	.13							
	Cult	.10	.37						
	Soc	.08	.26	.19					
	Persev	.14	.38	.32	.43				
	Appr	.12	.43	.30	.29	.29			
	Num	.16	.47	.50	.41	.34	.34		
	Read	.15	.37	.38	.31	.24	.67	.67	
	Writ	.15	.32	.43	.15	.29	.22	.63	.66
China									
	Comm	.20							
	Cult	.32	.48						
	Soc	.32	.43	.70					
	Persev	.12	.28	.34	.45				
	Appr	.23	.54	.57	.54	.34			
	Num	.27	.40	.47	.45	.25	.47		
	Read	.42	.39	.47	.45	.32	.44	.67	
	Writ	.16	.29	.32	.32	.26	.31	.62	.66
Kiribati									
	Comm	.40							
	Cult	.46	.61						
	Soc	.42	.53	.71					
	Persev	.21	.21	.27	.30				
	Appr	.31	.39	.48	.48	.21			
	Num	.43	.53	.61	.52	.24	.46		
	Read	.39	.47	.54	.45	.22	.41	.76	
	Writ	.37	.43	.51	.42	.20	.36	.72	.70

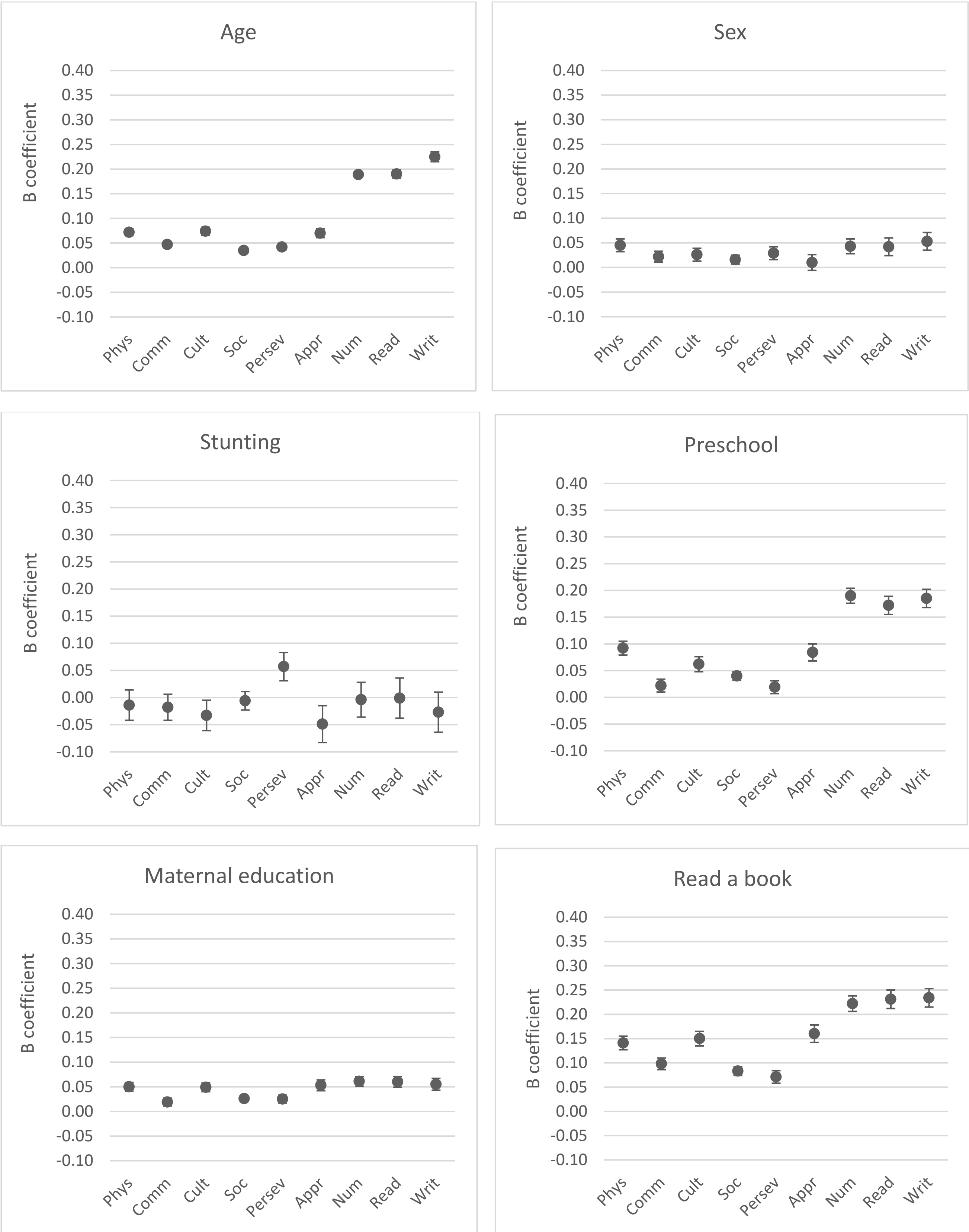
Lao PDR								
Comm	-							
Cult	-	.31						
Soc	-	.31	.46					
Persev	-	.09	.24	.27				
Appr	-	.43	.36	.45	.14			
Num	-	.30	.42	.50	.25	.37		
Read	-	.16	.38	.41	.20	.27	.52	
Writ	-	.12	.35	.42	.21	.28	.50	.55
Samoa								
Comm	.56							
Cult	.57	.61						
Soc	.57	.64	.81					
Persev	.16	.10	.25	.22				
Appr	.45	.63	.53	.61	.08			
Num	.48	.51	.69	.70	.26	.47		
Read	.46	.54	.68	.71	.20	.51	.78	
Writ	.46	.46	.63	.64	.25	.43	.75	.77
Tonga								
Comm	.34							
Cult	.48	.51						
Soc	.46	.38	.48					
Persev	.22	.21	.21	.24				
Appr	.34	.38	.46	.36	.20			
Num	.46	.37	.49	.40	.25	.43		
Read	.39	.33	.42	.36	.25	.38	.76	
Writ	.38	.26	.40	.32	.21	.34	.76	.73
Tuvalu								
Comm	.44							
Cult	.55	.67						
Soc	.48	.51	.67					
Persev	.36	.39	.45	.56				
Appr	.29	.36	.44	.41	.32			
Num	.52	.67	.69	.54	.44	.40		
Read	.56	.66	.69	.55	.42	.43	.85	
Writ	.47	.54	.57	.44	.42	.44	.74	.76

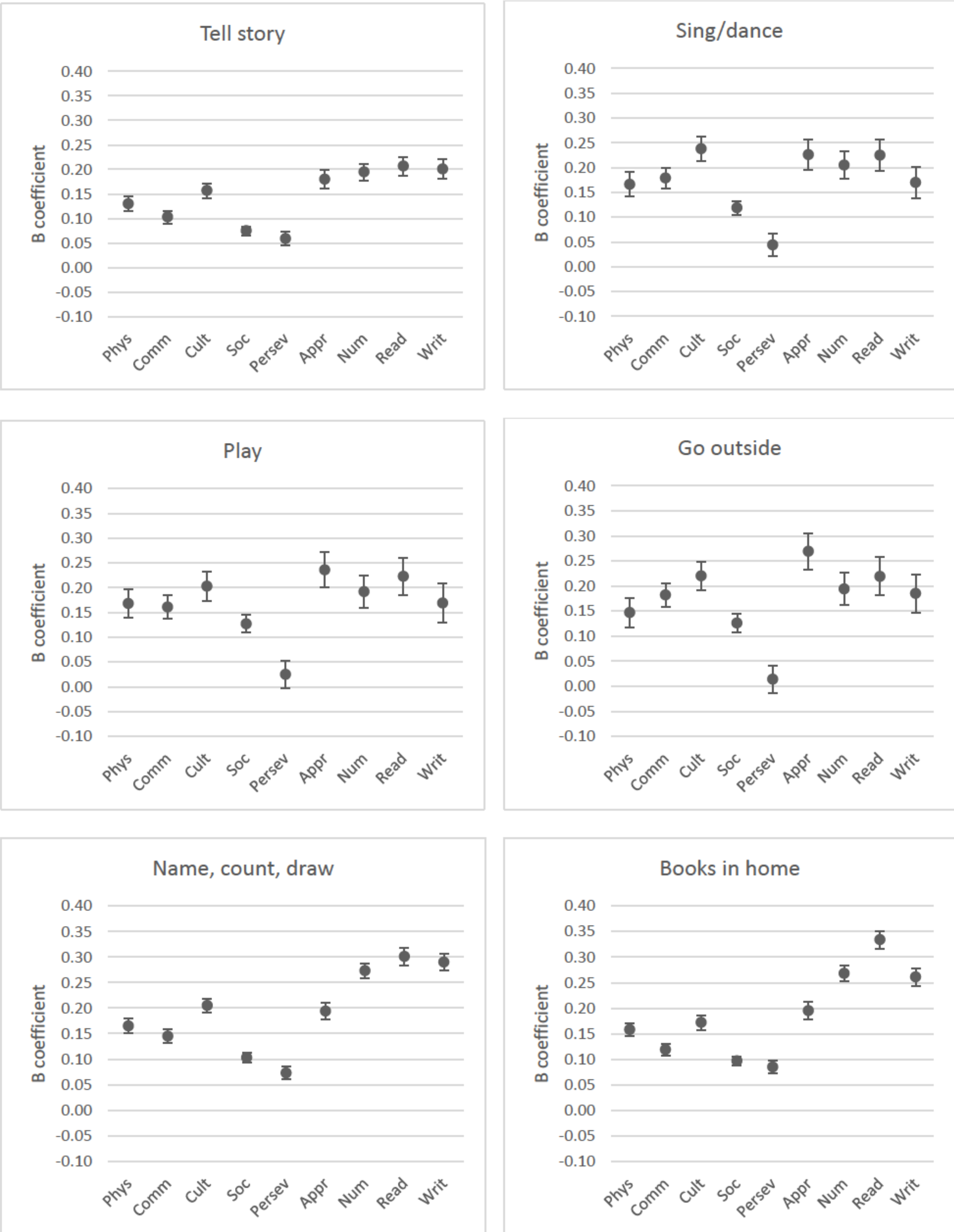
Note. $p < .001$ for all correlations (Spearman's rho). The Lao PDR version of the eHCI does not include a Physical Health domain. Phys=Physical Health, Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing

Mean eHCI domain scores according to child and family characteristics are presented in Supplementary Tables 4.3-4.8 (Appendix E). Regression coefficients demonstrating associations among eHCI domain scores and these characteristics are shown in Supplementary Table 4.9 (Appendix E), with Figures 4.2 and 4.3 providing examples for Tonga and Lao PDR.

Data for remaining countries are shown in Supplementary Figures 4.1-4.5 (Appendix E). Results show a positive association between child age and eHCI domain scores in all countries. Generally, the largest differences between older and younger children were found on Numeracy, Reading, and Writing domains, with smaller differences on remaining domains (e.g., a year increase in age was associated with a score 0.23 (95% CI 0.22, 0.24) points higher on the Writing domain and 0.04 (95% CI 0.03, 0.04) points higher on Social and Emotional Skills in Tonga). Figure 4.4 shows a visual representation of these age gradients for Numeracy and Social and Emotional Skills domains in each country. On average, girls scored slightly higher than boys across domains, except in Brazil. This association was strongest in Tuvalu (e.g., female sex was associated with a score 0.10 (95% CI 0.06, 0.14) points higher on Social and Emotional Skills and 0.11 (95% CI 0.05, 0.17) points higher on Reading). A negative association between child stunting and eHCI domain scores was observed, with stunted children scoring lower across domains in all countries, compared to children not stunted. Differences in scores tended to be larger on Numeracy, Reading, and Writing domains, and smaller on Social and Emotional Skills and Perseverance domains (e.g., being stunted was associated with a score 0.14 (95% CI -0.22, -0.07) points lower on the Writing domain and 0.02 (95% CI -0.07, 0.03) points lower on Social and Emotional Skills in Tuvalu), though this association was less clear in Tonga whereby prevalence of stunting was low. Children who attended preschool had better development than those who did not, with the largest differences in scores on Numeracy, Reading, and Writing domains (e.g., preschool attendance was associated with a score 0.44 (95% CI 0.43, 0.45) points higher on the Numeracy domain and 0.17 (95% CI 0.16, 0.19) points higher on Approaches to Learning in China). Figure 4.5 demonstrates these differences in eHCI scores for children who did and did not attend preschool on the Numeracy and Social and Emotional Skills domains in each country.

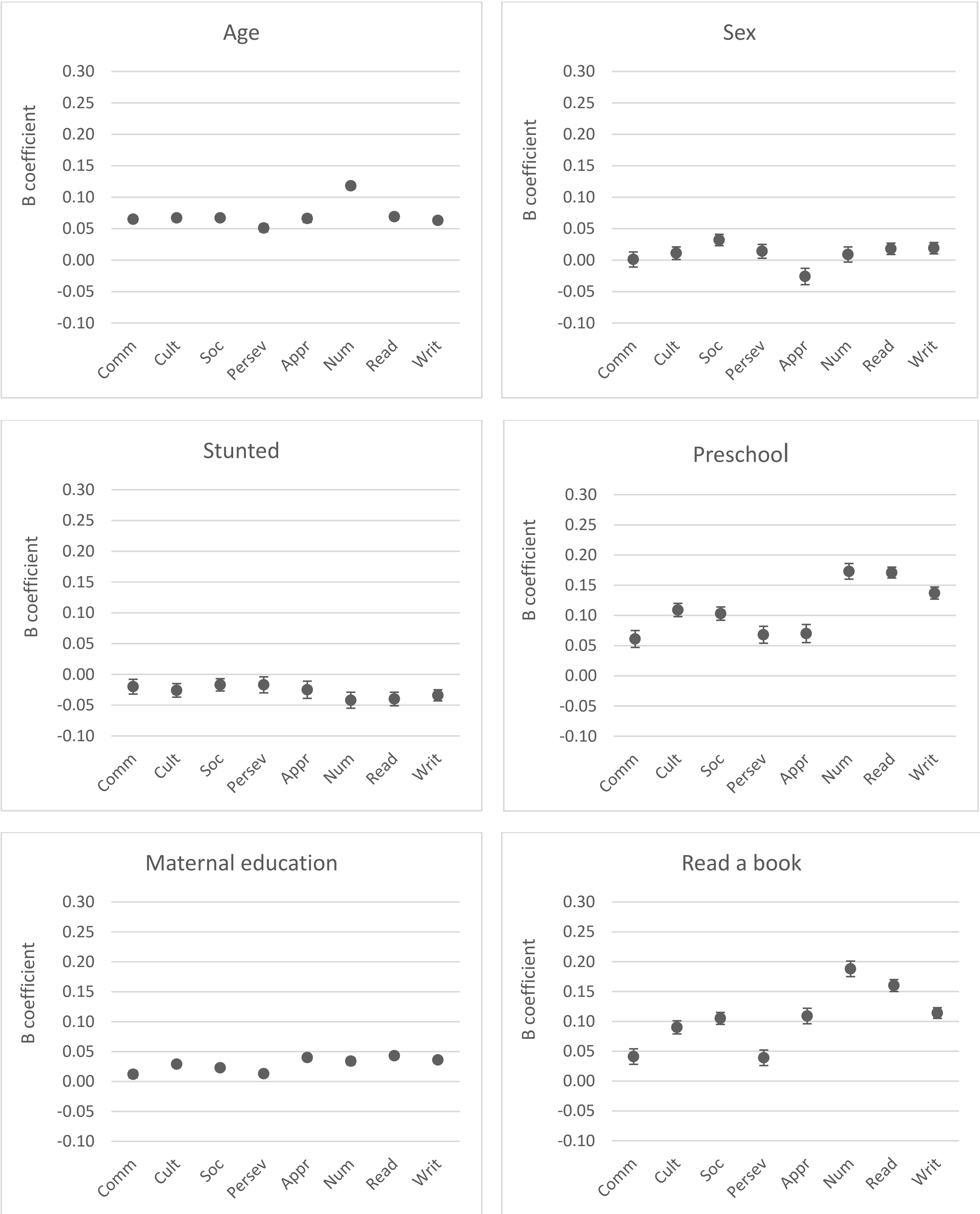
Figure 4.2. Discriminant validity: linear regression coefficients and 95% confidence intervals of eHCI domain scores on demographic and contextual variables in Tonga (n = 6,214, 2013/14)

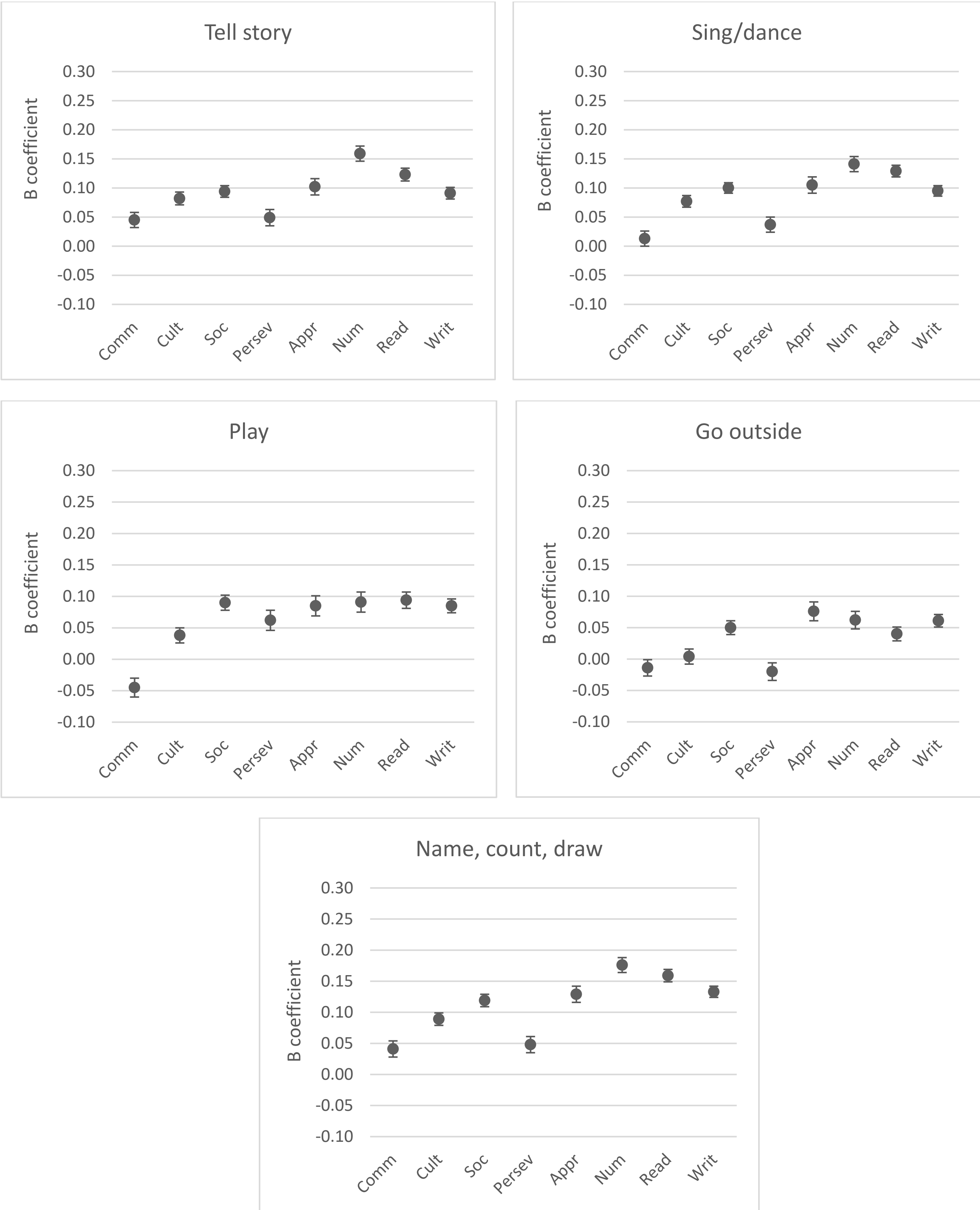




Note. Female, stunted, attended preschool, home learning environment yes = 1; male, not stunted, did not attend preschool, home learning environment items no = 0. Phys = Physical Health, Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

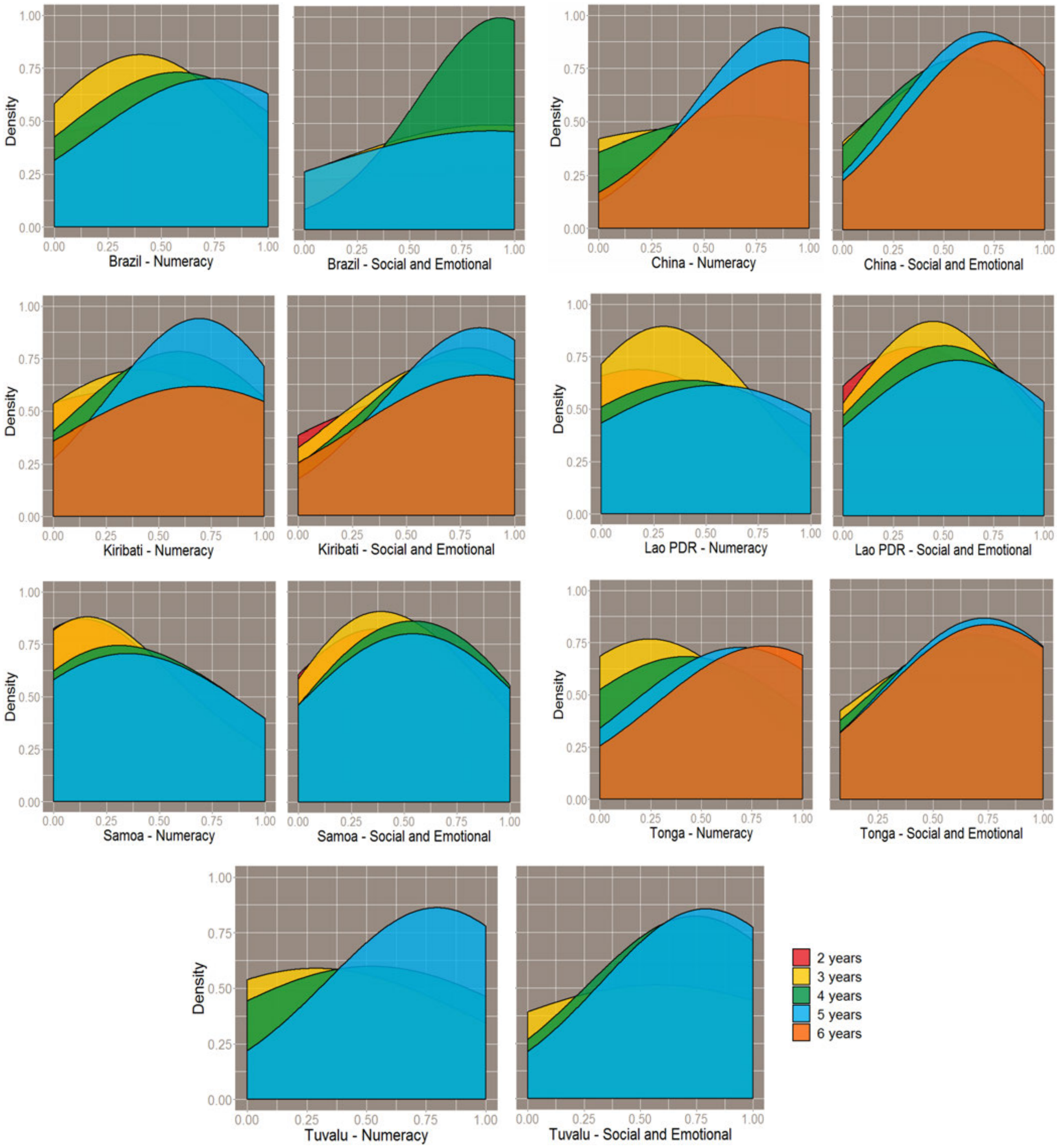
Figure 4.3. Discriminant validity: linear regression coefficients and 95% confidence intervals of eHCI domain scores on demographic and contextual variables in Lao PDR (n = 7,493, 2015/16)





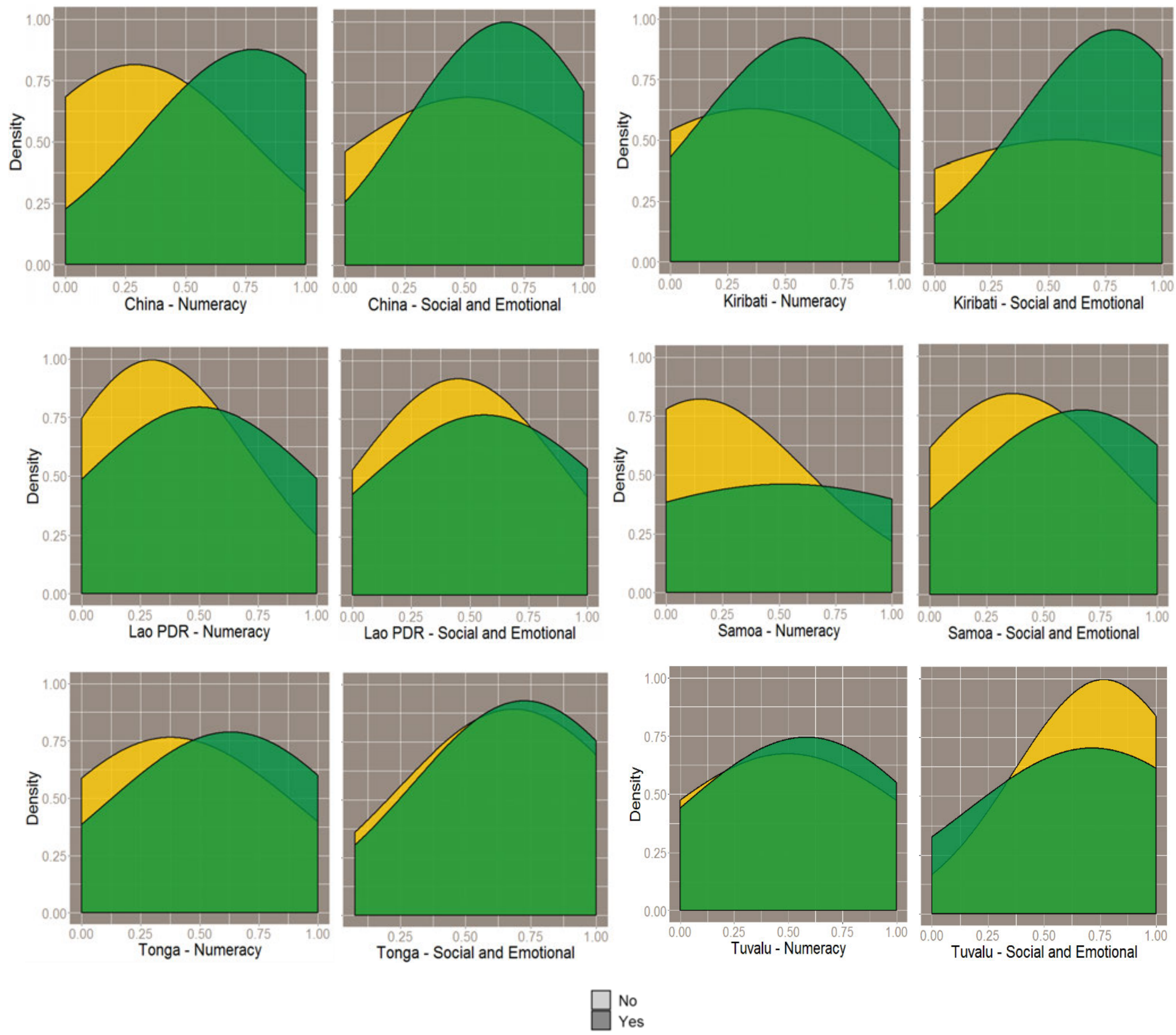
Note: The Lao PDR version of the eHCI does not include a Physical Health domain. Information regarding books in the home were not collected in Lao PDR. Female, stunted, attended preschool, home learning environment items yes = 1; male, not stunted, did not attend preschool, home learning environment items no = 0. Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

Figure 4.4. Discriminant validity: distribution of eHCI Numeracy and Social and Emotional Skills domains by age (years) in 7 LMICs



Note. X-axis = eHCI domain score, y-axis = proportion of children e.g., in Brazil the largest proportion of 3-year-olds scored approximately 0.40 on the Numeracy domain.

Figure 4.5. Discriminant validity: distribution of eHCI Numeracy and Social and Emotional Skills domains by preschool attendance in 7 LMICs



Note. X-axis = eHCI domain score, y-axis = proportion of children e.g., in China the largest proportion of children who attended preschool scored approximately 0.75 on the Numeracy domain. Results are not presented for Brazil as 100% of the sample were attending preschool.

Results also demonstrate a positive association between maternal education and eHCI domain scores, with children of more educated mothers scoring slightly higher across eHCI domains relative to children of less educated mothers in all countries but Tuvalu (e.g., a category increase in maternal education was associated with a score 0.09 (95% CI 0.07, 0.10) points higher on the Reading domain in Samoa). A positive association was also observed between home learning activities and eHCI domain scores in each country. Of all home environment items, the largest differences in scores were observed between children who did and did not have access to books at home on Numeracy, Reading and Writing domains (e.g., having books in the home was associated with a score 0.21 (95% CI 0.19, 0.22) points higher on the Writing domain and 0.12 (95% CI 0.11, 0.13) points higher on Physical Health in Kiribati). Differences in scores between children who were and were not read to, told stories, and named, counted and drew with were also larger across domains, while differences in scores were smaller for remaining activities (e.g. naming, counting, or drawing was associated with a score 0.18 (95% CI 0.16, 0.19) points higher on the Numeracy domain, while taking the child outside was associated with a score 0.16 (95% CI 0.05, 0.08) points higher on Numeracy in Lao PDR).

Table 4.3. Concurrent validity: correlations among eHCI domain scores and literacy and numeracy direct assessment scores in Lao PDR (n = 7,493, 2015/16)

	Total numeracy score	Total literacy score
Verbal Communication	.18	.06
Cultural Knowledge	.29	.24
Social and Emotional	.33	.26
Perseverance	.22	.18
Approaches to Learning	.24	.16
Numeracy	.50	.42
Reading	.40	.39
Writing	.37	.39

Notes: $p < .05$ for all correlations (Spearman's rho).

Correlations between eHCI domains and literacy and numeracy direct assessment in Lao PDR are shown in Table 4.3. Both direct assessment scores had the largest positive correlations with eHCI Numeracy, Reading, and Writing (r_s = .50 between direct assessment and eHCI numeracy domains), and smaller correlations with remaining domains (r_s = .18 between literacy direct assessment and eHCI Perseverance).

4.3.5 Discussion

We examined the convergent, divergent, and discriminant, validity of the eHCI in Brazil, China, Kiribati, Lao PDR, Samoa, Tuvalu, and Tonga, and the concurrent validity of the tool in Lao PDR. Despite different sampling methods and items used across countries, the overall pattern of results suggests the eHCI provides valid measurement of ECD in these contexts.

Evidence for validity of the eHCI

The strongest associations were observed amongst eHCI Numeracy, Reading, and Writing domains in all countries. Literacy and numeracy are often intertwined; indeed some ECD tools combine items measuring these skills into one domain because of their strong relationship.[187] Overall results provided evidence for the convergent and divergent validity of the eHCI.

eHCI domain scores discriminated between children's development according to age and sex in all countries. This is consistent with results of other measures of ECD in LMICs, including the East Asia Pacific Early Child Development Scales (EAP-ECDS), a direct assessment of development in children aged 3-5 years, across Cambodia, China, Mongolia, Timor-Leste, Papua New Guinea, and Vanuatu.[188] Although the eHCI was designed to measure ECD in children aged 3-5 years, results suggest it can be validly applied to children aged 2-6 years. Stunting impairs children's development with effects particularly detrimental to cognitive abilities.[155] In all countries eHCI scores were lower among stunted children, with the largest differences in scores on cognitive domains capturing literacy and numeracy. This is also aligned with results of other tools, such as the Caregiver Reported Early Development Instrument (CREDI) for children aged 0-2 years across 17 low, middle, and high middle income countries.[189] A positive association between early education and ECD was also observed, with differences in scores between children who did and did not attend preschool largest on Numeracy, Reading, and Writing. In contrast, results from a direct assessment, Measure of Development and Early Learning (MODEL), of children aged 4-8 years in Tanzania did not find a relationship between pre-primary education and children's development as was expected[87], whereas cognitive, language and socio-emotional development as measured by the EAP-ECDS was higher amongst children aged 3-5 years who had attended early education in Cambodia, China, Mongolia, and Vanuatu.[190] Higher eHCI scores were generally observed among children born to better educated mothers, except in Tuvalu. This sample represents poor families from a small island nation with a community-based economy. It is likely that advantages of maternal education are transferred to the whole community rather than biological children only and thus the association may not be seen. Similar results were observed when using the International Development and Early Learning Assessment (IDELA) in Ethiopia. A direct assessment of children aged 3-6 years, maternal education did not predict IDELA scores in this context.[191] Stimulating home environments have positive effects on ECD and this pattern was also evident among eHCI scores in all countries, with the strongest associations observed for literacy and numeracy. This finding was aligned with analysis of the CREDI in Brazil amongst children aged 0-2 years.[65]

The strongest associations with direct assessment of literacy and numeracy in Lao PDR were observed for eHCI Numeracy, Reading, and Writing domains as expected. Few similar ECD measures have published concurrent validity evidence with the IDELA an exception. Specifically, the IDELA and the Ages and Stages Questionnaire (ASQ) were used to measure development of children aged 4-5 years in Bangladesh, with results indicating medium correlations between respective IDELA and ASQ domains when examining children's literacy (r =.36) and larger correlations between domains measuring numeracy (r =.48).[82] Concurrent validity of eHCI scores were stronger than those reported for the IDELA. Together, results provide evidence for the concurrent validity of the eHCI in this setting.

Implications of findings

Inclusion of ECD in the Sustainable Development Agenda raises issues regarding how this goal should be monitored. Relative to other measures of ECD, the eHCI can be implemented feasibly in low resource settings, captures development across domains and abilities, and produces information relevant to local policy and practice. Results demonstrate that the eHCI discriminated between the development of children and captured the intended aspects of ECD within a range of LMICs. Together with previous research[184] findings indicate the tool can provide valid measurement of ECD in diverse contexts.

Although associations observed among eHCI domains and with child and family characteristics were relatively consistent overall, some variation in results across countries highlights the context-specific nature of ECD measurement. For instance, varied strength of the association between caregiver-child interactions and eHCI scores could be a reflection of cultural and/or contextual factors, including those related to caregiving practices and early years service provision. This lends support to the need for ECD measures to produce information that reflects local settings. Global comparability and cultural neutrality are the focus for tracking progress toward SDG 4.2[192], however this approach will not have the sensitivity to capture change in ECD as a result of local interventions or policy shifts, and will not reflect aspects of ECD important to the local context. Experts in ECD measurement continue to be challenged by striking a balance between producing globally comparable data and producing information relevant to local policy and practice.[93]

Conclusion

Results demonstrate the eHCI is psychometrically robust in diverse country contexts and could enable the evaluation of early years policies and programs, as well as monitoring of children's development to track progress towards the Sustainable Development Agenda in LMICs. Existing ECD measurement tools range from short adult-report tools designed for population monitoring (e.g., MICS ECDI), to detailed, multi-domain direct assessment batteries designed to aid program evaluation. Findings indicate the eHCI is suitable for both applications. Indeed, the eHCI is a feasible and valid population monitoring measure when applied through either a census or sample approach. The eHCI is free, requires minimal implementation resources, captures development across domains and abilities, and allows local culture and context to be reflected. Results presented in this paper together with previously published evidence demonstrate that the tool is able to provide valid measurement of ECD. The next step in assessing the validity of the eHCI is exploring how well it predicts later outcomes of interest such as academic achievement and social and emotional skills.

Chapter 5: Does the eHCI measured at 2 to 5 years have predictive validity for children's cognitive development (i.e., literacy, numeracy, and executive function) at 6 to 9 years?

5.1 Chapter outline

The first two studies used cross-sectional data from across several LMICs to demonstrate that the eHCI was psychometrically robust. Next, longitudinal data previously collected as part of the ECE Project in Lao PDR, enabled investigation of the predictive validity of the tool. This chapter presents the third paper contributing to this thesis, which was submitted to *PLOS Global Public Health* in April 2022 and is currently under review.

As described in Chapter 1, predictive validity (i.e., a form of criterion validity) refers to the extent to which scores on a measure are related with other constructs measured at a later time, as theoretically expected. Measures of children's development need to be predictive of meaningful outcomes throughout childhood and adolescence, considering results are intended to be used to inform investment in supports to promote children's outcomes. Despite this, limited child development measures have established predictive validity. More broadly, there exists little evidence regarding how early development influences later outcomes among children in LMICs. This study focused on exploring the ability of the eHCI to predict children's cognitive outcomes (i.e., literacy, numeracy, and executive function) four years later. Additionally, predictive ability of the eHCI was compared to that of measures of socioeconomic position and direct assessment measures of children's development. In doing so, findings from this study are also able to contribute to the debate surrounding whether direct assessment of early childhood development provides more accurate information on children's abilities, or if adult-reported measures can provide equally valid results.

Using Receiver Operator Characteristic (ROC) curve analyses, results from this study provide evidence that poor scores on the eHCI in early childhood predict poor cognitive outcomes once children reach primary school age in Lao PDR, establishing predictive validity of the tool in this setting. Specifically, eHCI overall development scores typically had the largest area under the ROC curve for all outcomes, indicating that the summary measure signalled risk for poor future cognitive development with similar ability to direct assessment of child development, as well as measures of socioeconomic position. These results represent the first predictive validity evidence for an adult report population-level measure of child development in LMICs. Indeed, respondents to the eHCI in Lao PDR were children's caregivers, many of which had low literacy and education. Based on findings from the first two studies (e.g., better fit of data to the theoretical structure of the eHCI amongst more educated caregivers, compared to those less educated), we would expect to see stronger predictive ability of the eHCI in settings where the tool is completed by children's teachers or more educated caregivers.

Together with the first two studies, findings indicate that costly and time-intensive measurement approaches (i.e., direct assessment) may not be necessary to inform investments and supports to promote children's outcomes. Rather, the eHCI appears to strike a balance between pragmatism (i.e., feasible for use at scale in low resource settings) and rigour (i.e., comprehensively capturing aspects of development that are predictive of future capabilities), while generating locally relevant information. Although it will be important for future research to further explore predictive validity in additional settings and when predicting difference outcomes (e.g., social and emotional skills, school completion), initial results suggest that data collected using the eHCI could be used to appropriately inform the investments required to promote children's early development and improve later outcomes.

5.2 Statement of authorship

Title of Paper	Child development at ages 2-5 and cognitive outcomes at ages 6-9 in Lao PDR: Measuring children's development using the early Human Capability Index
Publication Status	Under review
Publication Details	The manuscript is currently under review at *PLOS Global Public Health*

Principal Author

Name of Principal Author (Candidate)	Alanna Sincovich		
Contribution to the Paper	Contributed to the acquisition of data. Led study design, data analysis, interpretation of results, and drafting of the manuscript. Revised and finalised manuscript.		
Overall percentage (%)	85		
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	14/07/2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that: (i) the candidate's stated contribution to the publication is accurate (as detailed above); (ii) permission is granted for the candidate in include the publication in the thesis; and (iii) the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Tess Gregory		
Contribution to Paper	Supported study design, analysis plan, and interpretation of results. Reviewed the manuscript.		
Signature		Date	07/07/2022

Name of Co-Author	John Lynch		
Contribution to Paper	Supported study design, analysis plan, and interpretation of results. Reviewed the manuscript.		
Signature		Date	04/07/2022

Name of Co-Author	Sally Brinkman		
Contribution to Paper	Contributed to the acquisition of data. Supported study design, analysis plan, and interpretation of results. Reviewed the manuscript.		
Signature		Date	13/07/2022

5.3 Child development at ages 2-5 and cognitive outcomes at ages 6-9 in Lao PDR: Measuring children's development using the early Human Capability Index

5.3.1 Abstract

The Sustainable Development Agenda, together with burgeoning evidence for the need to support early development, has spurred increased use of population measures of early child development in low- and middle-income countries, despite lack of evidence on whether measures predict later outcomes. This study explored the predictive validity of the early Human Capability Index (eHCI), an adult-reported population measure of early child development. We used longitudinal data (n = 5,222) collected as part of the Early Childhood Education Project in Lao People's Democratic Republic. Receiver Operator Characteristic (ROC) curve analysis was used to estimate the ability of the eHCI and direct assessments of literacy, numeracy, and executive function at ages 2-5 to predict poor development as measured by the same direct assessments at ages 6-9, four years later. The eHCI overall development score typically had the largest area under the ROC curve for all outcomes, ranging from 0.60 to 0.68, indicating that the summary measure signalled risk for poor future cognitive development with similar ability to measures of socioeconomic position. Results highlight that time and resource intensive direct assessment of child development may not be necessary to identify where supports are required to improve children's outcomes. Rather, the eHCI, a pragmatic and freely available tool that generates contextually relevant child development information, could be used to guide early years interventions and policy, as well as evaluate their impacts.

5.3.2 Introduction

Pressing need for interventions to support early development has been propelled by recent estimates that 250 million children under 5 years in low- and middle-income countries (LMICs) are at risk of not achieving their developmental potential.[9] Child development assessment tools are used to monitor the impact of specific interventions, as well as health and education policy more broadly. Child development assessments thus serve as both lead and lag indicators between early years interventions and formation of capabilities later in life, but there is limited evidence on how well scores on such measures predict later outcomes.[29, 193] Measures of child development must have predictive validity for meaningful outcomes throughout school and into adulthood, if the expectation is that investing early to enhance these capabilities will reap longer-term benefits. It is critical that governments, researchers, and communities use measurement tools that generate child development data in a way that is not only feasible and locally relevant[93] but also predictive of future cognitive and social-emotional capabilities.

Although there are various lines of evidence from high income countries that early development sets trajectories throughout childhood and adolescence and across the life course[2, 157] there is much less evidence of these effects in LMICs. Studies in LMICs have typically focused on specific pathology and developmental delay. For instance, considerable research has demonstrated the association between low birth weight and later cognitive outcomes[194], and stunted linear growth and both cognitive and non-cognitive outcomes.[195, 196]

The Sustainable Development Agenda has recently shone a spotlight on tracking children's holistic development. Indicator 4.2.1 requires global monitoring of the proportion of children aged 24-59 months who are developmentally on track in health, learning and psychosocial well-being. UNICEF, custodian agency for reporting progress toward indicator 4.2.1, has promoted the Early Childhood Development Index (ECDI2030) as the global indicator for early child development.[69] However, there have also been other recent efforts to establish multidimensional and psychometrically robust assessments to monitor children's development internationally. Examples include the Global Scale for Early Development (GSED) for use among children aged 0-23 months, and a set of items that capture development of children aged 48-83 months identified through the Measuring Early Learning Quality and Outcomes

(MELQO) initiative.[66] These tools were selected based on their reliability and validity across diverse contexts (e.g., scores on items/tools demonstrate associations with children's demographic and contextual characteristics as expected according to existing evidence). However, the predictive validity of these instruments has yet to be explored.

Use of multidimensional assessments of children's development is becoming more prevalent in LMICs, despite lack of evidence on whether measures predict children's later abilities or academic outcomes. One exception is the International Development and Early Learning Assessment (IDELA), a direct assessment of children's literacy, numeracy, motor, and social-emotional development, for which predictive validity has been recently established.[158] Scores on the IDELA in pre-primary school predicted literacy and numeracy skills 2.5 years later among 2,080 children in Ghana. These findings provide the first evidence of how directly assessed early child development predicts later academic performance in a LMIC. Although direct assessments (i.e., conducted one-on-one with a child) are often considered more objective than adult reported assessments[89], the implementation of population-wide direct assessment of children's development poses significant barriers in resource poor settings.

Building on global efforts advancing population monitoring of child development, this study examines the predictive validity of the early Human Capability Index (eHCI). The eHCI is a population measure of development across nine domains among children aged 2-6 years. Collected via adult report, the eHCI was designed to be feasible for use in low resource settings, be adapted to produce locally relevant results, and predict human capability relevant outcomes. The eHCI is psychometrically robust in diverse contexts and has been adapted to support early childhood education and development projects in several LMICs[99, 102-104, 106, 183, 184]. The eHCI has also contributed items to development of both the ECDI2030 and the GSED. Indeed, many of the items within these instruments overlap, as through continuous improvement, the best performing items from former instruments are selected for use in new measures.

This paper explores the ability of the eHCI as well as direct assessments of literacy, numeracy, and executive function at ages 2-5 to predict literacy, numeracy, and executive function outcomes at ages 6-9 in Lao People's Democratic Republic (PDR) among over 5,000 children. Lao PDR has a population of approximately 7.3 million and is one of the poorest countries in East Asia with a Gross Domestic Product (GDP) ranked 112 internationally.[121]

5.3.3 Method

Data sources

We used data collected as part of the World Bank Early Childhood Education Project (ECE Project; P145544) impact evaluation in Lao PDR.[127] The project aimed to increase access to quality early childhood education (ECE) in disadvantaged villages in Northern Lao PDR. Baseline data were collected from November 2015 to March 2016 as part of three clustered randomized control trials evaluating the impact of ECE on developmental outcomes. Five northern provinces were selected by the Government of Lao PDR based on high levels of poverty, and 14 districts within these provinces were selected based on presence of a district level education office. Randomly selected villages within these districts in which a minimum of 20 children aged 3-5 years resided were selected, and 20 households in each village were randomly sampled.[102] Follow-up data were collected from February to March 2020.

Participants

The current study population included children aged 2-5 years at baseline, for whom data were collected at both time points (n = 5,275). A small number of children were excluded due to missing data on child development measures (i.e., the caregiver and/or child did not take part in assessments) (n = 53), resulting in an analysis sample of 5,222 children.

Measures

Demographic characteristics

Children's demographic characteristics were collected via primary caregiver report at baseline, and included child age (calculated using child's date of birth and date of data collection, or caregiver reported age in years when date of birth was unknown), sex (male, female), highest level of caregiver education (no school, primary school (incomplete), primary school (completed), lower secondary school, upper secondary school, vocational training or tertiary education), and caregiver level of literacy (cannot read at all, can read a little, can read well, vision problem; measured through asking the caregiver to read a short sentence).

Early Human Capability Index (eHCI)

The eHCI was collected via primary caregiver report through a household survey administered by trained enumerators at baseline. The Lao PDR adaptation of the eHCI includes 56 items designed to measure development across 8 domains: Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, Approaches to Learning, Numeracy, Reading, and Writing (Supplementary Table 5.1; Appendix F).[184] Response options for each item are binary (yes/no). Most items are positively worded so that "yes" responses were scored as 1. Negatively worded items were reverse scored. Individual item scores were averaged so that children received a score ranging from 0 to 1 for each domain, with higher scores indicative of better development. An overall development score was derived by taking the average of 8 domain scores, also ranging from 0 to 1, with higher scores indicative of better development. Overall development is the summary indicator most often used in the reporting of eHCI results. The eHCI is psychometrically robust across diverse contexts, including Lao PDR.[83, 105, 183, 184]

Literacy, numeracy, and executive function direct assessment

Children's literacy, numeracy, and executive function were measured via 92 direct assessment items at baseline and follow up. Assessment was administered by a trained enumerator who asked children to complete a series of activities/tasks in a one-on-one setting. Items formed an early working version (2014/5) of the Measurement of Development Early Learning (MODEL;

part of the MELQO initiative)[83], prior to validation studies.[84, 87] The key differences between earlier and later versions of the tool are that the version used in the current study did not include some of the more advanced executive function (e.g., mental transformation) and some of the social and emotional items (e.g., perspective taking) used in the later versions. Literacy and numeracy assessments were based on items from the Early Grade Reading Assessment and the Early Grade Maths Assessment and adapted to the context in Lao PDR.[102] Both assessments have been adapted for use in a number of countries, with well-established reliability and validity.[136] Executive function assessments were based on items from two sub-tests from the Wechsler Intelligence Scale for Children (WISC-IV; digit span forward and digit span backward) and the Head-Toes-Knees-Shoulders task.[137, 138] Taken together, children were assessed on 6 sub-domains of literacy, 9 sub-domains of numeracy, and 3 sub-domains of executive function (Supplementary Table 5.2; Appendix F). Correct item responses were scored as 1, and sub-domain scores were averaged to provide a total domain score for literacy, numeracy, and executive function, ranging from 0 to 1, with 1 indicating better development.

Statistical analysis

Analyses were stratified by age in years at baseline (2-5), as age has a strong positive association with development. Descriptive statistics (mean, standard deviation, interquartile range) for eHCI domain and overall development scores as well as direct assessment domain scores were calculated to explore distribution of scales and the extent to which they captured a range of children's abilities.

Receiver Operator Characteristic (ROC) curves were used to explore the ability of eHCI overall development and direct assessment literacy, numeracy, and executive function domain scores at baseline (ages 2-5) to predict poor development on the same direct assessments at follow up (ages 6-9). Poor development in outcome measures was defined as scores in the bottom 10th percentile (for each age in years stratum), with sensitivity analysis conducted by varying this definition to scores in the bottom 20th percentile. Additional ROC curve analyses explored the ability of 8 individual eHCI domains at baseline (ages 2-5) to predict direct assessment literacy, numeracy, and executive function scores in the bottom 10th percentile at follow up (ages 6-9). To help interpret the size of effects in all ROC curve analyses, we also explored the relationship between two measures of socioeconomic position (caregiver education and literacy) at baseline and poor development at follow up, as these factors would also be expected to predict developmental outcomes. Analyses were conducted using IBM SPSS Statistics 27.[162]

ROC curves have commonly been used to evaluate diagnostic accuracy of a clinical assessment for a dichotomous outcome.[159, 160] In this paper, ROC curves summarise the predictive validity of the eHCI and are presented with the area under the ROC curve (AUC), or C-statistic, to determine the ability of measures at ages 2-5 to predict (discriminate) outcomes at ages 6-9. A series of univariate ROC analyses were conducted for each outcome, with results presented on a single graph to aid the presentation of the predictive ability of each measure. An AUC of 0.5 indicates the development measure at ages 2-5 is no better than chance at discriminating between outcomes at ages 6-9 (i.e., bottom 10% of the various developmental outcomes). The ROC curve graphs sensitivity (% true positive outcomes) as a function of 1-specificity (% true negative outcomes) with the diagonal reference line equal to chance. An AUC of 1.0 indicates the predictor perfectly discriminates between outcomes. Typically, for clinical diagnostic purposes an AUC between 0.5 to 0.7 is considered low, 0.7 to 0.9 moderate, and greater than 0.9 indicates high predictive ability for the measure to discriminate between two outcome groups.[161] However, these should not be applied as strict "rules" and are context dependent.

5.3.4 Results

Table 5.1 shows that, at baseline (ages 2-5) the majority of children had caregivers who had either not attended school (28.6%), started but not completed primary school (26.8%) or had completed primary school (31.0%). Almost half of all children had caregivers who could not read (44.1%), 16.8% had caregivers who could read a little, and 39.0% had caregivers who could read well. Descriptive statistics of child development measures (Supplementary Table 5.3; Appendix F) showed that, at baseline (ages 2-5), children scored lowest on measures of literacy (both eHCI and direct assessment), and mean scores at follow up (ages 6-9) indicated the expected increases in development with age. Generally, there was evidence of ceiling effects (i.e., items too easy) on some eHCI domains among older children (e.g., verbal communication), and floor effects (i.e., items too hard) on some direct assessment domains among younger children (e.g., literacy).

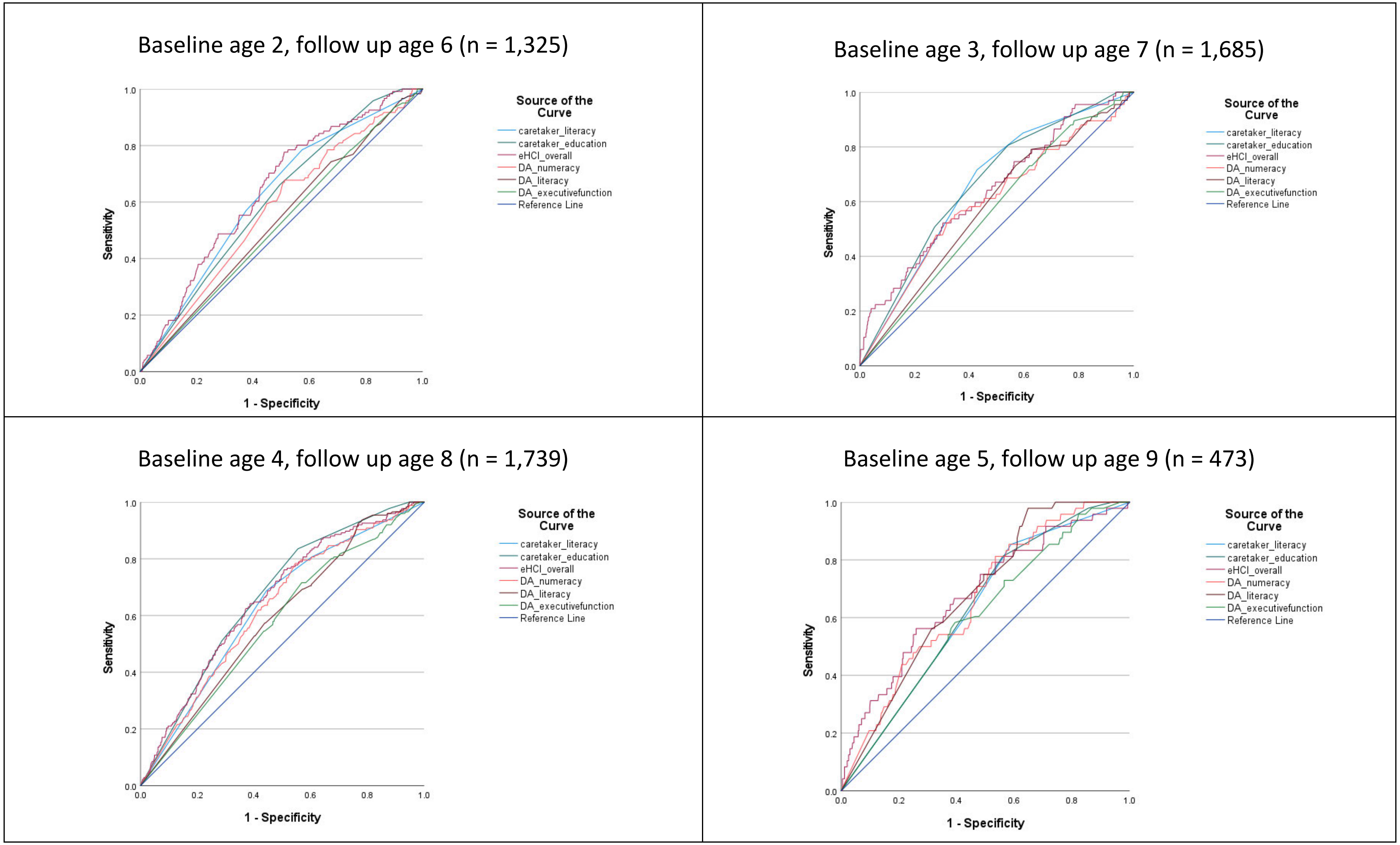
Table 5.1. Sample characteristics at baseline (ages 2-5) (n = 5,222)

		n (%)
Child age		
	2 years	1,325 (25.4)
	3 years	1,685 (32.3)
	4 years	1,739 (33.3)
	5 years	473 (9.1)
Child sex		
	Male	2,686 (51.4)
	Female	2,536 (48.6)
Caregiver education		
	No school	1,492 (28.6)
	Primary school (incomplete)	1,402 (26.8)
	Primary school (completed)	1,617 (31.0)
	Lower secondary school (completed)	411 (7.9)
	Upper secondary school (completed)	118 (2.3)
	Vocational training or tertiary education	179 (3.4)
	Don’t know	3 (0.1)
Caregiver literacy		
	Cannot read at all	2,301 (44.1)
	Can read a little	878 (16.8)
	Can read well	2,036 (39.0)
	Vision problem	7 (0.1)

Figures 5.1 to 5.3 show ROC curves depicting the ability of eHCI overall development and direct assessment literacy, numeracy, and executive function at baseline (ages 2-5) to predict direct assessment scores in the bottom 10th percentile at follow up (ages 6-9). Table 5.2 shows that, across ages, AUCs ranged from 0.52 to 0.68 when predicting literacy, 0.55 to 0.70 when predicting numeracy, and 0.52 to 0.65 when predicting executive function.

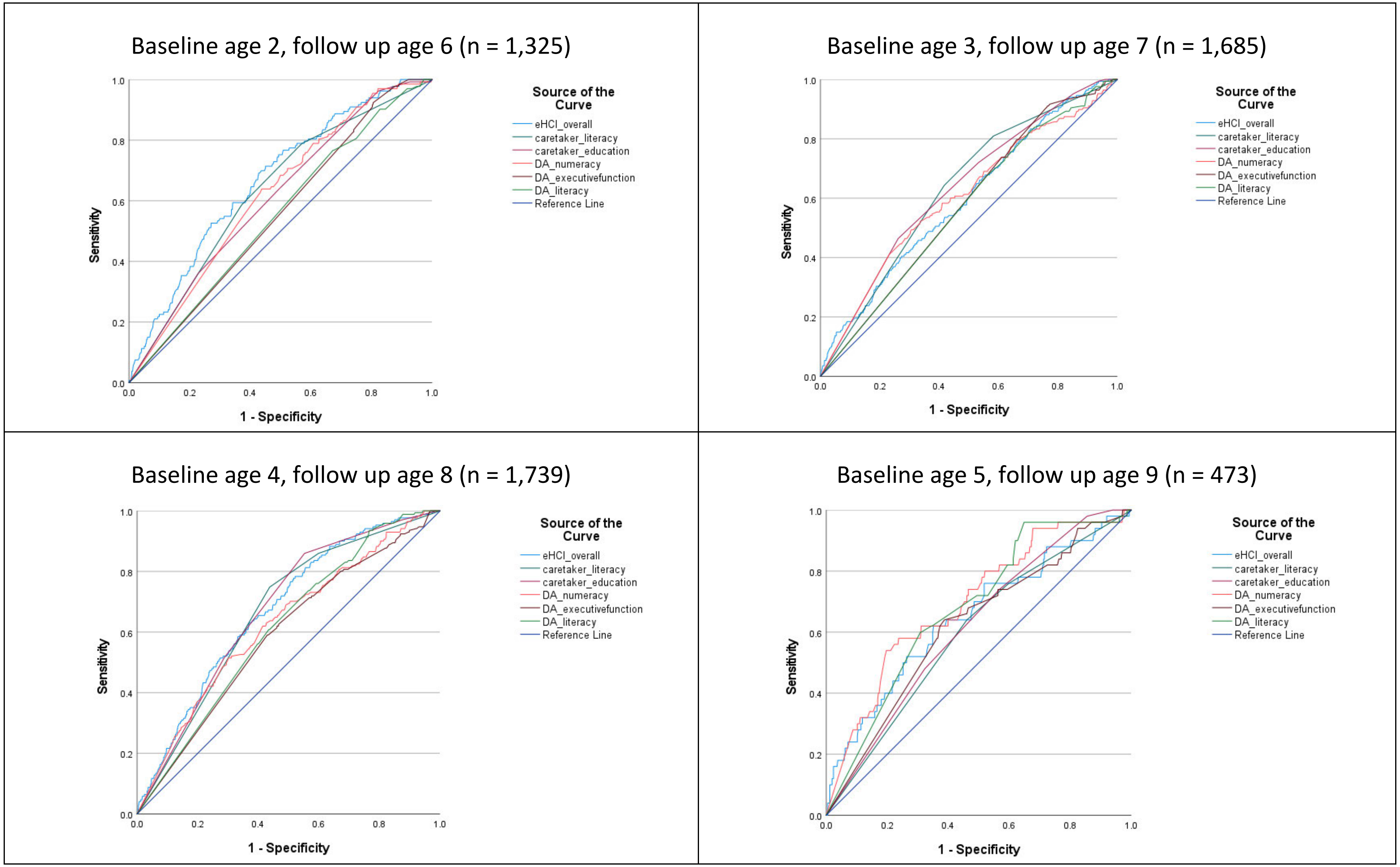
When examining the predictive ability of eHCI overall development scores, there was some variation in results dependent on children's age as well as outcome measures. AUCs were lowest among children aged 3 years (ranging from 0.60 (95% CI 0.55-0.64) when predicting direct assessment numeracy to 0.64 (95% CI 0.57-0.71) when predicting direct assessment literacy), while slightly higher among children aged 2, 4 and 5 years (ranging from 0.64 (95% CI 0.55-0.73) when predicting direct assessment executive function at age 9 to 0.68 (95% CI 0.60-0.75) when predicting direct assessment literacy at age 9). AUCs tended to increase with children's age when predicting direct assessment literacy, though this pattern was not evident on other outcome measures. Further, across ages, AUCs for eHCI overall development scores were typically slightly higher when predicting children's literacy and numeracy, relative to executive function. While C-statistics for eHCI overall development scores ranged from 0.60 to 0.68 and therefore would be considered to have predictive ability in the "low" range according to rules of thumb designated for diagnostic testing[161], interpretation of these results in conjunction with the predictive ability of socioeconomic measures helps to contextualise findings. Indeed, caregiver education and literacy C-statistics had a similar range to that for eHCI overall development, from 0.61 (95% CI 0.56-0.66) when predicting literacy at age 6, to 0.68 (95% CI 0.64-0.71) when predicting numeracy at age 8, with little difference in predictive ability between both measures of socioeconomic position.

Figure 5.1. ROC curves for eHCI, DA, and socioeconomic measures at ages 2-5 predicting literacy scores in the bottom 10% at ages 6-9



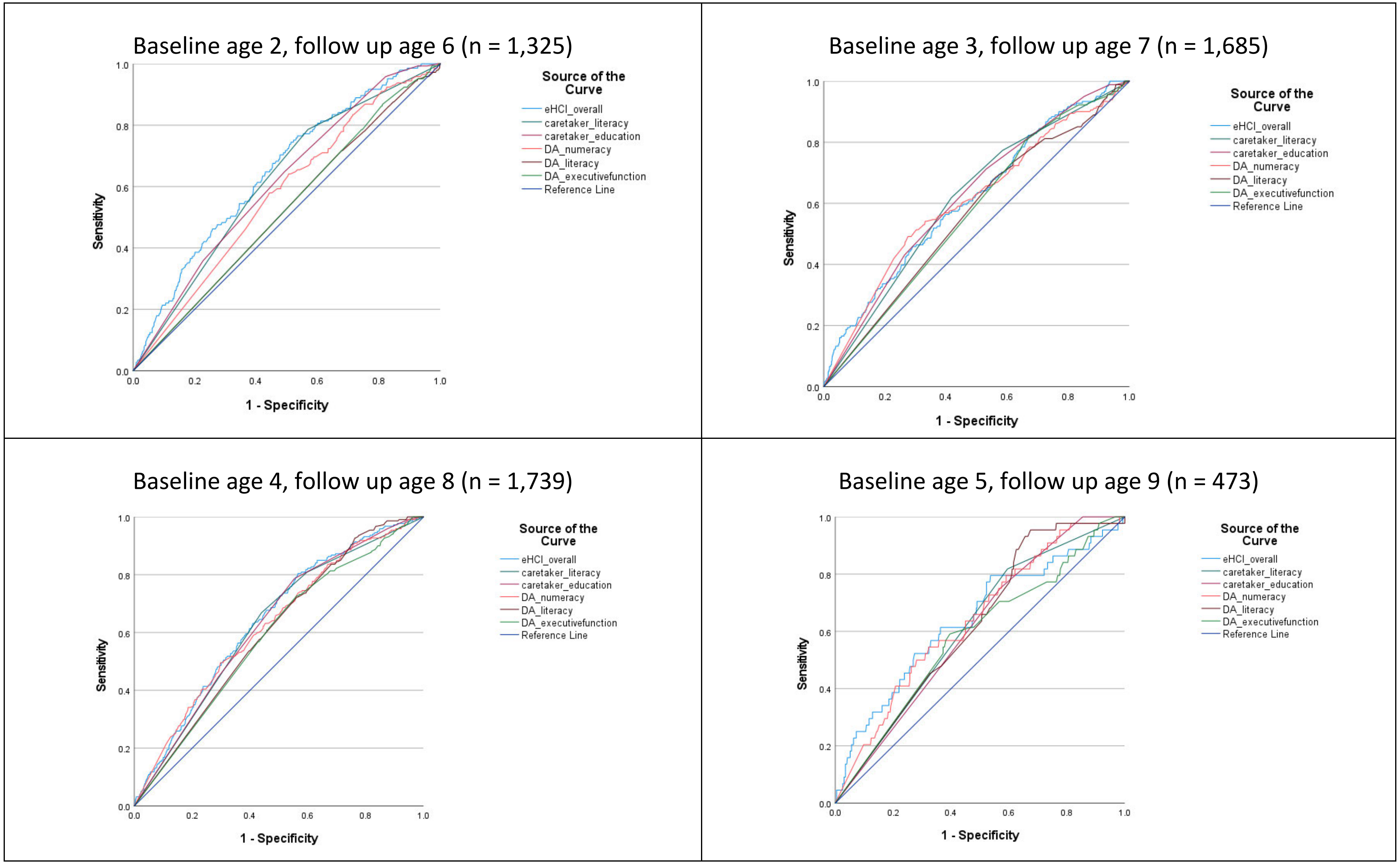
Note. eHCI = early Human Capability Index, DA = direct assessment.

Figure 5.2. ROC curves for eHCI, DA, and socioeconomic measures at ages 2-5 predicting numeracy scores in the bottom 10% at ages 6-9



Note. eHCI = early Human Capability Index, DA = direct assessment.

Figure 5.3. ROC curves for eHCI, DA, and socioeconomic measures at ages 2-5 predicting executive function scores in the bottom 10% at ages 6-9



Note. eHCI = early Human Capability Index, DA = direct assessment.

Table 5.2. AUC (95% CI) for eHCI, DA, and socioeconomic measures at ages 2-5 predicting DA scores in the bottom 10% at ages 6-9

	DA literacy	DA numeracy	DA executive function
Age 2 (n = 1,325)		Age 6	
eHCI overall development	0.64 (0.60-0.67)	0.67 (0.62-0.71)	0.65 (0.60-0.70)
DA literacy	0.53 (0.48-0.59)	0.55 (0.50-0.60)	0.52 (0.47-0.57)
DA numeracy	0.57 (0.52-0.62)	0.62 (0.57-0.67)	0.58 (0.53-0.62)
DA executive function	0.52 (0.47-0.58)	0.55 (0.51-0.60)	0.53 (0.48-0.57)
Caregiver education	0.61 (0.56-0.66)	0.61 (0.58-0.68)	0.61 (0.57-0.64)
Caregiver literacy	0.62 (0.57-0.67)	0.63 (0.62-0.71)	0.62 (0.57-0.67)
Age 3 (n = 1,685)		Age 7	
eHCI overall development	0.64 (0.57-0.71)	0.60 (0.55-0.64)	0.61 (0.57-0.66)
DA literacy	0.58 (0.51-0.65)	0.57 (0.52-0.61)	0.56 (0.52-0.60)
DA numeracy	0.60 (0.53-0.67)	0.60 (0.56-0.65)	0.61 (0.56-0.65)
DA executive function	0.56 (0.50-0.63)	0.58 (0.53-0.62)	0.57 (0.53-0.61)
Caregiver education	0.67 (0.60-0.73)	0.63 (0.59-0.68)	0.62 (0.58-0.66)
Caregiver literacy	0.66 (0.60-0.72)	0.63 (0.59-0.67)	0.61 (0.57-0.66)
Age 4 (n = 1,739)		Age 8	
eHCI overall development	0.66 (0.62-0.70)	0.67 (0.63-0.71)	0.64 (0.61-0.68)
DA literacy	0.59 (0.55-0.63)	0.61 (0.57-0.65)	0.61 (0.57-0.64)
DA numeracy	0.63 (0.59-0.67)	0.63 (0.58-0.67)	0.63 (0.59-0.66)
DA executive function	0.57 (0.53-0.62)	0.59 (0.55-0.63)	0.59 (0.55-0.63)
Caregiver education	0.67 (0.63-0.71)	0.68 (0.64-0.71)	0.63 (0.59-0.66)
Caregiver literacy	0.63 (0.59-0.67)	0.67 (0.63-0.71)	0.63 (0.59-0.67)
Age 5 (n = 473)		Age 9	
eHCI overall development	0.68 (0.60-0.75)	0.65 (0.56-0.73)	0.64 (0.55-0.73)
DA literacy	0.68 (0.61-0.75)	0.68 (0.60-0.75)	0.62 (0.54-0.69)
DA numeracy	0.66 (0.59-0.74)	0.70 (0.62-0.77)	0.64 (0.56-0.72)
DA executive function	0.60 (0.52-0.68)	0.62 (0.54-0.70)	0.58 (0.49-0.67)
Caregiver education	0.63 (0.56-0.70)	0.62 (0.54-0.69)	0.61 (0.53-0.69)
Caregiver literacy	0.63 (0.55-0.70)	0.60 (0.52-0.68)	0.61 (0.55-0.73)

Note. AUC = Area Under the Curve, CI = confidence interval, eHCI = early Human Capability index, DA = Direct Assessment.

Of all child development measures, eHCI overall development typically had the highest (or equal highest) AUC when predicting all outcomes four years later. For example, eHCI overall development at age 2 had the highest AUC when predicting direct assessment numeracy at age 6 (AUC 0.67, 95% CI 0.62-0.71), followed by direct assessment numeracy at age 2 (AUC 0.62, 95% CI 0.57-0.67). One exception to this finding was among children aged 5 at baseline; direct assessment literacy and numeracy yielded the highest AUCs when predicting numeracy scores (direct assessment numeracy AUC 0.70 (95% CI 0.62-0.77), direct assessment literacy AUC 0.68 (95% CI 0.60-0.75), eHCI overall development AUC 0.65 (95% CI 0.56-0.73)).

Sensitivity analyses using direct assessment scores in the bottom 20th percentile (Supplementary Table 5.4 and Supplementary Figures 5.1-5.3; Appendix F) produced results similar to the main analysis, with slightly lower C-statistics observed across all measures. Results from additional ROC curve analyses including individual eHCI domains are presented in Supplementary Table 5.5 and Supplementary Figures 5.4-5.6 (Appendix F). Developmental domains that yielded the highest AUCs varied across children's age as well as outcome measures. For instance, at age 2, the eHCI cultural knowledge domain had the largest AUCs across outcomes, while at age 4, eHCI reading, writing, and numeracy domains had the highest AUCs. However, C-statistics of individual eHCI domains did not exceed that of the eHCI overall development score when predicting direct assessment outcomes.

5.3.5 Discussion

Scores on the eHCI at ages 2-5 demonstrated some ability to discriminate between children with and without poor cognitive development at ages 6-9 in Lao PDR. Results represent the first predictive validity evidence for a population-level measure of child development in LMICs.

While we used techniques (ROC curves) that are typically applied to screening tools in healthcare, the eHCI was not designed to provide an individual diagnostic test but rather an indication of where a child sits on a developmental continuum. Information collected via the eHCI was intended to be used for purposes including population monitoring, evaluation of early years policies and programs, as well as longitudinal studies seeking to predict children's future capabilities.[171] The ROC curve analyses in this study were designed to show that the eHCI contained developmental 'signal' for later cognitive outcomes. Findings supported this aim, with results indicating the ability of eHCI overall development scores at ages 2-5 to discriminate between children with and without poor cognitive development, approximately four years later. Predictive ability was considered low according to diagnostic accuracy rules of thumb.[161] However, results indicated eHCI overall development scores signalled risk for poor future cognitive development with similar ability to that of caregiver literacy and education; both well-established predictors of children's developmental outcomes.[197] Further, caregivers in this study had low literacy and education, with the majority of respondents having not completed primary school. Based on existing evidence for the tool (e.g., better fit of data to the theoretical structure of the eHCI amongst more educated caregivers, compared to those less educated)[183, 184], we anticipate stronger predictive performance of the eHCI when completed by children's teachers or more educated caregivers.

There was some variation in findings dependent on children's age as well as outcome measures. Results indicated eHCI overall development scores had slightly lower predictive ability among 3-year-olds relative to children of other ages, particularly when predicting poor scores on direct assessment numeracy and executive function domains. It may be that growth in development (i.e., trajectories) between ages 3 (baseline) and 7 (follow up) was particularly varied among children in the study sample, which could be unique to this age cohort because of the interventions implemented as part of the ECE Project (e.g., introduction of playgroups for children aged 3-4 years) that underlay the data used in this study.[127] Results also suggested eHCI overall development scores were better able to predict direct assessment literacy and numeracy, relative to executive function which may be due to differences in the range of skills captured by these domains. Executive function included three sub-domains (two measuring working memory) and 16 items, while literacy and numeracy domains included 6 and 9 sub-domains and 37 and 39 items, respectively. It is possible that with fewer items and sub-domains, the executive function domain captured less developmental variation among children and therefore was more difficult for measures to predict low scores.

These results highlight that findings from our study are limited by the measures used. That is, there are no gold standard measures of child development and direct assessment measures included here had not been validated for use in Lao PDR prior to this study. Future research could explore the ability of the eHCI to predict children's academic outcomes and school completion; both of which are important indicators for future labour market participation.[198]

Of all child development measures included in this study, eHCI overall development scores demonstrated the greatest ability to predict poor cognitive development at follow up. Direct assessments of children's development are time and resource intensive and usually require existing qualifications and/or significant training of enumerators. The eHCI is an adult-report (e.g., caregiver, teacher) measure of child development that takes less than 10 minutes to complete and requires minimal training of enumerators. Our findings suggest the eHCI is not inferior to direct assessment of children's capabilities and offers a valid and reliable measurement choice, particularly in low resource settings. Further, the eHCI is a multidimensional measure capturing children's development across nine domains (or eight in the case of the Lao PDR adaptation).

Many direct assessment measures are limited to literacy and numeracy as these are the skills that can be most easily assessed objectively through activities/tasks with a child (versus measuring children's social and emotional skills, for instance). Aspects of development beyond literacy and numeracy are important for children's learning and development trajectories, and this is reflected in our results.

Implications

Little is known about how a child's early development influences later outcomes in LMICs, in part due to limitations in valid, feasible measurement tools and comprehensive monitoring systems. Results provide evidence that poor scores on the eHCI in early childhood predict poor cognitive outcomes once children reach primary school age in Lao PDR. Together with previous research[105, 183, 184], our findings indicate that costly and time-intensive direct assessment of child development may not be necessary to identify where (e.g., particular developmental domains or geographies) and among whom (e.g., particular sub-populations or children with poor eHCI scores) investment and support are required to promote children's outcomes. Rather, the eHCI, a pragmatic tool that is freely available for anyone to use, could be used to guide early years interventions and policy, as well as evaluate their impacts.

An important strength of this study is that it sampled from a well-defined population enumeration that contained children from households, rather than selected populations of children attending early childhood education, as is common with many studies exploring early development in LMICs. Further, we examined predictive ability of the eHCI after a follow up period of four years; considerably longer than that of existing research exploring predictive validity of multidimensional assessment of child development in LMICs.[158] While results build on existing psychometric evidence for the eHCI, further investigation of the predictive ability of the tool across diverse settings, with varied respondents, among children of different ages, using different outcome measures, and over varied periods of time is required for results to be generalised beyond the setting and study sample included in the current study.

Conclusion

Current global efforts to monitor child development in LMIC are limited by a lack of evidence demonstrating that scores on early development measures have predictive validity for meaningful social and economic outcomes as children grow up. Governments, researchers, and communities must have access to measurement tools that are not only feasible for use in low resource settings and generate locally relevant information but are also predictive of future capabilities. Our results support that the eHCI may be able to fulfil this need and is at least as predictive as direct assessments. Data collected using the eHCI could be used to inform the investments required to promote children's early development to improve later cognitive outcomes, though further research is required to generalise results beyond this study.

Chapter 6: Are scores on the eHCI sensitive to and therefore different based on the quality of the early childhood education children attend?

6.1 Chapter outline

The first three studies in this thesis established that the eHCI was psychometrically robust across diverse contexts and predicted children's later outcomes. As described in Chapter 1, a challenge specific to population measures is having the sensitivity to adequately detect variation in children's development so that results can meaningfully inform supports to promote children's outcomes. Such measures should cover a range of domains that capture holistic development, as well as levels of children's ability, from low to high, as is appropriate to the target age range. This seeks to ensure that scores have the sensitivity to detect both variation in development between children, as well as shifts in patterns of capabilities among cohorts of children over time, which is an essential aspect of population monitoring for informing policy and targeting supports. The eHCI was designed not only to be feasible for use at scale, but to be sensitive to variation in children's development with the aim to facilitate evaluation of programs and interventions. This represents a key measurement ideal, outlined in Chapter 1, and is the focus of the final study presented in this thesis.

This chapter includes the fourth paper contributing to this thesis. The manuscript was submitted to *Early Childhood Research Quarterly* in May 2022 and is currently under review.

This study used cross-sectional data from the ECE Project in Lao PDR to investigate the sensitivity of the eHCI to detect variation in children's development, based on differences in the quality of early childhood education (ECE) children attended. It was hypothesised that, on average, children attending higher quality ECE would have better development, after adjusting for confounding variables. The study also compared sensitivity to variation in development between the eHCI and the same direct assessment of children's cognitive development used in Chapter 5. Existing evidence demonstrates mixed findings regarding the relationship between ECE quality and children's development, particularly in LMICs, with the selection of child development measures a potential factor contributing to differences in results. In this way, beyond strengthening evidence for the utility of the eHCI, this study also contributes to advancing evidence around ECE quality measurement approaches in LMICs.

Results from this study showed that ECE quality was more strongly related to children's development as measured by the eHCI, compared to direct assessments, after adjustment for a range of child, household, and village level confounding variables. This may be partly attributable to the fact that the eHCI captures holistic development, while the direct assessments captured cognitive aspects of children's development only. Dosage of ECE attendance (e.g., frequency, intensity, duration) will have an important influence on the relationship between service quality and children's development. While incorporating this information was not possible in this study, future research should seek to investigate how this might influence results observed.

Overall, this study suggests that the eHCI had adequate sensitivity to capture differences in children's development based on the quality of the ECE they attended in Lao PDR. Together with results focused on the psychometric robustness of the tool presented in the first three studies of this thesis, findings indicate that the eHCI may be an appropriate tool to measure impact of the investment in supports in the early years, such as ECE. In this way, the eHCI appears appropriate for both population monitoring and program evaluation purposes. This, together with the implications of these finding, are discussed in Chapter 7.

6.2 Statement of authorship

Title of Paper	Early childhood education quality and child development in Lao PDR
Publication Status	Under review
Publication Details	The manuscript is currently under review at *Early Childhood Research Quarterly*

Principal Author

Name of Principal Author (Candidate)	Alanna Sincovich		
Contribution to the Paper	Contributed to the acquisition of data. Led study design. Supported data analysis. Led interpretation of results and drafting of the manuscript. Revised and finalised manuscript.		
Overall percentage (%)	75		
Certification	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	14/07/2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that: (i) the candidate's stated contribution to the publication is accurate (as detailed above); (ii) permission is granted for the candidate in include the publication in the thesis; and (iii) the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Benjamin Lam		
Contribution to Paper	Supported study design and analysis plan. Conducted analysis. Supported interpretation of results. Reviewed the manuscript.		
Signature		Date	15/07/2022

Name of Co-Author	Tess Gregory		
Contribution to Paper	Supported interpretation of results. Reviewed the manuscript.		
Signature		Date	07/07/2022

Name of Co-Author	Nathalie Rodilosso		
Contribution to Paper	Supported study design and analysis plan.		
Signature		Date	14/07/2022

Name of Co-Author	Sally Brinkman		
Contribution to Paper	Contributed to the acquisition of data. Supported study design, analysis plan, and interpretation of results. Reviewed the manuscript.		
Signature		Date	13/07/2022

6.3 Early childhood education quality and child development in Lao PDR

6.3.1 Abstract

Globally, as scale up of early childhood education (ECE) continues, monitoring ECE quality is imperative to identify and promote service aspects that drive positive outcomes for children. Monitoring of ECE quality in low- and middle-income countries (LMICs) is scarce, limited by challenges in conceptualisations of quality, lack of measurement tools that reflect local culture and context, and implementation difficulties in low resource settings. This study used data from the World Bank Early Childhood Education Project to conduct a secondary, cross-sectional investigation focused on ECE quality and children's development in Lao PDR. Participants were children aged 2-6 years attending ECE in 172 classrooms across 131 villages (n = 1,168). Linear regressions explored the association between ECE quality, captured using the Measure of Early Learning Environments (MELE), and both adult reported and directly assessed child development. Models were adjusted for confounding variables including child age, sex, ethnicity, stunting, household socioeconomic status, and village remoteness. Although associations between ECE quality and children's development were small, consistent with existing research in LMICs, findings reiterate the need for continued investment not only in expansion of access to ECE for all children, but also for policy to ensure high quality service provision. Further, results highlight the emphasis on quality needs to extend beyond facilities and materials, to adequately equipping schools and teachers with the training and resources required to facilitate learning through what are deemed high quality interactions and responses to children's needs within any given context.

6.3.2 Introduction

The benefits of early childhood education (ECE) for children's development have been well established.[199] Early learning environments that provide children with stimulating, responsive, and developmentally appropriate interactions and activities can help to mitigate adverse developmental consequences of growing up in disadvantage.[9] More recently, positive impacts of sustained, high quality ECE have been extended to long-term education achievement, labour market outcomes, and poverty reduction.[200-203] Together with this body of research, decades of advocacy and government investment have spurred significant increases in ECE attendance, particularly in low- and middle-income countries (LMICs) where access to early education has historically been scarce. Indeed, rates of ECE enrolment in LMICs almost doubled throughout 2000 to 2019, from 33% to 62%.[204] Despite this, increased attendance has not brought about the same growth in children's learning and development than that observed in high-income countries. Poor quality ECE has been argued to be one of the causes of smaller positive effects, or indeed no impact or even adverse effects, for children's learning and development often observed in LMICs.[163, 205] Emphasis on monitoring ECE quality is imperative to identify and promote service aspects that drive positive outcomes for children. As the scaling up of ECE services continues, this study seeks to build on current efforts regarding measurement and monitoring of ECE quality in LMICs.

ECE Quality in the Spotlight

Globally, with rapid increases in children's access to ECE has come a shift in focus beyond provision and attendance to the quality of services and the learning that occurs through participation. Indeed, this emphasis on quality and outcomes is reflected in the United Nations Sustainable Development Agenda, among other global frameworks and indicators. Sustainable Development Goal (SDG) 4 prioritises inclusive and equitable quality education to promote life-long learning opportunities for all, with target 4.2 stating that by 2030, countries must ensure all girls and boys have access to quality early childhood development, care, and pre-primary education so that they are ready for primary education.[61] Through access to quality ECE, target 4.2 seeks to maximise the proportion of children who are developmentally on track which will not only have a positive effect on individual outcomes throughout the life course, but also promote future human capability within nations. Increasingly, research has explored the quality of ECE relative to children's development[206], which requires a detailed understanding of the characteristics of ECE quality and how they can be measured. While the global goals have set out indicators to measure progress towards education for all, what constitutes quality ECE is not defined, and this highlights a key challenge when looking to measure and promote quality of children's early learning.

Aspects of ECE Quality

The quality of ECE environments is typically conceptualised in terms of structural and process characteristics.[207] Structural quality indicators include operational factors such as teacher-child ratios and length of the program day, as well as resources and facilities such as learning materials and drinking water. Structural factors are often readily defined, require no interpretation or inference, and thus are easily observable and recordable. As a result, indicators of structural quality have been widely used by governments to monitor classroom quality.[208] Process factors encompass interactions between participants and materials in the classroom environment, and include teacher-child interactions, the schedule and organisation of class time, and teacher responses to the learning, physical, and psychosocial needs of children. Aspects of process quality are more difficult to define and observe, requiring collection of high-inference data by thoroughly trained enumerators. Thus, implementation is complex, often resulting in heterogeneity of application and inconsistent results regarding quality outcomes.[209] both of which limit monitoring of process quality in low resource settings.

While it is thought that aspects of process quality tend to be the primary driver of the positive influence of ECE on children's development[210, 211], structural and process quality are intertwined in that high quality structural aspects of the ECE environment enable high quality processes to occur.[212] For instance, research in Ghana demonstrated that structural quality was predictive of process quality, which in turn predicted children's academic outcomes and social and emotional development.[213] This is particularly relevant in LMICs as access to materials and facilities is limited, coupled with large class sizes.[214] For the purposes of monitoring ECE quality, measuring structural or process aspects alone is insufficient. Rather, measures that incorporate both structural and process aspects of early learning environments are needed to adequately capture ECE quality.

Understanding ECE Quality in LMICs

The majority of research, including conceptualisation and description of ECE quality as well as investigation of associations between ECE quality and children's development, has been conducted in high-income countries. Globally, the tool most commonly used to capture both structural and process quality is the Early Childhood Environment Rating Scales (ECERS)[215]. The latest revision, ECERS-R, includes seven sub-scales: space and furnishings, personal care routines, language-reasoning, activities, interaction, program structure, and parents and staff.[216] Although the ECERS has been adapted for use in some LMICs[217-219], the tool was developed based on ECE quality as it is conceptualised in high-income countries. As a result, findings, though limited, suggest the tool does not operate in the same way in LMICs (e.g., differing factor structures).[220]

Indeed, whether Western conceptualisations of ECE quality are applicable to children and classrooms in LMICs is not well known.[221] Cultures and societies shape children's development and the environments in which learning occurs.[93] Thus, definitions of high quality, best practice ECE will vary across contexts. Further, how different aspects of classroom quality influence children's outcomes may also vary. For instance, in China, teachers reported valuing strict classroom management strategies to ensure order and quiet to facilitate children's learning.[222] This is not, however, a feature of high quality ECE according to tools developed in Western settings (e.g., ECERS). Thus, in LMICs, poor scores on measures of ECE quality developed for high-income countries may reflect lack of relevance to local context, rather than lack of quality.

The Measure of Early Learning Environments

Monitoring of ECE quality in LMICs is scarce, limited by challenges in conceptualisations of quality, lack of measurement tools that reflect ECE quality according to local culture and context, and difficulties implementing measurement in low resource settings. Recent measurement efforts have worked to overcome these limitations. One such example is the Measure of Early Learning Environments (MELE), developed as part of the Measuring Early Learning Quality and Outcomes (MELQO) initiative.[83] The MELE was designed for use in LMICs, to support feasible and actionable measurement of early education settings to provide locally relevant data, and inform global monitoring.[223] Through classroom observation, the tool intends to measure quality constructs that predict children's development. The MELE exists in several iterations and adaptations that have been implemented across several LMICs including Brazil, China, Colombia, Ethiopia, Indonesia, Peru, South Africa, and Tanzania (see ecdmeasure.org).

Emerging evidence has shown applicability of the MELE across diverse settings as well as acceptable psychometric properties.[84] However, associations between ECE quality as measured by the MELE and children's development, are limited and inconsistent. For example, research using the MELE in Tanzania demonstrated a small association between only one of three aspects of ECE quality (materials/activities; representing structural quality) and children's development measured via direct assessment, after controlling for a series of child, family, teacher, and classroom characteristics.[86] The same study showed no association between scores on the MELE and caregiver or teacher reported measures of children's development. In China, ECE quality on three of four MELE domains (learning activities; classroom arrangement, space and materials; facilities and safety; representing structural quality) had small associations with children's development measured by direct assessment, after adjustment for child, family, and program characteristics.[139] Additionally, this research showed a stronger association between ECE quality and child development in rural versus urban settings.

Provision of high quality ECE through supports targeting structural and process aspects of the classroom environment offers an important mechanism for intervention. It is intended that with improved ECE quality will come improved child learning and development outcomes. However, inconsistent results regarding this relationship in LMICs raise important questions around whether this is the case.[207] Without monitoring of ECE using measures of quality that are associated with children's learning and development, investments in such interventions to boost service quality may fail to deliver expected results. This study explores application of the MELE in Lao People's Democratic Republic (PDR), advancing current understanding of the relationship between ECE quality as measured by the MELE, and children's development.

Early Childhood Education in Lao PDR

As one of the fastest growing economies in East Asia and the Pacific, Lao PDR recently moved to lower middle-income classification and has halved rates of poverty, reduced malnutrition, and improved health and education outcomes over the past two decades.[224] Significant health, education, and economic disparities remain, however, particularly across ethnicities and geographies.[126] Globally, ECE is one of the most cost-effective and equitable interventions.[17] Despite the government's commitment to education through adoption of the Education for All National Plan of Action 2003-2015[225], upon ratification of the United Nations 2030 Sustainable Development Agenda in 2015[61], investment in ECE remained insufficient. Existing programs officially recognised by the Lao PDR Ministry of Education and Sports, primarily kindergarten and pre-primary school, largely catered for 5-year-olds and this was reflected in enrolment rates. In 2013, enrolment in ECE among 3–5-year-olds was 33%, while enrolment among 5-year-olds only was 52%.[226] As a result, children aged 3 and 4 years had particularly limited access to ECE. Further, there existed large disparities in ECE attendance based on a child's ethnicity, household socioeconomic position and geography. For instance, less than 8% of children from the poorest households, from rural areas, or from non-Lao Tai communities (i.e., ethnic minorities), had access to ECE services.[226] Together with additional threats to children's development including poverty and poor nutrition, lack of access to early learning has hindered human development.

Against this background, the World Bank Early Childhood Education Project was launched by the Government of Lao PDR. The project supported expansion of quality ECE with the aim of improving development among children aged 3 to 5 years.[226] Project interventions consisted of three key components designed to increase demand for, coverage/access, and quality of ECE, as well as strengthening monitoring and evaluation in the early childhood sector. Overall, the project sought to contribute to reduction in poverty and inequalities through improving educational outcomes among the most disadvantaged. Project intervention included the introduction of two new forms of ECE, with research design enabling analysis to determine which type was most effective in promoting children's development.[102] Briefly, Community Child development Groups (CCDGs) involved training local caregivers in the provision of community-based playgroups. CCDGs, delivered in purpose-built huts, were intended to provide an informal learning environment for children aged 3 to 4 prior to entering kindergarten or a pre-primary classroom at age 5. Multi-Age Teaching (MAT) provided training and resources to existing, qualified pre-primary teachers in an ECE curriculum. MAT, therefore, provided a learning environment for children aged 3 to 5 years, before entering primary school.

Spanning 2015 to 2020, results from the project's impact evaluation included greatly increased ECE attendance and modest associated gains in child development, with differential impacts based on child age, ethnicity, and ECE type.[128] Further, the original study compared the quality of different types of ECE services implemented as part of the project interventions, with findings indicating little difference in the quality of services despite differences in teacher training and qualifications, as well as physical classroom characteristics.

Current Study

This study used data from the ECE Project to conduct a secondary, cross-sectional investigation focused on ECE quality and children's development in Lao PDR. With rapid increase in ECE attendance as a result of the ECE Project, the need to investigate quality of ECE services through this scale up is paramount. To do so, we used the MELE classroom quality observation tool, together with outcome measures that provided data from multiple informants on children's development. This research sought to describe quality of ECE across Northern Lao PDR and explore the association between ECE quality with caregiver reported and directly assessed child development. Advancing our understanding of the quality of ECE provision in Lao PDR, how quality may be captured in this LMIC setting using a built-for-purpose tool, and how ECE quality relates to children's development, are all necessary steps toward informing interventions and policy aimed at improving ECE quality.

6.3.3 Method

Data Source

This study used data collected for the World Bank Early Childhood Education Project (P145544) (herein referred to as the ECE Project) in Lao PDR.[226] An impact evaluation of the ECE Project was conducted through three clustered randomized control trials across 376 villages in Northern Lao PDR, with three waves of data collection spanning five years from 2015 to 2020.[128] Five northern provinces with high levels of poverty were selected and 14 districts within provinces with a district level education office were selected to take part. All villages within these districts in which a minimum of 20 children aged 3 to 5 years resided were selected, with 20 households in each village randomly sampled and children within households randomly selected. In villages with fewer than 20 households with children in the eligible age range, multiple children from the same household were included in data collection. Villages, households, and children were followed over the course of the project, with an additional cohort of children aged 3 to 5 years (sampled using the same methodology) included in the third wave of data collection to enable cross-sectional as well as longitudinal analyses. While focus was on children aged 3 to 5 years, some children outside of this age range (i.e., 2- and 6-year-olds) were captured in data collection.

Relevant to the current study, measurement of ECE quality was employed in the third wave of data collection in selected villages. Specifically, the 135 villages that participated in the arm of the impact evaluation which compared effectiveness of different ECE programs (study one, see Supplementary Figure 6.1; Appendix G). Investigation of ECE quality was employed in response to large increases in ECE attendance observed at the project's second wave of data collection in

2017, yet little improvement in children's development outcomes.[227] Thus, while the ECE Project had a longitudinal design, the current study used the third data wave as a cross-sectional analysis.

Participants

The study population included children aged 2-6 years for whom data were collected in the third wave (February to March 2020) of the ECE Project (full sample; N = 3,322). The sample was restricted to children currently attending ECE (eligible sample; n = 1,309), enabling investigation of the association between the exposure (ECE quality) and outcome (child development) among the targeted age range of ECE services in Lao PDR. Missing data on the exposure (n = 96) and covariates (n = 45) resulted in a final analysis sample of n = 1,168. Supplementary Table 6.1 presents the characteristics of full, eligible, and analysis samples.

Measures

Early Childhood Education Quality

The Measure of Early Learning Environments (MELE) captures constructs related to the quality of early learning environments.[83] The MELE constitutes a suite of tools including a classroom observation, teacher interview, caregiver interview, and school director interview. The MELE does not intend to serve as an evaluation of teachers or schools. Rather, it was designed to enable tracking of the quality of children's learning environments at a population level and identify areas in which training and supports might be needed to promote early learning outcomes. To date, several adaptations of the MELE have been developed and field-tested, with validation conducted in selected LMICs.[84, 86, 139]

The MELE classroom observation (version 04.01.2018) was adapted for use in Lao PDR for the ECE Project.[140] Adaptation was led by stakeholder consultation as well as field testing of the tool, and revisions ensured the measure was aligned with the culture and context in Lao PDR. A summary of adaptations is provided in Supplementary Table 6.2 (Appendix I). The Lao PDR version of the MELE measures four key aspects (domains) of quality: Learning Activities (8 items; $\alpha = 0.63$); Classroom Interactions and Approaches to Learning (10 items; $\alpha = 0.51$); Classroom Arrangement, Space, and Materials (14 items; $\alpha = 0.56$); and Facilities and Safety (6 items; $\alpha = 0.53$). Table 6.1 describes the characteristics each domain captures, including example items. With the exception of Classroom Arrangement, items in all domains were scored from 1 to 4, with 4 representing high quality. Items in the Classroom Arrangement domain were either binary (yes/no) or multiple choice to indicate presence of materials (no materials present/materials present but children do not use/materials present and children use). Domain scores were calculated by summing scores on all items. Domain scores were transformed to range from 0 to 10 (with 10 indicating high quality) to aid interpretability, and an overall quality score ($\alpha = 0.63$) was derived by taking the average of three domains (Interactions and Approaches, Classroom Arrangement, Facilities and Safety), also ranging from 0 to 10, with higher scores indicating better quality.

The Learning Activities domain was not included in the overall quality score. Weak (and sometimes negative) correlations between the Learning Activities domain and remaining domains as well as overall quality were observed, in contrast to that among other quality domains (Supplementary Table 6.3; Appendix G). Unlike remaining domains, scores on the Learning Activities domain (e.g., whether learning opportunities to support development of maths, literacy, language, or fine motor skills occurred) in the context of Lao PDR were highly dependent on the day and time at which the classroom observation occurred and thus which lesson was being conducted, and therefore quality scores. We provide further commentary on this in the discussion.

Table 6.1. MELE domain descriptions and example items

Domain	Domain description	Example item
Learning Activities	Learning opportunities to develop math, literacy, expressive language, fine and gross motor skills, as well as book reading, music/ movement, and free play	Book reading to support children's listening and speaking skills: 1= Instructor does not read book to children OR reads a book that is not age-appropriate (i.e., text or schoolbooks for older children or adults; religious text for adults; books with no pictures, or books for younger children). 2= Instructor reads to the class without discussion OR without questions about the story. 3= Instructor reads to the class using one of the following strategies: asks children basic or close-ended questions about what happened in the story; encourages children to discuss the story through open-ended questions; talks about vocabulary learned in the book; connects the story to the children's own experiences or context. 4= Instructor reads to the class using two or more of the following strategies: asks children basic or close-ended questions about what happened in the story; encourages children to discuss the story through open-ended questions; talks about vocabulary learned in the book; connects the story to the children's own experiences or context.
Interactions and Approaches	Classroom management, learning environment, instruction methods, supervision, and child engagement	Verbal disciplinary strategies: 1= Instructor does not discipline when there is disruptive or bad behavior 2= Instructor uses negative verbal interactions (yelling, harsh tone, threats, humiliation) with children to control child behavior 3= Instructor redirects children to using more appropriate behavior (for example, "sit down" or "use a quiet voice") BUT is inconsistent with redirection (e.g., only uses with some situations or some children) OR is ineffective with redirection (e.g., does not provide appropriate redirection or does not follow-through). 4= Instructor uses positive techniques for guiding children's behavior consistently (explains reasons for rules, consistently applies rules) AND teacher

		consistently addresses behavior problems OR no behavior problems were observed.
Classroom Arrangement	Equipment, furniture, space, and learning materials present, such as books and writing utensils	All children have a seat and access to a writing surface that are appropriately sized (yes/no)
Facilities and Safety	Facilities available for water and sanitation, as well as presence of any safety issues or hazards	Handwashing practices: 1= Children do not wash hands or only a few children wash hands (but only use water). 2= Handwashing is sporadic (some do and some do not) and procedures are inconsistent. 3= More than half of children wash hands after toileting and most of those children wash with soap. The teacher supports handwashing. 4= All children wash hands with soap after toileting with a few (less than 5) lapses and there is a system or process in place for supporting hand washing (teacher supervises, encourages, it is part of routine, etc.).

Note. MELE = Measure of Early Learning Environment.

Child Development

This study utilised two measures of child development that served as outcomes: a caregiver report and a direct assessment. Prior to the ECE Project, measures of child development had not been adapted and/or validated for use in Lao PDR. Implementation of both tools enabled exploration of the applicability, feasibility, and validity and reliability of adult report and direct assessment measures in this context. Both tools are described below.

Caregiver Report: The early Human Capability Index (eHCI) is a population measure designed to capture holistic development among children aged 2 to 6 years. The eHCI is an adult-reported tool completed in less than 10 minutes by an adult who is familiar with the child (e.g., caregiver, preschool teacher). The tool was designed to be feasible for use in low resource settings and adapted to each context to produce locally relevant results. The eHCI has been demonstrated to be psychometrically robust in diverse contexts and has been adapted to support early education and development projects in several LMICs, including Lao PDR.[105, 183, 184]

The Lao PDR version of the eHCI includes 56 items measuring development across 8 domains: Verbal Communication (4 items; $\alpha = 0.82$), Cultural Knowledge (5 items; $\alpha = 0.79$), Social and Emotional Skills (14 items; $\alpha = 0.78$), Perseverance (4 items; $\alpha = 0.83$), Approaches to Learning (7 items; $\alpha = 0.80$), Numeracy (10 items; $\alpha = 0.77$), Reading (6 items; $\alpha = 0.79$), and Writing (6 items; $\alpha = 0.79$).[184] The eHCI was collected via primary caregiver report through a household survey administered by trained enumerators. Item response options were binary (yes/no), with "yes" responses to positively worded items scored as 1, and negatively worded items reverse scored. Item scores were averaged so that children received a score ranging from 0 to 1 for each domain, with higher scores indicating better development. An overall development score ($\alpha = 0.82$) was calculated by taking the average of 8 domain scores, also ranging from 0 to 1.

Direct Assessment: The Measurement of Development Early Learning (MODEL) was created drawing upon existing measures used in LMICs capturing aspects of child development considered globally relevant.[83] The MODEL, forming the MELQO initiative together with the MELE, consists of a child direct assessment, as well as teacher and caregiver reports of children's development. The ECE Project utilised an early working version of the direct assessment (circa 2014/5), adapted to the local context, prior to validation studies conducted in various LMICs.[84, 87] The key differences between this earlier version and later validated versions are that this version did not include some of the more advanced numeracy and executive function items (e.g., mental transformation), as well as items focused on social and emotional concepts (e.g., perspective taking, empathy, understanding feelings).

Literacy, Numeracy, and Executive Function were measured via 92 items. Literacy and numeracy items were based largely on the Early Grade Reading Assessment (EGRA) and the Early Grade Maths Assessment (EGMA) and adapted to the context in Lao PDR[102], both of which have been adapted for use in a number of countries with well-established reliability and validity.[136] Executive function assessments were based on items from the Wechsler Intelligence Scale for Children (WISC-IV) (e.g., Digit Span sub-test) and the Head-Toes-Knees-Shoulders task.[137, 138] Overall, children were assessed on 6 aspects of Literacy (37 items; $\alpha = 0.74$) (e.g., print familiarity, letter knowledge), 9 aspects of Numeracy (39 items; $\alpha = 0.69$) (e.g., number identification, spatial vocabulary), and 3 aspects of Executive Function (16 items; $\alpha = 0.78$) (e.g., forward and backward digit span). Correct item responses were scored as 1, and scores were averaged to provide a total score for literacy, numeracy, and executive function ranging from 0 to 1, with higher scores indicating better development. The average of these scores constitutes an overall development score ($\alpha = .81$), also ranging from 0 to 1.

Sociodemographic Characteristics

Various village, household and child level characteristics were included in analysis as potential confounders of the association between ECE quality and child development (Supplementary Figure 6.2; Appendix G). Village level information included an indicator of remoteness based on the item “Can cars access the village during rainy seasons?” (yes, no). Household socioeconomic status (SES) was derived using Principal Components Analysis and included the following items: number of household members, assets (e.g., television, motorcycle, chicken, cow), house construction materials, caregiver literacy, access to electricity, receipt of financial aid, and purchasing ability (e.g., school supplies, food.). A factor score was created using the first component for each household which was then divided into quintiles to create five levels of SES.[165] Caregiver-reported child level information included age (months), sex (male, female), ethnicity (Lao Tai, Khmu, Hmong, Other), and health status (very healthy, normal, unhealthy/often sick, or don’t know based on the item “What do you think is the current health status of the child?”). Further, children’s height and weight (measured by trained enumerators) were converted into World Health Organization Child Growth Standards height-for-age z-scores, and stunting was defined as a height-for-age z-score < -2.[134]

ECE type (CCDG versus MAT) did not meet assumptions of a confounding variable and therefore was not included as such in analyses. That is, ECE type was not believed to influence both ECE quality and children’s development, as depicted in the Directed Acyclic Graph in Supplementary Figure 6.2, based on findings from the original impact evaluation study.[128]

Procedure

Data collection was conducted by Indochina Research Limited. Fieldwork staff undertook training focused on data collection objectives and logistics, as well as implementation of measurement tools. Training, developed and supervised by authors (blinded for review), was conducted over a 10-day period, after which enumerators were assessed and those most competent were selected to conduct data collection. Together with the Ministry of Education and Sports, Indochina Research Limited coordinated a logistical plan for data collection in each village. Field work teams (consisting of a team leader, assistant, and four enumerators) visited villages to conduct data collection throughout February and March 2020. The MELE classroom observation was completed by a pair of enumerators over a two-hour period, typically beginning when class commenced for the day. Teachers were made aware of the observation and were requested to conduct teaching as they would typically do so. The eHCI was conducted with children's caregivers as part of a broader household survey led by an enumerator. The MODEL was conducted with children in a one-on-one format, framed as a series of activities, either while the child was at home or at ECE. All data were collected using electronic tablets, with various quality control measures employed throughout including random observation of assessments by team leaders and random checking of completed instruments. In March 2020, toward the end of the data collection period, school closures were announced in Lao PDR as a result of the COVID-19 pandemic. As a result, MELE observations were made in 131 of the targeted 135 villages.

As ECE quality was collected at the village level, MELE scores were allocated to each child based on their village of residence. Among children living in villages in which more than one ECE service was available (i.e., CCDG intended for children aged 3-4 years, and pre-primary for children aged 5 years and above), MELE scores were allocated to individual children based on the ECE service relevant to their age.

Statistical Analysis

Descriptive statistics for exposure (MELE) and outcome (eHCI, MODEL) variables were assessed to investigate distribution of scales. Visual inspection to assess model fit was conducted via kernel density curves and transformations were conducted for non-linearity where required. Correlations (Spearman's rho) first explored associations among study variables. A series of linear regressions were then conducted to explore the association between ECE quality (MELE overall quality and individual domain scores, excluding the learning activities domain) and both adult reported and directly assessed child development (eHCI and MODEL overall development and individual domain scores). Linear models were adjusted for confounding variables including child age, sex, ethnicity, stunting, household SES, and village remoteness. Standard errors were clustered on village as children living in the same village were likely to share similarities. Post-model checks for homoscedasticity, non-linearity, and influential observations were conducted via post estimation residual-versus-fitted plots, augmented partial residual plots, and added-variable plots. Model coefficients (unstandardised beta coefficients) were reported with 95% confidence intervals. Effect estimates with confidence intervals that did not overlap zero were interpreted as meaningful associations. Analyses were conducted using Stata/SE 17.[166]

6.3.4 Results

Children in the analysis sample were from 172 classrooms in 131 villages in Northern Lao PDR. The mean number of children per class was 12.26 (SD = 5.54). Table 6.2 shows that the majority of children were in the 3-to-5-year age range, with 39.0% of children aged 4 years. The study sample included slightly more females than males (50.8% and 49.2%, respectively). The largest ethnic group in Northern Lao PDR, Khmu, was the most common ethnicity (40.7%) among the study sample, followed by Lao Tai (22.2%) and Hmong (18.2%). In terms of children's health, 42.1% of children were stunted, though the majority of caregivers reported their children's overall health was either very healthy (57.1%) or normal (36.4%). The study sample was socioeconomically skewed, relative to the full and eligible samples, including fewer children from the most socioeconomically disadvantaged households as a result of restricting the sample to children attending ECE (Supplementary Table 6.1; Appendix G). For instance, 21.0% of children in the analysis sample were living in the least disadvantaged households compared with 16.0% of children in the full sample. Almost half (44.2%) of children lived in remote villages (i.e., not accessible by car in the wet season).

Table 6.2. Participant characteristics (n = 1,168)

Characteristic		n	%
Age	2 years	60	5.1
	3 years	285	24.4
	4 years	455	39.0
	5 years	319	27.3
	6 years	49	4.2
Sex	Female	595	50.9
	Male	573	49.1
Ethnicity	Lao Tai	259	22.2
	Khmu	475	40.7
	Hmong	212	18.2
	Other	222	19.0
Socioeconomic status	(Most disadvantaged) 1	172	14.7
	2	212	18.2
	3	260	22.3
	4	279	23.9
	(Least disadvantaged) 5	245	21.0
Stunted	Yes	500	42.8
	No	668	57.2
Village accessibility	Yes	652	55.8
	No	516	44.2
Child health status	Very Healthy	667	57.1
	Normal	425	36.4
	Unhealthy/Often Sick	73	6.3
	Don't know	3	0.3

Note. Village accessibility based on item "Can cars access the village during rainy seasons?”. Child health status based on item “What do you think is the current health status of the child?”

Mean scores on MELE quality domains indicate classrooms scored highest on Interactions and Approaches (M = 7.11, SD = 0.96) and lowest on Learning Activities (M = 4.47, SD = 0.96) (Table 6.3). Standard deviations indicate that the Classroom Arrangement and Facilities and Safety domains had the greatest variability. The mean Overall Quality score was 5.92 (SD = 0.78), with scores ranging from 3.48 to 7.31 across ECE classrooms (Figure 6.1). Table 6.3

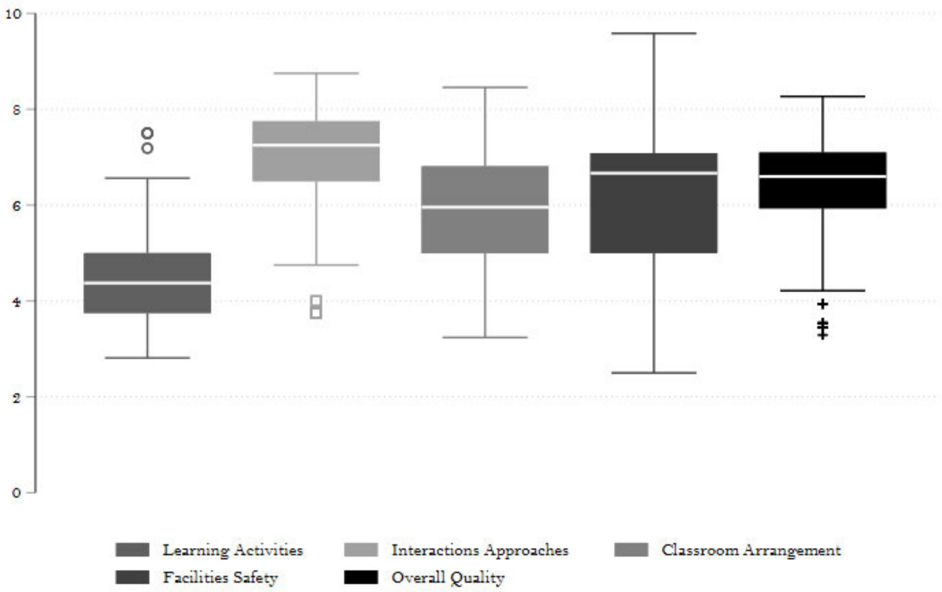
also presents descriptive statistics of child development measures. Children scored, on average, lowest on literacy domains (M = 0.25, SD = 0.29 eHCI Reading; M = 0.31, SD = 0.27 eHCI Writing; M = 0.11, SD = 0.15 MODEL Literacy) and higher on numeracy (M = 0.54, SD = 0.28 eHCI; M = 0.33, SD = 0.20 MODEL). Domain scores for the eHCI and MODEL had similar variance and most were normally distributed with skewness between ± 1.5 [228], though some eHCI domains (e.g., Verbal Communication) showed evidence of ceiling effects (i.e. items too easy) and some MODEL domains (e.g., Literacy) showed evidence of floor effects (i.e., items too hard).

Table 6.3. Descriptive statistics of early childhood education quality (MELE) and child development (eHCI, MODEL) measures (n = 1,168)

Measure	Domain	Mean	SD	Skewness	Kurtosis
MELE	Learning Activities	4.47	0.96	0.58	2.68
	Interactions and Approaches	7.11	0.95	-0.62	3.98
	Classroom Arrangement	5.89	1.22	-0.19	2.30
	Facilities and Safety	6.20	1.62	-0.36	2.36
	Overall Quality	5.92	0.78	-0.77	3.53
eHCI	Verbal Communication	0.85	0.27	-1.87	5.53
	Cultural Knowledge	0.74	0.23	-1.01	4.03
	Social and Emotional	0.56	0.24	-0.21	2.05
	Perseverance	0.45	0.26	0.05	2.70
	Approaches to Learning	0.72	0.26	-0.92	3.15
	Numeracy	0.54	0.28	-0.10	2.07
	Reading	0.25	0.29	1.10	3.22
	Writing	0.31	0.27	0.96	3.27
	Overall Development	0.55	0.17	-0.02	2.67
MODEL	Numeracy	0.33	0.20	0.63	3.14
	Literacy	0.11	0.15	2.01	7.32
	Executive Function	0.22	0.20	0.52	2.42
	Overall Development	0.22	0.15	0.94	3.64

Note. SD = standard deviation, MELE = Measure of Early Learning Environment, eHCI = early Human Capability Index, MODEL = Measurement of Development and Early Learning. MELE data were missing for 96 children. MELE domain scores can range from 0 to 10, scores on the eHCI and MODEL can range from 0 to 1. Higher scores represent better ECE quality and child development.

Figure 6.1. Box plots of MELE domain and overall scores (n = 1,168)



Correlations among study variables (Supplementary Table 6.3; Appendix G) indicate small negative associations between the Learning Activities domain and children's development, and small to moderate positive associations between Interactions and Approaches, Classroom Arrangement, and Facilities and Safety domains and children's development. Results from linear models estimating the association between ECE quality (MELE) and child development (eHCI, MODEL) are shown in Table 6.4. Assessment of model fit indicated models performed well, however transformations were tested on MODEL Literacy and eHCI Verbal Communication domains, based on post-model checks. No improvement was found with transformations of both outcomes thus transformations were not retained. Model coefficients (unstandardised beta coefficients) report mean score increase in outcome per point increase in the exposure. For example, Overall Quality had an effect estimate of 0.04 for eHCI Overall Development (95% CI 0.02, 0.05), which suggests an increase of 0.04 in eHCI Overall Development per point increase in over Overall Quality. Effect estimates (for which confidence intervals did not include zero and thus were considered meaningful) ranged from 0.02 to 0.05

for eHCI domains and Overall Development, and 0.01 to 0.02 for MODEL domains and Overall Development, indicating stronger associations between ECE quality and adult-reported child development, relative to direct assessment.

Overall Quality typically demonstrated the strongest association with child development, followed by the Interactions and Approaches domain. For instance, Overall Quality had an effect estimate of 0.05 (95% CI 0.03, 0.07) for eHCI Social and Emotional development, while Interactions and Approaches had an effect estimate of 0.04 (95% CI 0.02, 0.06). Across child development measures, ECE quality had the largest association with eHCI Numeracy, followed by eHCI Verbal Communication and eHCI Social and Emotional development. Associations between quality scores and MODEL domains were smaller and less variable (i.e., Literacy, Executive Function, and Overall Development each demonstrated the same effect estimates). Although effect estimates observed were small, results indicate all child development scores demonstrated a meaningful association with all or some measures of ECE quality, with the exception of eHCI Perseverance and MODEL Numeracy.

6.3.5 Discussion

This study builds on the limited, existing research that has investigated measurement of ECE quality and its association with children's development in LMICs. Findings demonstrated associations between ECE quality, as measured by the MELE classroom observation, and children's development in Lao PDR. Results help to inform both local and global efforts focused on ensuring ECE provision has an emphasis on quality.

Table 6.4. Adjusted linear regression results for the association between early childhood education quality (MELE) and child development (eHCI, MODEL) measures (n = 1,168)

	Effect	MELE											
		Interactions Approaches			Classroom Arrangement			Facilities Safety			Overall Quality		
		Estimate	95% CI		Estimate	95% CI		Estimate	95% CI		Estimate	95% CI	
			LL	UL		LL	UL		LL	UL		LL	UL
eHCI	Verbal Communication	.04	.02	.07	.03	.01	.05	.03	.01	.04	.05	.03	.08
	Cultural Knowledge	.03	.01	.05	.02	.00	.04	.02	.01	.03	.04	.02	.06
	Social and Emotional	.04	.02	.06	.03	.01	.05	.02	.01	.04	.05	.03	.07
	Perseverance	-.01	-.03	.01	-.00	-.02	.02	-.01	-.02	.00	-.01	-.03	.01
	Approaches to Learning	.04	.01	.06	.02	-.00	.04	.03	.01	.04	.04	.02	.06
	Numeracy	.05	.02	.07	.04	.02	.05	.02	.01	.04	.05	.03	.08
	Reading	.03	.01	.06	.01	-.01	.03	.01	-.00	.03	.03	.01	05
	Writing	.03	.01	.05	.03	.01	.04	.01	-.00	.02	.04	.02	.05
	Overall Development	.03	.02	.05	.02	.01	.04	.02	.01	.03	.04	.02	.05
MODEL	Numeracy	.01	-.01	.02	.00	-.01	.01	.01	-.00	.01	.01	-.01	.02
	Literacy	.02	.01	.02	.00	-.01	.01	.01	.00	.02	.02	.01	.02
	Executive Function	.02	.01	.03	.01	-.00	.02	.01	.01	.02	.02	.01	.04
	Overall Development	.02	.01	.02	.00	-.01	.01	.01	.00	.02	.02	.01	.02

Note. Adjusted for age, sex, ethnicity, socioeconomic status, village accessibility, and stunting. eHCI and MODEL scores range from 0-1, MELE scores range from 0-10. MELE = Measure of Early Learning Environment, CI = confidence interval, LL = lower limit, UL = upper limit, eHCI = early Human Capability Index, MODEL = Measurement of Development and Early Learning.

ECE quality and child development in Lao PDR

Findings from the current study indicated that ECE classrooms across Northern Lao PDR scored highest in terms of Interactions and Approaches (i.e., process quality), with greatest variation among scores measuring structural aspects of quality (i.e., Classroom Arrangement and Facilities and Safety domains). Results suggest inequality in the materials and facilities in ECE classrooms, which is likely to reflect both differences in levels of poverty across villages as well as differences in resourcing across types of ECE (e.g., CCDGs are delivered in purpose-built huts, while MAT is delivered in school buildings). Quality scores may also reflect measurement challenges in that structural factors are easily observable and recordable, while process factors are more difficult to observe and require interpretation.

The relationship between ECE quality and child development varied across domains of quality. The MELE summary measure, Overall Quality, which incorporated aspects of both structural and process quality, demonstrated the strongest association with child development. Following this was the Interactions and Approaches domain, which captured aspects of process quality including how the teacher instructs and manages the classroom, and how children engage in the learning environment. These findings are aligned with existing evidence in that, while aspects of process quality are considered the main driver of the benefits of ECE for child development[210, 211], quality processes cannot occur without structural quality in place.[212]

Findings regarding associations between ECE quality and child development also varied across domains of development. Broadly, ECE quality was most closely related to Numeracy, followed by Verbal Communication and Social and Emotional development (caregiver report indicators, measured by the eHCI). However, differences in the size of associations across child development domains were small. Little existing research has explored these relationships across specific aspects/domains of development in LMICs, with studies typically using a summary or global indicator of children's development. An exception is a recent investigation of ECE quality (MELE) and child development in China.[139] This study used the East Asia Pacific Early Child Development Scales (EAP-ECDS), a psychometrically robust regional direct assessment of child development. While the study's main results demonstrated a positive association between ECE quality and overall child development, supplementary analyses showed that, of the seven domains of the EAP-ECDS, only scores on the Social Emotional domain demonstrated significant associations with ECE quality. Despite differences across studies (e.g., child development measures, country settings), this combination of findings suggests structural and process ECE quality are particularly important for children's social and emotional development.

The MELE as a Measure of ECE Quality

This study builds on existing evidence regarding application of the MELE in diverse contexts. The MELE was designed to be adapted for use in each new context[84]; for instance, to align with early learning or education quality standards in a particular country, or to capture change in ECE quality or child development. As a result, the tool varies across applications, for example, quality domains differ across adaptations ranging from three to seven domains.[221] This flexibility is crucial in ensuring local relevance and ability to capture quality as it is defined in any given context. Though, it creates challenges for comparing results across studies and establishing global measures.

The current study demonstrated associations between both structural and process aspects of ECE quality and children's development. Previous research using the MELE has not indicated this relationship with process quality, but structural quality only. In Tanzania, scores on the MODEL direct assessment demonstrated a small association with the Materials/Activities domains of the MELE only.[86] In China, scores on the EAP-ECDS direct assessment demonstrated small associations with the three domains of the MELE that capture structural quality.[139] These findings contrast with the understanding that aspects of process quality drive the benefits of ECE for children's development. Authors of these studies suggested MELE measurement of teacher-child interactions may lack the detail, depth, or precision required to capture or be sensitive to variations in these processes important for children's development. Further, the MELE captures a one-off snapshot of the classroom and in doing so, fails to measure the stability of processes over time. In addition, evidence from other research suggests a threshold effect, where teacher-child interactions must reach a certain level of quality before being beneficial for children's learning.[229, 230]

Previous research studies exploring the relationship between the MELE and children's development have demonstrated varied results. Findings from the current study suggest this may be influenced by measures of child development employed. Specifically, ECE quality was more strongly related to children's development measured by the eHCI versus the MODEL. This may be a product of the properties of the measures themselves. For example, while the MODEL is a direct assessment and therefore captures a snapshot of children's development at one point in time, the eHCI is an adult report tool. Thus, the eHCI captures the respondent's report of a child's abilities over an extended period. Further, findings might be influenced by the domains of development measured. That is, the eHCI captured holistic development, including aspects of children's cognitive (e.g., literacy, numeracy) and non-cognitive (e.g., social and emotional) development. In contrast, in this study the MODEL captured aspects of cognitive development only. In Tanzania, researchers found no association between scores on the MELE and the MODEL caregiver and teacher report of children's development (focused on Social and Emotional development), but did detect an association between ECE quality and the MODEL direct assessment (Literacy, Numeracy, Executive Function).[86] In China, research reported an association between ECE quality and the EAP-ECDS direct assessment (Overall Development).[139] Variation in findings highlights the importance of using different measures of children's development when exploring ECE quality. While all measures in studies described were adapted to the local contexts, it appears holistic measures of child development have been most sensitive to variation in ECE quality, as measured by the MELE.

Challenges in the Application of the MELE in Lao PDR

Use of the MELE in Lao PDR has highlighted challenges specific to the Learning Activities domain which may be relevant in other settings. To receive the top score on the Learning Activities domain, opportunities to support children's development of maths, literacy, expressive language, and fine and gross motor skills, as well as book reading, free play, and music or movement activities must occur within the observation period. Uncertainty around feasibility of this occurring (i.e., all types of learning observed in a two-hour window) emerged through implementation. Further, unlike remaining MELE domains, scores on the Learning Activities domain were highly dependent on the day and time at which the classroom observation occurred, and thus the lesson being conducted. This is because the ECE curriculum in Lao PDR is nationally consistent (i.e., all forms and classes of ECE carry out the same lesson plan). Thus, unless all observations occurred on the same day and at the same time, this would result in bias among scores on the Learning Activities domain.

More recently established classroom quality measures, also designed for use in LMICs, have addressed this challenge. Teach ECE, developed by the World Bank based on the existing framework of Teach Primary[231, 232], captures time spent on learning activities and the extent to which children are on task.[233] In this way, quality scores are reliant on time spent engaged in learning rather than the type of activity covered in the observation period. This is measured in addition to quality of teaching practices (centred on classroom culture, guided learning, and socioemotional skills) as well as a checklist of aspects of structural quality. Measurement of time on learning and children's level of engagement is based on a growing body of evidence indicating the importance of teachers' ability to maximise learning opportunities in the classroom, including incorporating learning into routines and transitions throughout the day.[234]

Psychometric evidence of Teach ECE is yet to be established across diverse contexts. Moving forward, researchers seeking to measure ECE quality in LMICs could look to incorporating both tools (MELE, Teach ECE) to investigate how these differences play out in practice.

Limitations

Use of cross-sectional data brings about some limitations in our findings. Without longitudinal data and baseline measurement of children's development, growth in development as a result of differences in ECE quality (i.e., a causal relationship) cannot be determined. This is particularly important when investigating the relationship between ECE quality and children's development, considering previous research has shown quality classrooms can be differentially beneficial dependent on a child's existing ability (e.g., children with advanced development can better take advantage of the learning environment).[235, 236] Also, ECE quality, particularly aspects of process quality, are dynamic and thus measurement at one point in time does not capture this nuance. While findings from the current study would suggest increased ECE quality in Lao PDR would result in improved child development, future longitudinal research with measurement of ECE quality and child development at multiple time points is required.

While analyses adjusted for a range of child, household, and village level variables, additional unmeasured factors likely played a role in the relationship between ECE quality and child development. For instance, children are instructed in the country's official language, Lao. However, for many children (particularly considering the study sample included a majority of Khmu children), Lao is not their mother tongue (first language) and so this will influence the ability good quality ECE has to promote development. Children's experience of disability is also an important consideration when investigating the relationship between ECE quality and child development, to help inform strategies for inclusive quality education. Additionally, time spent in ECE is likely to influence the relationship between ECE quality and children's development. The current study was not able to incorporate ECE dose and the relationship between ECE quality and children's outcomes would likely differ for children who attended ECE, for instance, once versus every day of the week. Future studies should consider measuring language of instruction, disability, and ECE dose, as well as other factors likely to influence the ECE quality and child development relationship.

Finally, children in our sample resided in provinces throughout Northern Lao PDR attending three different forms of ECE. This does not include representation of kindergarten, one of the main forms of early education in Lao PDR. Our sample also was not representative of children living in the poorest communities and/or households in Lao PDR. While provinces were selected for participation in the ECE Project based on high levels of poverty, districts were only eligible if they had existing education infrastructure. As a result, research findings may not generalise beyond the setting of the current study.

Implications

Findings from this study advance understanding of ECE quality both in Lao PDR and LMICs more broadly, including measurement approaches and the relationship between quality and children's development. Using a psychometrically robust and locally relevant measure of children's development, the eHCI, this study is the first to demonstrate associations between process quality in ECE classrooms as measured by the MELE, and children's development. Indeed, aspects of both structural and process quality were associated with child development. Although associations observed were small, which is consistent with research in LMICs[237], findings reiterate the need for continued investment not only in expansion of access to ECE for all children, but also for policy to ensure high quality service provision. Further, the emphasis on quality needs to extend beyond facilities and materials, to adequately equipping schools and teachers with the training and resources required to facilitate learning through what are deemed high quality interactions and responses to children's needs within any given context.

Moving forward, future research avenues offer ways through which current challenges in the measurement of ECE quality in LMICs could be addressed. This includes studies that employ a longitudinal research design and those that consider language of instruction, disability, and ECE dose. Further, recent developments in tools capturing ECE quality in LMICs, such as Teach ECE, offer alternative measurement approaches that should be explored to address challenges faced by measures currently used. While SDG target 4.2 outlines that access to quality early childhood development, care, and pre-primary education is necessary for all children, what constitutes quality and how countries should track progress against this target lacks consensus. This is a reflection of where the field of evidence regarding ECE quality in LMICs currently stands and the research investment needed to advance understanding in this space. A more robust evidence base will inform targeted policies (e.g., national ECE quality standards and associated monitoring) and interventions (e.g., teacher professional development programs) to promote high quality ECE service provision and ultimately, improvement of children's development outcomes.

Conclusion

Monitoring of ECE quality in LMICs is scarce, limited by a lack of measurement tools that reflect local culture and context as well as implementation difficulties. With rapid expansion of ECE in Lao PDR and LMICs more broadly, the need to investigate quality of ECE through scale up is paramount. Using a built-for-purpose ECE quality measure, the MELE, findings from this study reiterate the need for continued investment not only in expansion of access to ECE for all children, but also to ensure high quality service provision through emphasis on both structural and process aspects of quality. Globally, continued efforts to strengthen understanding of the ECE service aspects that drive positive outcomes for children are required to appropriately inform policies and interventions and ensure children can benefit from investments in early learning environments.

Chapter 7: Discussion and conclusions

7.1 Chapter outline

The research presented throughout this thesis is brought together here in the final chapter. It begins by discussing key findings of the research studies conducted, highlighting how each study supports these learnings and how findings contribute to and advance existing evidence regarding population level measurement of early childhood development. Next, overall limitations of this thesis are discussed, including how future research may overcome these challenges. The chapter concludes with a discussion of the implications of findings for policy and practice, including how the eHCI may be used to facilitate population measurement and monitoring of holistic early childhood development in LMICs moving forward.

7.2 Key findings and contributions

Population level measurement of early childhood development in LMICs has been challenged by a lack of appropriate measurement tools that are feasible for use at scale, reflect local culture and context, and have the sensitivity required to detect variation in children's development and appropriately inform supports to promote children's outcomes. The eHCI was developed with the vision to overcome these challenges. This thesis investigated the psychometric properties and sensitivity of the eHCI in seven LMICs, including Brazil, China, Kiribati, Lao PDR, Samoa, Tonga, and Tuvalu. This research sought to establish reliability, validity, and sensitivity of the tool across diverse settings and advance understanding of how the eHCI can be used to facilitate population measurement of early childhood development. Pre-existing data collected using the eHCI among children aged 2 to 6 years in seven LMICs throughout 2013 to 2020 was utilised in four research studies presented in this thesis. Studies were designed to explore four research questions, and papers presented in Chapters 3 to 6 detailed how the research undertaken has explored and responded to these questions. In this section, the key findings and contributions from across these four research studies are summarised. Key contributions to existing evidence regarding population measurement of early childhood development, with a focus on LMICs, have been organised into five points below.

7.2.1 The eHCI is psychometrically robust

Findings demonstrated that the eHCI was psychometrically robust across the seven LMICs included in this thesis. Specifically, the tool demonstrated adequate internal reliability (Chapter 3), as well as various aspects of construct validity (Chapters 3 and 4) and criterion validity (Chapters 4 and 5). The eHCI maintained adequate psychometric properties despite the necessary processes of local adaptation, translation as well as different methods of implementation in each country. Beyond this, the instrument proved reliable and valid notwithstanding the great cultural and societal diversity in these countries. Few population measures of early childhood development have demonstrated comprehensive reliability and validity evidence across LMICs in this way. For instance, among available measures summarised in Chapter 1, while the EDI and IDELA have several published reliability and validity studies across countries[82, 238], psychometric evidence for remaining tools is limited.[29] Specifically, some instruments lack any published reliability and validity investigations. For other tools, such as the MELQO MODEL, while psychometric investigations have been conducted, strength of reliability and validity results across countries have been mixed.[84, 87]

Overall, results from this thesis indicate that the eHCI can be used across diverse, low resource settings to facilitate reliable and valid measurement of early childhood development. Prior to this, there did not exist psychometric evidence for the eHCI across countries beyond content and face validity through the process of local expert consultation in each setting. Moving forward, based on the findings presented in this thesis, countries, their governments, researchers, and communities can be confident in implementing the eHCI to accurately capture the state of their children's development. Further, although the eHCI was originally designed to measure development among children aged 3 to 5 years[99], results presented in Chapters 3 to 6 of this thesis indicate the tool can be validly applied to children aged 2 to 6 years. Being psychometrically robust for use among a wider age range of children expands applicability of the tool, and this is discussed along with other implications of findings for policy and practice in Section 7.4 below. Additionally, results across studies demonstrated that both caregivers and teachers provided responses to the eHCI that were reliable and valid. This included caregivers with low literacy and education (i.e., most caregiver in Lao PDR had not completed primary school). This is aligned with the instrument's intended pragmatic and feasible nature in that information can be collected through methods most convenient in any particular setting.

7.2.2 Evidence for predictive validity in LMICs

From a public health perspective, it can be argued that predictive validity is most important among all psychometric properties of a measure. This is because measures of child development should have the ability to predict meaningful outcomes later in life (e.g., cognitive and social-emotional capabilities, educational attainment) if scores on such measures are to be used to appropriately inform investments designed to promote children's development. Alternatively, using a measure with poor predictive validity can lead to ill-informed policy and misdirected resources (e.g., implementation of interventions designed to promote aspects of development that are inconsequential for later life outcomes). Despite this, limited early childhood development measures have established predictive validity, particularly in LMICs.[29, 193] This was the case prior to the commencement of this thesis in 2017 and remains true upon its completion. Population measurement and monitoring of early childhood development in LMICs has gained momentum in recent decades. Time is a critical factor in investigations of predictive validity (e.g., ideally several years will have passed since collecting baseline scores on a tool, before children's later outcomes can be measured and thus predictive validity explored), so it would be expected that recently developed measures would seek to establish predictive validity once longitudinal data becomes available. However, even many of the longer standing measures focused in LMICs have not published predictive validity evidence. For instance, the EAP-ECDS, developed in 2010 and implemented across several countries in East Asia Pacific, has not published predictive validity investigations for either the original long form instrument or the more recently development short form tool.

Chapter 5 of this thesis established adequate short term predictive validity of the eHCI. Specifically, poor scores on the eHCI among children aged 2 to 5 years in Lao PDR predicted poor cognitive development (i.e., literacy, numeracy, and executive function) four years later. Globally, this represents the first evidence for the predictive validity of an adult report population measure of early childhood development in LMICs, and thus a key contribution of the research undertaken within this thesis to existing literature. More broadly, these results also strengthen current evidence regarding how a child's early development influences later outcomes in LMICs, which is scarce. Further, use of Receiver Operator Characteristic curve analysis typically applied to clinical screening tools provided a more comprehensive approach to exploring the predictive ability of the eHCI, relative to existing studies exploring predictive ability of population measures only using single or multiple variable regression associations (e.g., studies focused on the EDI in Canada and Australia, and the IDELA in Ghana)[157, 158, 239, 240] which provide little information about overall predictive ability of a measure. Overall, this predictive validity evidence generated supports use of information collected via the eHCI to inform early years investments and/or shifts in policy and practice, with findings suggesting that employing strategies to promote children's scores on the eHCI have the potential to provide meaningful benefits for later outcomes. However, as outlined in Chapter 5 and again below in Section 7.3.2, further investigations are needed to strengthen these claims.

7.2.3 Importance of capturing children's holistic development

The eHCI was developed through a process of local expert consultation combined with existing theoretical conceptualisations to capture children's holistic development across domains including Physical Health, Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, Approaches to Learning, Numeracy, Reading, and Writing. While some of these domains are common in measurement instruments designed for use in Western, high-income settings (e.g., aspects of language, social emotional skills, literacy, and numeracy), others are not (e.g., knowledge of culture, perseverance), reflecting differences across cultures and contexts in terms of the skills and capabilities of a well developing child.[29, 93]

Emphasis on holistic development has gained momentum in recent decades, highlighted by inclusion of targets related to children's health, learning, and psychosocial wellbeing in the United Nations Sustainable Development Agenda.[61] The findings presented in this thesis provide unique additional evidence to the existing literature, which further strengthens the case for capturing children's holistic development in population measurement.

Research studies in this thesis indicated some variation in the strength of reliability and validity across individual eHCI domains measuring different aspects of children's development (e.g., Physical Health and Perseverance domains demonstrated less than satisfactory internal reliability). Importantly however, findings demonstrated that the instrument's summary measure, eHCI overall development, which combines information from all nine domains, was particularly valuable. Namely, in Chapter 5, eHCI overall development was consistently the best predictor of children's later cognitive outcomes, relative to individual eHCI domains (e.g., Reading, Writing etc). Further, findings from this study highlighted that a holistic measure of children's development (i.e., eHCI overall development) provided a better predictor of later cognitive outcomes, relative to other child development measures that captured literacy and numeracy alone (i.e., the version of the MELQO MODEL utilised in the ECE Project in Lao PDR). Considering the paucity of evidence regarding how early development influences later outcomes in LMICs, results indicating the predictive validity of children's holistic development make an important contribution to existing knowledge.

Additionally, in Chapter 6, children's eHCI overall development scores were associated with ECE quality, including both structural and process components of the classroom. These findings help to highlight another strength of a holistic measurement approach. When measuring development among a wide age range of children (e.g., 2 to 6 years), some aspects of development will have floor or ceiling effects (i.e., if items are too hard or too easy), dependent on child age. For instance, floor effects may be observed on literacy and numeracy among 2-year-olds, and ceiling effects may be observed on communication skills among 6-year-olds. A holistic measure combining domains (e.g., eHCI overall development), is less prone to this issue and enables greater sensitivity, as previously discussed. Overall, findings generated throughout this thesis highlight the importance of capturing children's holistic development, as locally relevant. Short tools designed to be added into household surveys for global comparison (e.g., ECDI2030) may be too blunt (i.e., capture limited information and therefore lack variability in scores) to be capable of this.

7.2.4 Importance of reflecting local culture and context

Acknowledging that children's development is shaped by culture and context[94, 95], the eHCI was designed to be adaptable across diverse settings to ensure the tool generates locally relevant results.[99] Specifically, users are able to revise, add, or remove items from the instrument as relevant to their setting. The commentary article presented in Appendix H, co-authored by the Candidate and published in February 2020[93], highlights the need for measures of early childhood development to reflect culturally-valued capabilities of children developing well. Thus, measures designed to be culturally "fair" or "neutral" (i.e., items or activities that seek to operate identically across cultures) for the purposes of global monitoring and international comparison do not reflect important attributes of children's development within local cultures and therefore have limited relevance and utility when it comes to informing intervention and policy within countries.[93] This is aligned with Section 7.2.3 above, in that a culturally fair or neutral tool will not capture children's holistic development, as it does not include culturally relevant capabilities.

Findings presented throughout this thesis support the need for tools to be aligned with local culture to produce information that both accurately reflects children's abilities and is relevant to local policy and practice. Further, use of a locally adapted tool and generation of context-specific data supports local ownership of results, which is critical for local action, driving investments in children's early development. In Chapter 3, while eHCI domains measuring children's literacy and numeracy yielded consistent internal factor structure results across countries, findings across remaining domains demonstrated greater variation across countries, reflecting the influence of culture and context on the capabilities measured by these domains including communication, health and hygiene practices, and social and emotional skills. In Chapter 4, variation in the strength of associations between eHCI domains and child and family characteristics also highlighted the context-specific nature of early childhood development measurement. At the outset of this thesis, in 2017, emphasis in the field of early development measurement centred on global comparability and cultural neutrality, in light of ratification of the Sustainable Development Agenda in 2015 and the requirement for global monitoring of children's health, learning, and psychosocial wellbeing.[61] Now, in 2022, increasingly measurement approaches are seeking to combine both a culturally neutral set of core items that can be used across countries for the purposes of global comparison, together with items that capture holistic development as relevant to local culture and context (e.g. AIM-ECD).[66] Results from the research studies included in this thesis, as well as the arguments presented in the commentary article described, have served as proponents in this global debate and strengthened the argument for the need for locally relevant measurement of early childhood development.

7.2.5 The eHCI can be used for dual purposes

Selection of an instrument is largely driven by the purpose of measurement.[29] Existing population measures of early childhood development range from brief, adult report tools intended for global population monitoring (e.g., MICS ECDI), to detailed direct assessment instruments designed to aid program evaluation (e.g., IDELA, EAP-ECDS). Results presented throughout this thesis indicate the eHCI may strike a balance between the two, with findings demonstrating the tool is suitable for both applications. In Pacific Island Countries, including Kiribati, Tonga, Tuvalu, and Samoa, the eHCI was employed using a census approach for population monitoring purposes. The tool was feasible for collection of information among all children aged 3 to 5 years nationally, with Chapters 3 and 4 demonstrating reliability and validity of results. In Brazil, China, and Lao PDR, countries with considerably larger populations, the eHCI was applied through a sample approach for varied purposes, including identification of inequalities in children's development in China, or measuring impact of early years interventions in Lao PDR. Chapters 3 and 4 also demonstrated reliability and validity of information collected through this approach. Further, Chapter 5 showed that the eHCI predicted children's later cognitive outcomes, and Chapter 6 demonstrated that scores on the tool had adequate sensitivity to variations in children's development according to ECE quality. Together, results have demonstrated that the eHCI has achieved its aims of being feasible and valid for purposes including population monitoring, program evaluation, as well as longitudinal studies seeking to predict children's future capabilities.[99]

Efforts in the instrument's design to ensure feasibility for use at scale in low resource settings, as well as adaptability to each new context, have not hindered adequate reliability and validity. Taken together, findings suggest the eHCI may be validly applied for the dual purposes of population monitoring and program evaluation. Prior to this thesis, no population measure of early childhood development had demonstrated such applicability in LMICs. For instance, while the teacher-reported EDI is feasible for population monitoring, as has been shown in Canada and Australia, demonstrating predictive validity in both settings[157, 239, 240], the tool cannot be adapted to be appropriate for use in in diverse settings including LMICs due to copyrighting and licensing restrictions. On the other hand, while the IDELA has demonstrated predictive validity in Ghana[158], the instrument does not share the same pragmatic nature of the eHCI as it is administered via direct assessment; a method that is not typically feasible for implementation at scale in low resource settings.

Indeed, though direct assessment measures are often argued to be the gold standard approach to measuring children's development, relative to adult report measures[89], findings from Chapters 5 and 6 indicate that the eHCI (i) better predicted children's later outcomes and (ii) exhibited greater sensitivity to variation in children's development according to ECE quality, relative to a direct assessment of children's development. As discussed in Chapters 5 and 6 (and reiterated in Section 7.2.3), this is argued to be partly attributable to the holistic nature of the eHCI (i.e., capturing both cognitive and non-cognitive development), versus direct assessment measures focused on cognitive development only. Nevertheless, results cast some doubt over the argument for the superiority of direct assessment of children's development and warrant further investigation into this widely accepted gold standard. In this way, the eHCI and the findings presented in this thesis make a unique contribution to current knowledge.

7.3 Limitations and recommendations for future research

While limitations of individual research studies have been detailed in corresponding chapters, this section discusses two overarching limitations of the research presented in this thesis.

7.3.1 Respondent reliability

As described in Chapter 1, two aspects of reliability that are important to establish when exploring the psychometric properties of a measurement tool include consistency in scores over time (test-retest reliability) and across respondents (inter-rater reliability).[110, 111] The data required to conduct these investigations had not been collected and/or was not accessible at the time of undertaking this research[d] and could not be collected within the scope of this thesis given the context within which the eHCI was deployed in LMICs (i.e., as part of independent research programs). Specifically, data collected for the same child via the same respondent at two timepoints within a short timeframe (e.g., a few weeks to a month) to evaluate test-retest reliability, and data collected for the same child at the same time via two different respondents (e.g., a child's caregiver and teacher) to assess inter-rater reliability.

Overall, results from this thesis show that the eHCI is reliable and valid through administration via both caregivers and teachers as respondents. However, future research should explore test-retest reliability and inter-rater reliability to establish consistency of the tool over time and across respondents. Considering the eHCI is typically completed by children's caregivers or teachers, it is particularly important to gain a better understanding of the level of agreement between how caregivers and teachers respond to the tool (i.e., inter-rater reliability). Existing research has demonstrated that caregivers and teachers tend to rate children differently on measures of early learning and development[241-243], with results dependent on children's age,

[d] In Tonga and Tuvalu, some children had eHCI records completed by both their caregiver and preschool teacher. In Tuvalu this duplication was planned while in Tonga it was not. These records had limited respondent information, and so it was decided not to use this data for inter-rater reliability investigations.

developmental domains assessed, as well as context of measurement. Although the majority of this research has been conducted in high-income countries, this phenomenon may be particularly applicable in LMICs whereby it is more likely that a larger proportion of caregivers have received little to no education, or teacher training is conducted with varying standards.

Research investigating the reliability and validity of the eHCI in Shanghai, China, was published in 2020, including test-retest and inter-rater reliability among samples of 183 and 168 children, respectively.[105] Test-retest reliability among caregivers demonstrated strong results for eHCI overall development scores, with less consistency in scores on Social and Emotional Skills, Perseverance, and Approaches to Learning domains. Inter-rater reliability analyses showed moderate agreement between caregivers and teachers on eHCI overall development, with stronger agreement for measures of children's literacy and numeracy compared to that for other domains. Although this research provides initial insight into test-retest and inter-rater reliability of the eHCI, Shanghai, China's urban capital, is a context quite different to other countries in which the eHCI has been implemented. Therefore, it is particularly important that future research explores whether there exists an adequate level of agreement between caregiver and teacher ratings of children's capabilities on the eHCI; not with the intention of determining a "more reliable" respondent as the eHCI seeks to be pragmatic including implementation via the most convenient method, but rather to establish adequate agreement and further bolster confidence in the tool.

7.3.2 Generalising findings globally

Recently, McCoy proposed a model or framework of "developmental universality with specificity".[95] In this article, current understanding of the extent to which early childhood development is similar (universal) and different (specific) across cultures was summarised. For instance, there is evidence regarding the homogeneity of children's mastery of basic motor development milestones across countries, with a study demonstrating that 90% of samples of infants in Ghana, India, Norway, Oman, and the United States followed the same sequence of sitting, crawling, standing, and walking within a similar age range.[244] The same has been demonstrated for children's language and cognitive milestones.[95] Alongside this there exists evidence for differences in the form (what it looks like), timing (when it emerges), and relevance (how it is valued) of early childhood development across cultures. One example provided is variation across countries in the extent to which children exhibit prosocial behaviour such as sharing or helping others (i.e., form), with differences largest among older children versus younger children (i.e., timing).[245, 246] Specificity is described to derive from differences in societal norms as well as caregiving practices[247]. This evidence highlights an important consideration when reflecting on the research settings in which the eHCI was conducted and thus the children and countries that were included in this thesis.

Specifically, the research presented here is based on data collected in seven LMICs. These countries capture some of the diversity that exists across LMICs, including differences in population size, income classification, culture, and early years service landscape. However, they are also homogenous in other ways. For instance, six of the countries are situated in East Asia and Pacific, with one country (Brazil) in Latin America. While results have demonstrated that the eHCI is psychometrically robust across these seven countries, this somewhat limited geographical spread and therefore cultural diversity is important to consider in terms of the generalisability of findings to world regions beyond those included in the studies presented here. It is important that future research seeks to adapt the eHCI, replicating reliability, validity, and sensitivity findings across countries in other world regions including Europe and Central Asia, Middle East and North Africa, South Asia, and Sub-Saharan Africa, particularly the latter two regions where a large proportion of the world's children at risk of not achieving their development potential reside.[9] Perhaps additional developmental aspects/domains will emerge as important for a well-developed child in different settings, or related to Section 7.3.1 above, perhaps implementation in other world regions will highlight important considerations for reliability of different respondents to the tool. This additional research will instil further confidence in the eHCI as a feasible, reliable, and valid tool for use globally.

Further, investigations of the predictive validity and sensitivity of the eHCI were restricted to Lao PDR. This was enabled through existing data collected as part of the World Bank ECE Project which included collection of the eHCI alongside various other child, family, and village level information longitudinally.[226] To date, data has not been collected in other countries to enable similar investigations. However, as such data becomes available, research should seek to extend findings presented throughout this thesis to other countries and world regions. Again, doing so will further strengthen validity of the tool for global application.

7.4 Implications for policy and practice

This thesis provides evidence for how the eHCI can be utilised to facilitate measurement of holistic early childhood development in LMICs moving forward. Specifically, the tool can be reliably and validly used among children aged 2 to 6 years, through either a census or sampling approach, for the purposes of locally relevant population measurement and monitoring, as well as program evaluation. Ultimately, both applications (i.e., census, sample) collect information that can be used to inform policy and practice in terms of resourcing and supports to promote children’s development. In this section, implications of these findings are discussed, organised into two overarching points, namely (i) next steps for the field of population measurement of early childhood development, including the eHCI, as well as (ii) next steps for using population data on early childhood development to promote children’s outcomes.

7.4.1 Next steps for population measurement of early childhood development

Ratification of the Sustainable Development Agenda in 2015 and the requirement for countries to monitor children's health, learning, and psychosocial wellbeing (i.e., SDG 4.2.1)[61] saw an emphasis in the field of early childhood development measurement on cultural neutrality to enable global comparability. In 2019, the ECDI2030, a 20-item caregiver report tool that determines the percentage of children aged 24 to 59 months developmentally on track in health, learning, and psychosocial wellbeing, was established as the measure for global monitoring against SDG 4.2.1.[72-74] Countries will utilise this tool to track progress against the SDG indicator, with the ECDI2030 a standard inclusion in UNICEF MICS7 (national household surveys) from 2023 onward.[75, 76] Importantly, this will enable a comparison of the percentage of children developmentally on track within and across countries and world regions (see for example a comparison of ECDI results across 66 countries[248]), highlighting inequalities in children's outcomes and informing targeting of global resources and supports. However, it is important that countries understand what population monitoring using the ECDI2030 will not provide. It will not holistically capture the skills and capabilities of what is considered a well-developed child as locally or culturally relevant in any particular context. Further, due to its brevity together with universality of items eliminating aspects of culture from measurement, it is likely that the tool will not provide results that reflect adequate sensitivity to variation in children's development as a result of interventions or supports, and therefore will not be appropriate to facilitate evaluation of programs or policy. Whether the ECDI2030 is valid for such an application currently remains untested. If countries, particularly LMICs with limited resources, are to invest in population measurement of early childhood development, it would be remiss not to ensure the endeavour is most efficient and could serve all of these purposes.

To this end, now in 2022, increasingly measurement approaches are seeking to combine both a culturally neutral set of core items that can be used across countries for the purposes of global monitoring, together with items that capture holistic development as relevant to local culture and context. Such a solution was proposed in the commentary article by Rao and colleagues (Appendix H), arguing that global monitoring of children's development does not need to come at the cost of producing locally relevant results.[93] A recently developed example of such an approach is the AIM-ECD; a core set of items that measure early childhood development among children aged 4 to 6 years through both caregiver report and direct assessment modules.[66, 88] The instrument includes 20 caregiver report items and 84 direct assessment items spanning domains including early literacy, early numeracy, executive functioning, and social-emotional competencies. As described in Chapter 1, the AIM-ECD intends to be used in conjunction with items that capture local contexts, to facilitate scale up of early childhood measurement and advance understanding of both universal and context-specific factors that influence children's development to inform policy and practice.[66, 88]

The AIM-ECD was developed through a process of harmonisation of data from existing measures of child development from across LMICs. To highlight overlap between measures, authors compared the AIM-ECD core items with items in other instruments including the eHCI, EAP-ECDS, and IDELA. Of all measures examined, the eHCI had the greatest overlap. Specifically, it contains all early literacy and numeracy items (11), and five of the nine executive functioning and social-emotional competencies included in the AIM-ECD caregiver report core items.[66] This overlap means that few items would need to be added to measures to ensure they contain the core items for global comparison, while simultaneously achieving aims of the original measure (e.g., locally relevant program evaluation). Indeed, perhaps the same approach could be applied to the ECDI2030 among children aged 2 to 4 years as the eHCI also shares items with this measure. As the eHCI is freely available for anyone to use, it may facilitate such novel, multi-purpose strategies. Findings regarding how this approach works in practice are not yet available, and such investigations will be important in guiding future population measurement of early childhood development.

The eHCI was developed with a focus on being feasible, locally relevant, and psychometrically robust.[99] The instrument was not originally intended to be used for international comparison of children's development. Indeed, reliability and validity investigations required for such a tool (e.g., measurement equivalence across cultures[249]) would need to be conducted in addition to investigations presented in this thesis. However, considering the findings demonstrated throughout this thesis, it may be that the eHCI could be applied using the same approach. While the tool is intended to be adaptable for different cultures, many items have remained the same across country versions of the tool to date. For instance, a comparison of versions of the eHCI utilised in the seven countries included in this thesis demonstrated that approximately 30 items (about 45-55%, dependent on adaptation) were maintained across all versions. In using the eHCI moving forward, it may be that all countries could maintain a core set of the instrument's items in addition to those that fulfil their measurement needs as identified through local adaptation, whether that be cultural-specific items or items designed for policy or program evaluation. Such an approach may have broader utility than that using the AIM-ECD considering the wider age range captured by the eHCI (2 to 6 years), as well as greater flexibility in terms of respondents (caregivers and teachers).

Overall, while the ECDI2030 will continue to be implemented for the purposes of tracking progress towards the Sustainable Development Agenda, novel approaches incorporating core items for global comparison of children's development as well as locally relevant items, such as the AIM-ECD or the eHCI, should be explored in future early childhood development measurement endeavours. Implementing an approach that seeks to serve dual purposes would likely be a more efficient use of resources in LMIC settings, and thus better facilitate scale up of early childhood development measurement.

7.4.2 Next steps for using population data to promote children's outcomes

This thesis has centred on approaches to generating population early childhood development data. Such data are important in demonstrating impact of interventions and supports designed to promote children's developmental outcomes, which is a necessary step in bringing programs to scale and maximising benefits across populations.[192] It is important to acknowledge that data generation plays one part in the complex landscape of the systems required to support children's development, and indeed health and development across the life course.

Currently, there exist training and resources to support adaptation and implementation of early childhood development measurement tools, as well as guides to facilitate analysis of data collected and dissemination of results. What is missing from the suite of supporting materials for all measures designed for use in LMICs discussed, including the eHCI, are guidelines on how data generated can be used to inform implementation of interventions or policy. The 2022 Lancet series on Optimising Child and Adolescent Health and Development provides an array of what are deemed effective and affordable interventions to promote child health and development in LMICs, spanning preconception and pregnancy to 20 years of age.[13] However, while this summary includes many health interventions (e.g., promoting adequate nutrition or the prevention of infectious diseases), only a small number of interventions identified target the promotion of children's development including integrated responsive caregiving and early learning interventions[5, 250-252], as per the Nurturing Care Framework, and school-based social and emotional learning and psychosocial interventions.[253, 254]

Currently, over seven years since ratification of the Sustainable Development Agenda, gaps in children's health and development persist within and across countries, reflecting inadequate coverage and implementation of high-quality supports required for nurturing care.[12, 255] Coupled with the effects of the COVID-19 pandemic on economies, services, systems, and populations undermining previous progress, evidence-based interventions must be brought to scale, with integration across both systems and the life course, in order for countries to achieve the Sustainable Development Agenda.[248] Data generated through locally relevant, valid, and reliable population measurement of early childhood development, such as the eHCI, can be used to inform the targeting of supports for children's development and formation of human capital. However, the next decade of the Sustainable Development Agenda era must not only encourage ongoing monitoring, but emphasise supports for countries, their governments, and communities to operationalise information collected into implementation of evidence-based interventions and policy designed to ensure that all children can survive and thrive.

7.5 Thesis conclusion

The eHCI was designed to overcome challenges of existing population measures of early childhood development, particularly in low resource settings. Prior to this thesis, the eHCI had been implemented in several LMICs as part of various independent projects seeking to promote children's health, development, and education, though there did not exist evidence for the reliability, validity, and sensitivity of the tool across countries. The research presented in this thesis has demonstrated that the eHCI is psychometrically robust across diverse settings for use among children aged 2 to 6 years for the purposes of population monitoring as well as program evaluation. Research learnings bolster confidence in continued use of the tool to facilitate generation of locally relevant population level early childhood development data. Moving forward, freely available for anyone to use, the eHCI will continue to feed into new approaches that seek to simultaneously achieve measurement of locally relevant holistic development alongside monitoring for global comparison; the latter required for tracking progress toward the Sustainable Development Agenda. Investigations into how this approach works in practice will be important in guiding future population measurement of early childhood development.

Appendices

Appendix A: University of Adelaide Human Research Ethics Committee exemption

Human Research Ethics Committee (HREC)

THE UNIVERSITY of ADELAIDE

ETHICAL ISSUES CHECKLIST FOR HUMAN RESEARCH

Who should use this checklist and why?

University staff, students enrolled in the University and persons in any way associated with or sponsored by the University who are involved in a human research project must ensure that the project has undergone the appropriate level of ethical review before it can commence. Ethical review is undertaken by the University's Human Research Ethics Committee (HREC) at various levels. The following information and checklist is designed to assist researchers when determining the appropriate level of ethical review for the project.

What is a human research project?

Human research is research conducted with or about people, or their data or tissue. It can be broadly understood to include, but is not limited to: taking part in surveys, interviews or focus groups; undergoing psychological, physiological or medical testing or treatment; being observed by researchers; the collection and use of participants' body organs, tissues or fluids; and accessing health and medical records, personal documents or other materials.

Research involving people must comply with current ethical standards. In Australia, these ethical standards are set by the National Health and Medical Research Council (NHMRC). Researchers should be familiar with the relevant publications of the NHMRC and all human research must comply with the:

- National Statement on Ethical Conduct in Human Research (2007)
- NHMRC Values and Ethics - Guidelines for Ethical Conduct in Aboriginal and Torres Strait Islander Health Research
- NHMRC Guidelines Under section 95 and section 95A of the Privacy Act 1988.

University staff and students in the practice of their research must be aware of and adhere to the Australian Code for the Responsible Conduct of Research 2007.

Note: all researchers proposing to conduct surveys with University staff, students or Alumni as participants must also comply with the University's Survey Framework administered by the Survey Reference Group (SRG). This approval process is independent of the HREC. For more information refer to https://www.adelaide.edu.au/learning/reviews/surveys/

Levels of ethical review

Research projects with different levels of risk to participants are subject to different requirements for review. References to 'NS' are to the National Statement on Ethical Conduct in Human Research (2007).

A risk is a potential for harm, discomfort or inconvenience. It involves:

- the likelihood that any harm (or discomfort or inconvenience) will occur; and
- the severity of the harm, including its consequences.

Guidance on different levels of risk, and the differences between 'inconvenience', 'discomfort' and risks that are 'more serious than discomfort' is contained in NS Chapter 2.1 (NS 2.1). Researchers should be familiar with these definitions when considering the level of review that is required to be undertaken.

There are 3 review categories; research that is exempt from HREC review, research eligible for low risk review and research requiring full review.

Research exempt from HREC review

Research which has undergone a thorough review by the researchers against the NS and is considered to be 'negligible risk research' is exempt from HREC review.

Research that satisfies both of the following conditions is exempt from ethical review:

- It is 'negligible risk' research: there is no foreseeable risk of harm or discomfort; and any foreseeable risk is no more than inconvenience
- It involves the use of existing collections of data or records that contain only non-identifiable data about human beings.

Research that uses qualitative methods e.g. face-to-face interactions between participant and researcher; personal interviews; focus groups; oral histories; observation; and on-line research that identifies participants, carries more than negligible risk and therefore requires HREC review (NS 3.1).

Researchers are to keep an auditable record of any research that has been exempted from review (NS 5.2.9).

Research eligible for low risk review

'Low risk research' describes research which satisfies both of the following conditions:

- There is no foreseeable risk of any harm to participants and others; and any foreseeable risk is no more than discomfort. Discomfort can include minor side-effects of medication, discomfort when taking blood pressure, and anxiety induced by an interview (NS 2.1).
- It is not included in the following categories:
    - the use without consent of personal information in medical research, or personal health information (NS 2.3.5).
    - interventions and therapies: including clinical and non-clinical trials, and innovations (NS 3.3)
    - human genetics (NS 3.5)
    - human stem cells (NS 3.6)
    - women who are pregnant and the human fetus (NS 4.1)
    - people highly dependent on medical care who may be unable to give consent (NS 4.4)
    - people with a cognitive impairment, an intellectual disability, or a mental illness (NS 4.5)
    - children and young people (with some exceptions) (NS 4.2)
    - people who may be involved in illegal activities (NS 4.6)
    - Aboriginal and Torres Strait Islander peoples (NS 4.7)
    - people in other countries (with some exceptions) (NS 4.8).

Low risk research may be reviewed either through a subcommittee of the HREC or executively approved by the HREC Convener.

Where the risk, even if unlikely, is more serious than *discomfort*, the project will require full review. Research timetables should allow for the possibility that a project submitted as a low risk application may be deemed to involve more than low risk, or to raise other issues, therefore requiring full review. Researchers may also be required to provide additional information.

Research requiring full review

Research that is included in the NS categories listed above, or where the level of risk is more serious than discomfort i.e. harm, or includes concealment or deception, must be submitted for full review.

Harm includes physical, psychological, devaluation of personal worth, social, economic and legal harm (NS 2.1). Research where the true purpose, or the collection of data itself, is concealed from participants or where participants are deceived is not considered ethical unless compelling reasons are given for its use.

The use without consent of personal information in medical research, or personal health information requires full review, whether the research is low risk or not (NS 2.3.5).

Using the Ethical Issues Checklist

In the next section you are given a checklist and asked to answer some questions about your research project. This list contains areas and methods of research that have previously raised ethical or other issues by the HREC. Where applicable these are cross referenced to the corresponding chapter of the NS for you to read.

Important: this checklist is a general guide only and does not replace the need for researchers to do their own thorough review against the NS. If you require additional information or advice as to the appropriate level of ethical review for your research after completing the checklist, reading the NS or discussing your project with colleagues or supervisors you should contact the HREC Secretariat.

Note, all researchers proposing to conduct surveys with University staff, students or Alumni as participants must also comply with the University's Survey Framework administered by the Survey Reference Group (SRG). Please refer to https://www.adelaide.edu.au/learning/reviews/surveys/

THE UNIVERSITY of ADELAIDE

Human Research Ethics Committee (HREC)

ETHICAL ISSUES CHECKLIST FOR HUMAN RESEARCH

Project Title: *(insert)*

Validating a population measure of early childhood development in low and middle income countries: The early Human Capability Index

To activate a tick box, double click left mouse button and a 'check box form fields' box appears. Select 'checked' and OK. To untick do the same again and select 'unchecked' and OK.

Does your research involve?

Research Areas and Methods	**Yes**	**Possibly**	**No**
The use without consent of personal information in medical research, or personal health information?	☐	☐	☒
Interventions and therapies: including clinical and non-clinical trials, and innovations? (NS 3.3)	☐	☐	☒
Human genetics? (NS 3.5)	☐	☐	☒
Human stem cells? (NS 3.6)?	☐	☐	☒
Assisted reproduction technology?	☐	☐	☒
Ionizing radiation (x-rays, fluoroscopy or radioisotopes)? (NS 3)	☐	☐	☒
Controversial topics in the community (e.g. AIDS, IVF)?	☐	☐	☒
Women who are pregnant and the human fetus? (NS 4.1)	☐	☐	☒
Children or young people under the age of 18? (NS 4.2)	☐	☐	☒
Persons in dependent or unequal relationships e.g. between participants (line manager-staff, employer-employee) or between participants and researchers (teacher-student)? NS 4.3)	☐	☐	☒
People with a cognitive impairment, an intellectual disability, or a mental illness? (NS 4.4, 4.5)	☐	☐	☒
People who may be involved in illegal activities? (NS 4.6)	☐	☐	☒
Separate identification of, or focus on, Aboriginal and Torres Strait Islander peoples? (NS 4.7)	☐	☐	☒
Separate identification of, or focus on, other minority or culturally sensitive groups of people or their communities?	☐	☐	☒
Research on people living overseas (NS 4.8)?	☐	☐	☒
Research methods that identify participants, e.g. questionnaires, surveys, personal interviews (group or individual), focus groups, oral histories, observation and on-line research?	☐	☐	☒
Involves collection of data or interviews not in a public place i.e. in people's homes?	☐	☐	☒
Research where the true purpose, or the collection of data itself, is concealed from participants or that participants are in any way deceived?	☐	☐	☒
Any situation where the informed consent of participants cannot be given or assured e.g. observation of participants without their knowledge?	☐	☐	☒
Payments or inducements other than reasonable reimbursement for travel, time spent participating in the research?	☐	☐	☒
Any perceived, possible or actual conflicts of interest e.g. where a researcher or organisation is in a commercial relationship?	☐	☐	☒

Go to next page…

Is HREC review required?

If you answered **Yes** or **Possibly** to any of these questions the research will require HREC review and you will need to submit an application for low risk or full review.

If you answered **No** to **every** question this is a good indicator that your research is 'negligible risk research' and is exempted from HREC review remembering that this checklist is a general guide only and does not replace the need for researchers to do their own thorough review against the National Statement.

Researchers satisfied that the research is negligible risk research should sign the declaration below and retain this checklist and declaration.

This document can be used in the auditable record of any research undertaken that has been exempted from HREC review (NS 5.2.9).

Researcher's Declaration:

I/we have read the National Statement on Ethical Conduct in Human Research (2007).

I/we, the researcher(s) agree:

- That the research to be undertaken in the above named project is 'negligible risk research' and is exempted from review by the Human Research Ethics Committee (HREC);
- To only carry out this research project where adequate funding and personnel is available to enable the project to be carried out according to good research practice and in an ethical manner;
- To maintain the confidentiality of all data from or about project participants and to store the data according to the Australian Code for the Responsible Conduct of Research 2007 and other legislative requirements;
- To notify the HREC in writing immediately if any change to the project is proposed that might require review by the HREC and await approval before proceeding with the proposed change.

Principal researcher's signature:		Name:	A/Prof Sally Brinkman	Date:	15/11/2017
Researcher signature:		Name:	Ms Alanna Sincovich	Date:	15/11/2017

Note: add additional signature lines as required.

Appendix B: Published journal article 1

Sincovich et al. BMC Pediatrics (2019) 19:471
https://doi.org/10.1186/s12887-019-1852-5

BMC Pediatrics

RESEARCH ARTICLE Open Access

Measuring early childhood development in multiple contexts: the internal factor structure and reliability of the early Human Capability Index in seven low and middle income countries

Alanna Sincovich[1,2*], Tess Gregory[1,2], Cristian Zanon[3], Daniel D. Santos[4], John Lynch[1,5] and Sally A. Brinkman[1,2]

Abstract

Background: The fourth year of the Sustainable Development Agenda era calls for countries to continue to invest not only in interventions and policies that will promote global equity and sustainability, but also in the monitoring systems required to track progress against these targets. A more pragmatic solution to measuring children's early development in low and middle income countries in particular, is required. This study explores the psychometric properties of the early Human Capability Index (eHCI), a population measure of holistic development for children aged 3–5 years, designed with the vision of being flexible and feasible for use in low resource and capacity settings.

Methods: Utilizing data from seven low and middle income countries: Brazil ($n = 1810$), China ($n = 11,421$), Kiribati ($n = 8339$), Lao People's Democratic Republic ($n = 7493$), Samoa ($n = 12,191$), Tonga ($n = 6214$), and Tuvalu ($n = 549$), analyses explored the internal factor structure and reliability of scores produced by the tool within each country.

Results: Confirmatory factor analyses and internal consistency coefficients demonstrated that after local adaptation, translation, and different implementation methods across countries, the eHCI maintained the same factor structure of nine theoretically-based developmental domains: Physical Health, Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, Approaches to Learning, Numeracy, Reading, and Writing.

Conclusions: Findings support the aims of the eHCI in being adaptable and applicable for use within a range of low and middle income countries to facilitate measurement and monitoring of children's early development, as is required for the tracking of progress towards the Sustainable Development Agenda.

Keywords: Child development, early Human Capability Index, Low and middle income countries, Program evaluation, Population monitoring

Background

Global endorsement of the United Nations Sustainable Development Goals (SDGs), an agenda for which the healthy development of children is central [1], together with burgeoning evidence regarding the value of investing in children's early years [2], have highlighted the need for services and supports that provide children with the opportunity not only to survive, but to developmentally thrive [3]. In turn, the creation and utilization of instruments that measure such development in children has gained momentum. The early Human Capability Index (eHCI), a population measure designed to capture the holistic development of children aged 3–5 years, represents one such effort. This paper presents preliminary evidence of the psychometric properties of scores produced by the eHCI, and highlights how the tool could make an important contribution to the

* Correspondence: alanna.sincovich@adelaide.edu.au
[1]School of Public Health, University of Adelaide, Level 9, Adelaide Health and Medical Science Building, 57 North Terrace, Adelaide, South Australia 5005, Australia
[2]Telethon Kids Institute, University of Western Australia, Level 15, 31 Flinders St, Adelaide, South Australia 5000, Australia
Full list of author information is available at the end of the article

task of evaluating early childhood policies and programs as well as monitoring children's development in the early years.

Tracking progress towards healthy child development

Ratified by United Nations member countries in 2015, the Sustainable Development Agenda specifies 17 goals and 169 targets to end poverty, mitigate inequality, and protect the planet for a better future [4]. The fourth year of the SDG era calls for countries to continue to invest not only in interventions and policies that will promote global equity and sustainability, but also in the monitoring systems required to track progress against these targets and thus identify those at risk of falling behind. Of particular relevance to early childhood development, SDG 4.2 states that by 2030, countries must ensure that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education. To track progress against this target, countries are required to monitor (i) the percentage of children under 5 years of age who are developmentally on track in health, learning and psychosocial wellbeing, and (ii) national participation rates in early childhood education.

While many countries monitor national enrolment rates in early childhood education [5], few track the status of children's early development. Measuring progress toward SDG 4.2 calls for population monitoring of children's early health and development outcomes. Indeed, child development contributes to a number of SDG targets, including those related to health, gender equity, and poverty reduction, and thus global monitoring of children's early development is key to supporting progress toward the broader Sustainable Development Agenda [6]. Faced with limitations in terms of the measurement instruments available as well as the resources and capacity required to implement such monitoring systems, tracking children's health and development in low and middle income countries will be a challenge. In this, however, lies an important opportunity to promote and address the current obstacles associated with measuring children's early development.

A call for new measurement solutions

Measurement of children's development is influenced by culture, language, and theory. What are considered important aspects of and appropriate goals for children's development, as well as what are deemed suitable assessment techniques to capture this information, can vary considerably across cultures and contexts [7–9]. Consequently, although tools need to capture aspects of child development that are important to outcomes throughout the life course, they should also be aligned with local culture and early learning and development frameworks, so that they not only accurately reflect children's capabilities, but also produce information relevant to local policy and practice [1, 8].

A number of measurement initiatives are currently underway to monitor children's early development at national, regional, and global levels. Some examples include the Early Development Instrument (EDI), the Early Childhood Development Index (ECDI), the Caregiver Reported Early Development Instrument (CREDI), the International Development and Early Learning Assessment (IDELA), the East Asia-Pacific Early Chid Development Scales (EAP-ECDS), the Malawi Developmental Assessment Tool (MDAT), the Measurement of Development and Early Learning (MODEL), and the Regional Project on Child Development Indicators (PRIDI) [3]. Various characteristics of the instruments in use however, including the cost of licensing fees, the level of enumerator training required prior to administration, the time they take to administer, how they are administered, and their applicability and adaptability within different contexts, constitute considerable barriers to their utilization. This is especially the case in contexts where resources and capacity are limited. To overcome these challenges, international leaders in early childhood have called for a more pragmatic and reliable solution to measuring children's early development in low and middle income countries in particular. It was against this background that the eHCI was developed.

The early Human Capability Index

Designed to capture key aspects of holistic development in children aged 3–5 years, the eHCI was developed with the vision of being feasible for use in low resource and capacity settings while having the ability to capture change in children's development over time [10]. The tool includes approximately 60 items (dependent upon country adaptation) spanning nine developmental domains (Physical Health, Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, Approaches to Learning, Numeracy, Reading, and Writing) and can be completed via adult report (e.g. by children's caregivers, teachers, or early childhood practitioners) in less than 10 min. The eHCI requires minimal resources to be implemented; the tool is available for anyone to use free of charge, little enumerator training is required, and it can be completed quickly and easily by any adult who knows the child. Further, the tool was designed so that it can be easily adapted and utilized within diverse contexts for a range of purposes, including population monitoring, the evaluation of effects of early childhood policies and programs, as well as longitudinal studies seeking to predict children's future capabilities.

Development of the eHCI

The eHCI was originally developed in 2013 for the purposes of evaluating a program designed to support children and families to be better prepared for school in Tonga [11]. Consultations were undertaken to understand locally, what good child development at school entry looks like, first establishing broad areas (i.e. domains) of development and then identifying specific items within these areas. On the basis of consultations and the child development literature, a draft instrument was developed and independently reviewed by child development experts. Once translated into Tongan, stakeholders originally consulted reviewed the instrument to ensure content and face validity. Piloting was then conducted to determine respondent understanding of items, efficiency of data collection methods, if scale distributions were discriminating between children as theoretically expected, as well as any floor or ceiling effects (i.e. if items were too hard or too easy) for the targeted range of children aged 3–5 years. After revisions based on findings from the pilot as well as a final review by local stakeholders, the eHCI was implemented nationally. Exploration of the psychometric properties of scores produced by the eHCI census in Tonga demonstrated adequate discriminant validity (comparison of mean scores across children grouped by demographic characteristics met theoretical expectations, i.e. older children scored higher than younger children, girls received slightly higher scores than boys, children of more educated mothers received higher scores than those of less educated mothers) and internal scale reliability (through Rasch analysis) [10].

Utilization of the eHCI

The eHCI has since been adapted and utilized to support a range of early childhood development projects in several low and middle income countries, predominantly across the Asia-Pacific region [12–17]. Similar to a number of tools designed to measure children's development [18–20], development and adaptations of the eHCI in each new context were informed by a combination of both theoretical conceptualization as well as local expert consultation regarding the key aspects of children's development that are predictive of future capabilities. Through these consultative processes, content and face validity of the instrument were established and adaptations and translations were ensured to be capturing the true intent of each item [21]. The internal factor structure of the instrument, however, is yet to be explored within multiple countries.

Evaluating the psychometric properties of scores produced by a measurement tool is fundamental for its future utilization and effectiveness [22]. An instrument measuring children's development that lacks in reliability and validity could produce biased scores that lead to ill-informed decisions. With eHCI data now available in multiple countries, work is needed to explore the tool's validity and reliability. An instrument with the properties of the eHCI that produces scores that are psychometrically robust and appropriate for use within diverse settings has potential for global applicability. Indeed, such a tool could better enable population monitoring of children's development, as is required for SDG 4.2, particularly in low and middle income countries, with the ultimate goal of shaping services and policy to promote global equity of children's health and development.

The current study

This research is a first step in working to establish the psychometric properties of scores produced by the eHCI within different cultures and contexts. Utilizing data previously collected from seven low and middle income countries, Brazil, China, Kiribati, Lao People's Democratic Republic (PDR), Samoa, Tonga, and Tuvalu, analyses sought to explore the internal factor structure and reliability of scores produced by the tool within each country. Findings will be used to guide recommendations regarding the reporting of eHCI results moving forward.

Methods

Measures

The early Human Capability Index

Completed by an adult who knows the child, the eHCI includes approximately 60 items (dependent upon country adaptation) measuring children's development across nine domains: Physical Health, Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, Approaches to Learning, Numeracy, Reading, and Writing. Response options to each item are binary ("yes"/"no" or "able"/"unable"). The majority of items are positively worded so that the "yes"/"able" responses are scored as 1, and the "no"/"unable" responses are scored as 0. A small number of items (ranging from 4 in Kiribati and Lao PDR to 6 in Tonga) are negatively worded and thus are reverse scored. Individual item scores in each domain are averaged so that children receive a score for each developmental domain ranging from 0 to 1, with higher scores indicative of better development.

The eHCI underwent a local adaptation process to ensure the tool's content and face validity in each country. Thus, although many items are similar across different adaptations of the eHCI, some items and domains differ across countries. To illustrate, the Perseverance domain is measured by the same 4 items across all adaptations of the eHCI. In contrast, the Physical Health domain varies from 2 items in Brazil, 3 items in Samoa and China, 4 items in Tonga, and 5 items in Tuvalu and Kiribati, while the Laotian version of the eHCI does not capture physical health as a result of local expert consultation. Table 1 presents the Tongan eHCI (in English) and the percentage of children for whom respondents reported yes/able for each item, while Additional file 1: Tables S1, Additional file 2: Table S2, Additional file 3: Table S3, Additional file 4: Table S4, Additional file 5: Table S5 and Additional file 6: Table S6 present the same information for remaining countries, highlighting similarities and differences between adapted versions of the instrument.

Demographic characteristics

In addition to measuring children's development, the eHCI collects information about children's demographic characteristics as well as relevant contextual information. Specifically, respondents provide information about children's age, gender, and special needs status, and then dependent upon country adaptation, they also provide information about children's height and weight, their mother's highest level of education, whether the child has attended preschool, if there are reading materials (i.e. books) in the child's home, and caregivers' engagement in six different types of stimulating activities with their children in the home (e.g. reading a book, playing, counting etc). Variables presented in this manuscript include children's age, gender, preschool attendance, as well as maternal education.

Data collection procedures

Data included in this manuscript were collected from seven countries between 2013 and 17, utilizing different sampling techniques and data collection methods. Contextual information regarding data collection procedures in each country are summarized in Table 2.

Participants

Characteristics of each country sample are presented in Table 3. Samples ranged in size from 549 children in Tuvalu to 12,191 in Samoa, with children ranging in age from 2 to 6 years. Though the eHCI was designed to capture the development of children aged 3–5 years, the tool has also been used to collect data on children who fall slightly outside of this age range. This is a result of varied data collection purposes across countries. For instance, in Lao PDR, 2 year olds were included in data collection as this dataset serves as the baseline measure for a randomized control trial; younger children needed to be included at baseline to ensure they also fall into midline and endline data collections in years to come. Each country sample had a relatively even split of males and females; maternal education ranged from the majority of children with mothers who had never attended school (30.2%), started (27.3%), or finished primary school in Lao PDR (29.1%), to the majority of children with mothers who had completed secondary school (42.2%) or tertiary studies in Tonga (17.8%); while the percentage of children who had attended preschool ranged from 23.2% in Lao PDR to 100.0% in Brazil.

Statistical analysis

First, confirmatory factor analyses (CFAs) were conducted separately for each country to determine the fit of eHCI data to the theoretical structure of the instrument (i.e. nine developmental domains, or eight domains in the case of Lao PDR). Next, CFAs were conducted separately for children aged 3, 4, and 5 years old in each country (as this was the age range consistent across all countries) to explore any variation in fit based on children's age. Children with missing age data were excluded from this analysis (Brazil $n = 2$, China $n = 56$, Kiribati $n = 884$, Lao PDR $n = 852$, Tonga $n = 53$, Tuvalu $n = 3$). Additional CFAs were conducted for the Lao PDR sample stratified by maternal education, to explore if fit of data to the theoretical structure of the eHCI varied by respondent's level of education. Specifically, the sample was split into two groups: low maternal education (i.e. no school, started primary, finished primary) and high maternal education (i.e. finished secondary, tertiary) and CFAs were conducted separately for each group. Children with missing maternal education data were excluded from this analysis ($n = 5$). This analysis was conducted for the Lao PDR sample only, as it was the sole sample for which data were available on the education level of all respondents.

Goodness-of-fit indices including χ^2 ($p > 0.05$ indicates good fit; Brown, 2006), root mean square error of approximation (RMSEA; values ≤0.06 indicate good model fit while those between 0.06 and 0.08 indicate satisfactory fit; Hu & Bentler, [24]; Yu, [25]), the Comparative Fit Index and the Tucker-Lewis Fit Index (CFI and TLI, respectively; values ≥0.95 indicate good fit while values between 0.90 and 0.95 indicate satisfactory fit; Hu & Bentler, [24]; Yu, [25]; Schreiber, Nora, Stage, Barlow, & King, [26]), as well as standardized factor loadings (≥ 0.4 considered high and thus deemed a good indicator of the underlying construct; Costello & Osborne, [27]) were used to evaluate model fit. CFAs were conducted in Mplus [28] utilizing polychoric correlation matrices and the weighted least squares mean and variance adjusted (WLSMV) estimation method, both of which are deemed most appropriate for use with binary-type data such as that of the eHCI [29–31], as well as oblique (geomin) factor rotation which assumes correlations amongst factors [31].

The internal reliability of eHCI domains was also examined for each country, which is often conducted in conjunction with factor analysis to measure how interrelated a set of items are and thus how well they, collectively, measure the underlying construct of focus [32]. Although the majority of similar research assesses internal reliability using Cronbach's alpha [19, 20], increasingly, ordinal reliability coefficients, specifically ordinal alpha, are deemed to be more appropriate in the case of evaluating the internal reliability of scales including items with binary response options in particular [33]. As such, ordinal alpha coefficients were calculated as well as Cronbach's alpha (≥ 0.70 deemed acceptable for both coefficients; Bland & Altman, [34]; Gadermann et al., [33]) to allow for comparison with previous research.

Appendix B

Table 1 Tongan eHCI items and n (%) children for whom their caregiver/teacher reported yes/able

Domain	Item	Yes/able	Missing
Physical Health	1. Is this child sickly or looked after poorly?[a]	826 (13.3)	5 (0.1)
	2. Does this child have good hygiene (i.e. always wash their hands after toileting)?	4929 (79.3)	13 (0.2)
	3. Does this child have positive habits?	4073 (65.5)	16 (0.3)
	4. Does this child know good foods from bad foods?	3710 (59.7)	25 (0.4)
Verbal Communication	1. Can this child use a group of words in talking?	5705 (91.8)	8 (0.1)
	2. Can this child converse with others?	5988 (96.3)	8 (0.1)
	3. Can this child talk about something that he/she has done?	5619 (90.4)	10 (0.2)
	4. Can this child give detail with good Tongan words?	3844 (61.9)	10 (0.2)
	5. Can this child hold an adult like conversation (for example talkative, always questioning)?	5078 (81.7)	16 (0.3)
Cultural Knowledge	1. Shows compassion, understanding and tolerance of others?	5229 (84.1)	15 (0.2)
	2. Can this child identify two culturally important foods/dishes?	5043 (81.1)	11 (0.2)
	3. Can this child identify two local plants that provide food/fruits?	4741 (76.3)	14 (0.2)
	4. Does this child show the Tongan cultural values of humility?	2892 (46.5)	21 (0.3)
	5. Does this child show the Tongan cultural values of devotion/ commitment/obligation/responsibility?	2745 (44.2)	16 (0.3)
	6. Does this child show the Tongan cultural values of reciprocity in relationships?	2849 (45.8)	14 (0.2)
	7. Does this child participate in cultural routines?	4764 (76.7)	11 (0.2)
	8. Is this child able to say a short prayer?	5503 (88.5)	16 (0.3)
Social and Emotional	1. Is this child happy to share their toys and belongings?	5296 (85.2)	9 (0.1)
	2. Does this child take care of their own things?	4579 (73.7)	4 (0.1)
	3. Does this child demonstrate respect for adults?	3870 (62.3)	11 (0.2)
	4. Does this child demonstrate respect for other children?	3945 (63.5)	11 (0.2)
	5. Does this child accept responsibility for their actions?	4264 (68.6)	9 (0.1)
	6. Is this child considerate of other people's feelings?	4274 (68.8)	6 (0.1)
	7. Does this child repeatedly do something wrong even though he/she has been told to stop?[a]	3612 (58.1)	9 (0.1)
	8. Is this child always helpful?	5445 (87.6)	6 (0.1)
	9. Is this child friendly to other children?	5644 (90.8)	9 (0.1)
	10. Does this child kick, bite or hit adults or other children?[a]	2294 (36.9)	12 (0.2)
	11. Is this child impatient?[a]	3868 (62.2)	11 (0.2)
	12. Does this child always understand the difference between acceptable and unacceptable behaviour?	4522 (72.8)	8 (0.1)
	13. Does this child follow simple directions on how to do something?	5210 (83.8)	6 (0.1)
Perseverance	1. Does this child always perform tasks independently?	4588 (73.8)	8 (0.1)
	2. Does this child always keep at a task until they are finished?	3348 (53.9)	9 (0.1)
	3. Does this child need constant reminding to finish something off?[a]	4468 (71.9)	10 (0.2)
	4. Does this child get easily distracted from a task?[a]	4664 (75.0)	14 (0.2)
Approaches to Learning	1. Does this child show more curiosity about something new in comparison to something familiar?	4976 (80.1)	10 (0.2)
	2. Does this child investigate/explore the function of a new toy/game/puzzle or object?	4900 (78.8)	9 (0.1)
	3. Is this child always wanting to learn new things?	4850 (78.0)	12 (0.2)
	4. When in an unfamiliar environment with a familiar person present, does this child feel free to explore?	3611 (58.1)	15 (0.2)

Table 1 Tongan eHCI items and n (%) children for whom their caregiver/teacher reported yes/able *(Continued)*

Domain	Item	Yes/able	Missing
	5. Is this child diligent in their approach to a new job or task?	3840 (61.8)	19 (0.3)
Numeracy	1. Can this child recognise geometric shapes (e.g. triangle, circle, square)?	3322 (53.5)	13 (0.2)
	2. Can this child name and identify at least 3 colours?	4278 (68.8)	8 (0.1)
	3. Can this child sort and classify objects by common characteristics (e.g. shape, colour, size)?	3218 (51.8)	12 (0.2)
	4. Can this child name and recognise the symbol of all numbers from 1 to 10?	3020 (48.6)	15 (0.2)
	5. Can this child count to 10?	5053 (81.3)	7 (0.1)
	6. Can this child count to 20?	2073 (33.4)	9 (0.1)
	7. Can this child count to 100?	430 (6.9)	13 (0.2)
	8. Does this child know that a horse is taller than a dog?	4464 (71.8)	11 (0.2)
	9. Does this child know the order of the day (e.g. morning, then afternoon and then evening)?	2411 (38.8)	17 (0.3)
	10. Does this child understand the concepts of yesterday, today and tomorrow?	1812 (29.2)	14 (0.2)
	11. Does this child know that a vehicle weighs more than a cup?	4197 (67.5)	11 (0.2)
	12. Does this child know that the number 8 is bigger than the number 2?	2258 (36.3)	16 (0.3)
Reading	1. Does this child know the sounds of three letters of the alphabet? (phonics)	3932 (63.5)	16 (0.3)
	2. Can this child identify at least 3 letters of the alphabet?	3263 (52.5)	13 (0.2)
	3. Can this child identify at least 10 letters of the alphabet?	2232 (35.9)	14 (0.2)
	4. Can this child hold a book and turn the pages in the right way?	3811 (61.3)	10 (0.2)
	5. Can this child follow reading directions (i.e. left to right, top to bottom)?	1799 (28.9)	9 (0.1)
	6. Can this child read at least 4 popular words?	2263 (36.4)	13 (0.2)
Writing	1. Can this child draw something identifiable? (e.g. a stick person)	3531 (56.8)	9 (0.1)
	2. Can this child copy (trace) the shape of a letter (e.g. A, E, F)?	3392 (54.6)	5 (0.1)
	3. Can this child write their own name?	2428 (39.1)	9 (0.1)
	4. Can this child write short and simple words?	1452 (23.4)	9 (0.1)
	6. Can this child write short and simple sentences?	778 (12.5)	9 (0.1)

Note: *reverse scored items

These analyses were conducted using the package 'psych' in R-Studio [35].

Results

Model fit indices yielded from a CFA for each country are presented in Table 4. Although a statistically significant χ^2 indicates poor model fit, this test is sensitive to sample size and thus large samples are likely to yield significant χ^2 values [24, 36]. Considering country samples were large, we instead rely more heavily on other fit indices. While χ^2 values were statistically significant in all countries, RMSEA values indicated good model fit (i.e. ≤ 0.06) in each country with the exception of Samoa, for which the RMSEA value (0.07) indicated satisfactory fit. CFI and TLI values were ≥ 0.90 in all countries, again indicating satisfactory fit of eHCI data to the theoretical structure of the instrument (i.e. 9 developmental domains, or 8 domains in Lao PDR). Table 4 also presents model fit indices from CFAs conducted separately for 3, 4, and 5-year-old children, demonstrating relatively inconsistent results across countries. For instance, although model fit indices indicated better fit of data to the theoretical structure of the eHCI for 5-year-old children in Brazil and Samoa, slightly better fit was observed for 3-year-olds in Tonga and Tuvalu, relative to that of older children.

Standardized factor loadings yielded from full sample CFAs are presented in Table 5 for Tonga, and Additional file 7: Tables S7, Additional file 8: Table S8, Additional file 9: Table S9, Additional file 10: Table S10, Additional file 11: Table S11 and Additional file 12: Table S12 for remaining countries (factor loadings yielded from CFAs for 3, 4, and 5-year-old children are available from authors upon request). Items had factor loadings ≥0.40 across domains and countries with few exceptions. Items that form Numeracy, Reading, and Writing domains in particular had high factor loadings (≥ 0.80 on average) consistently across countries with

Table 2 Data collection contexts and procedures

	Country context	Year/s	Respondent/s	Method	Data collection purpose	Sample
Brazil	Occupies a large area on the eastern coast of South America. In 2017, population was approx. 209,000 and GNI was USD8600 per capita.	2015	Preschool teachers	Pen and paper	Adaptation of the eHCI for the Brazilian context [16]	Children from one city in Southwest Brazil
China	In East Asia, the world's most populous country. In 2017, population was approx. 1.4 billion and GNI was USD8690 per capita.	2015/16	Preschool teachers and caregivers	Pen and paper	Evaluate inequality in children's outcomes across population groups [17]	Children from two provinces across Northern China
Kiribati	Comprised of 33 coral atolls and isles in the Central Pacific. In 2017, population was approx. 116,000 and GNI was USD3010 per capita.	2017	Preschool teachers and caregivers	Tablet	National baseline of child development to guide policy and programs [14]	Population; aimed to collect data for all children aged 3–5 years nationally
Lao PDR	A landlocked country in the Southeast of Asia. In 2017, population was approx. 6.9 million and GNI was USD2270 per capita.	2015/16	Caregivers	Tablet	Baseline data for an RCT designed to support children's early development [12]	Children from five provinces across Northern Lao PDR
Samoa	Comprised of 2 main islands in the South Pacific. In 2017, population was approx. 196,000 and GNI was USD4090 per capita.	2016	Preschool teachers and caregivers	Tablet	National baseline of child development to guide policy and programs [13]	Population; aimed to collect data for all children aged 3–5 years nationally
Tonga	An archipelago of more than 170 South Pacific islands (36 inhabited). In 2017, population was approx. 108,000 and GNI was USD4010 per capita.	2013/14	Preschool teachers and caregivers	Pen and paper	Baseline data for an RCT designed to support children's school readiness [10]	Population; aimed to collect data for all children aged 3–5 years nationally
Tuvalu	An island country in the South Pacific, comprising 9 small islands. In 2017, population was approx. 11,000 and GNI was USD4970 per capita.	2015	Preschool teachers and caregivers	Pen and paper	National baseline of child development to guide policy and programs [15]	Population; aimed to collect data for all children aged 3–5 years nationally

Note. *GNI* Gross National Income, *RCT* Randomized Control Trial. Population and Gross National Income figures sourced from World Bank [23]

few exceptions, while the strength of factor loadings for remaining developmental domains tended to be more varied across countries. For example, factor loadings for items in the Verbal Communication domain in China ranged from 0.52–0.90, while those in Lao PDR ranged from 0.73–0.94. Similarly, factor loadings for items in the Cultural Knowledge domain in Lao PDR ranged from 0.58–0.83, while those in Brazil ranged from 0.76–0.99. Reverse-scored items in Physical Health, Social and Emotional Skills, and Perseverance domains had weak factor loadings in all countries but Brazil. In contrast, Brazil was the only country in which some non-reverse-scored items had weak factor loadings. Specifically, Reading item 6 for which just under 2% of children were reported to be able to read complex sentences had a factor loading of 0.36, and Writing item 1 for which all but just under 2% of children were reported to be able to scribble on paper had a factor loading of 0.21.

Table 6 presents model fit indices for CFAs for low versus high maternal education in Lao PDR, with standardized factor loadings presented in Additional file 13: Tables S13 and Additional file 14: Table S14. RMSEA, CFI and TLI values indicated better fit of eHCI data to the theoretical structure of the instrument (i.e. 8 domains in Lao PDR) when respondents had a higher level of education. Factor loadings for reverse-scored items in Social and Emotional Skills and Perseverance domains, however, were weak across both education groups.

Finally, Table 7 presents internal consistency coefficients for eHCI domains in each country, demonstrating varied results across domains. The Numeracy domain had consistently high internal reliability across countries, with ordinal $\alpha \geq .91$, which is considered high [33]. Verbal Communication (ordinal $\alpha \geq 0.87$), Cultural Knowledge (ordinal $\alpha \geq 0.79$), Social and Emotional Skills (ordinal $\alpha \geq 0.70$), Approaches to Learning (ordinal $\alpha \geq 0.86$), Reading (ordinal $\alpha \geq 0.87$), and Writing (ordinal $\alpha \geq 0.78$) domains also yielded internal consistency coefficients deemed to be acceptable across countries. In contrast, the remaining two domains, Physical Health and Perseverance, demonstrated less than satisfactory internal reliability with ordinal $\alpha < 0.70$ in all countries with the exception of Tuvalu and Kiribati on the Physical Health domain (ordinal α = 0.77 and 0.76, respectively) and Tuvalu (ordinal $\alpha = 0.75$) and Brazil (ordinal α = 0.75) on the Perseverance domain.

Discussion

The current study presents the psychometric properties of scores produced by the eHCI in seven low and middle income countries. Results demonstrated adequate fit of eHCI data to the theoretical structure of the instrument measuring children's development across 9 domains (or 8 domains in the case of Lao PDR). Overall, findings lend support to the aims of the eHCI in being adaptable and applicable for use within a range of low and middle

Table 3 Sample descriptive characteristics

	Brazil n (%)	China n (%)	Kiribati n (%)	Lao PDR n (%)	Samoa n (%)	Tonga n (%)	Tuvalu n (%)
Child gender							
Male	853 (47.1)	5587 (48.9)	4269 (51.2)	3824 (51.0)	6293 (51.6)	3247 (52.3)	272 (49.5)
Female	855 (47.2)	5338 (46.7)	3915 (46.9)	3669 (49.0)	5898 (48.4)	2967 (47.7)	277 (50.5)
Missing	102 (5.6)	496 (4.3)	155 (1.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Child age							
2 years	57 (3.1)	54 (0.5)	948 (11.4)	1410 (18.8)	1159 (9.5)	13 (0.2)	1 (0.2)
3 years	571 (31.5)	3396 (29.7)	2185 (26.2)	1749 (23.3)	4377 (35.9)	1609 (25.9)	163 (29.7)
4 years	760 (42.0)	3329 (29.1)	2136 (25.6)	1867 (24.9)	4616 (37.9)	2058 (33.1)	180 (32.8)
5 years	420 (23.2)	3360 (29.4)	2013 (24.1)	1599 (21.3)	2089 (16.7)	2038 (32.8)	195 (35.5)
6 years	0 (0.0)	1226 (10.7)	173 (2.1)	16 (0.2)	0 (0.0)	443 (7.1)	7 (1.3)
Missing	2 (0.1)	56 (0.5)	884 (10.6)	852 (11.4)	0 (0.0)	53 (0.9)	3 (0.5)
Child preschool attendance							
Yes	1810 (100.0)	9159 (80.2)	7665 (91.9)	1738 (23.2)	4657 (38.2)	2701 (43.5)	498 (90.7)
No	0 (0.0)	2176 (19.1)	674 (8.1)	5755 (76.8)	7534 (61.8)	3513 (56.5)	51 (9.3)
Missing	0 (0.0)	86 (0.8)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Maternal Education							
No school	–	307 (2.7)	–	2265 (30.2)	–	–	–
Started Primary	–	1242 (10.9)	714 (8.6)	2045 (27.3)	–	–	–
Finished Primary	–	3372 (29.5)	1438 (17.2)	2182 (29.1)	222 (1.8)	81 (1.3)	47 (8.6)
Started Secondary	–	3563 (31.2)	2319 (27.8)	–	603 (4.9)	2399 (38.6)	89 (16.2)
Finished Secondary	–	1358 (11.9)	2710 (32.5)	754 (10.1)	10,037 (82.3)	2621 (42.2)	107 (19.5)
Tertiary	–	751 (6.6)	785 (9.4)	242 (3.2)	1329 (10.9)	1107 (17.8)	117 (21.3)
Missing	–	828 (7.2)	373 (4.5)	5 (0.1)	0 (0.0)	0 (0.0)	189 (34.4)
Total n	1810	11,421	8339	7498	12,191	6214	549

Note. Maternal education data were not collected in Brazil. In Samoa, Tonga, and Tuvalu, when responding to the maternal education item, respondents could select only one response option pertaining to primary school and so the proportions represented against the 'finished primary school' category for these countries may include a combination of children for whom their mother either started or finished primary school

income countries to facilitate measurement of children's early development [10].

Psychometric findings

Samples utilized in this research differed considerably across countries in terms of children's demographic backgrounds, data collection methods and purposes, as well as sampling techniques and sizes. Although it might be argued that such differences present a challenge to exploring and comparing the validity and reliability of the eHCI within multiple contexts, this is the pragmatic nature in which the instrument was intended to be used; for a range of purposes and across varied contexts. As Yapa and Bärnighausen [37] discuss, the resource constraints that come with research in low and middle income countries are often the driving force behind creative solutions. As such, we argue that there is strength in that the eHCI was found to demonstrate a common underlying factor structure within the varied contexts in which the instrument has been implemented.

Numeracy, Reading, and Writing domains in particular were found to be working consistently across countries. Items that form these domains had high factor loadings and these scales had high internal reliability across countries. Similar results have been reported for other measures of children's development, for instance, factor analyses of domains that constitute the EDI, a teacher-completed checklist measuring children's holistic development in their first year of school, have demonstrated the Language and Cognitive Development domain (which captures children's literacy and numeracy skills) to have the best fit across multiple countries [38]. Examination of items that form Numeracy, Reading, and Writing domains in the eHCI highlight that little adaptation of these items was required across countries. For instance, the Numeracy domain covers the same concepts of shape, colour, and number recognition, counting ability, and knowledge of numerical concepts such as time and weight, across countries. This might suggest these domains to be the more universal aspects of children's development, indeed such

Table 4 Model fit indices from confirmatory factor analyses in each country

		X^2 (df)	RMSEA (90% CI)	CFI	TLI
Brazil	Full sample	4595.48 (1503)	0.034 (0.033–0.035)	0.94	0.94
	3-year-olds	3772.55 (1394)	0.055 (0.053–0.057)	0.75	0.74
	4-year-olds	4413.833 (1503)	0.050 (0.049–0.052)	0.82	0.80
	5-year-olds	2157.71 (1448)	0.034 (0.031–0.037)	0.92	0.92
China	Full sample	35,272.80 (1616)	0.043 (0.042–0.043)	0.92	0.92
	3-year-olds	10,356.63 (1616)	0.040 (0.039–0.041)	0.94	0.93
	4-year-olds	12,039.11 (1616)	0.044 (0.043–0.045)	0.87	0.86
	5-year-olds	9892.98 (1616)	0.039 (0.038–0.040)	0.84	0.83
Kiribati	Full sample	28,191.32 (1979)	0.040 (0.039–0.040)	0.93	0.93
	3-year-olds	8200.71 (1979)	0.038 (0.037–0.039)	0.92	0.92
	4-year-olds	7057.17 (1682)	0.039 (0.038–0.040)	0.91	0.91
	5-year-olds	6245.19 (1979)	0.033 (0.032–0.034)	0.91	0.91
Lao PDR	Full sample	22,116.62 (1456)	0.044 (0.043–0.044)	0.90	0.90
	3-year-olds	4744.53 (1456)	0.036 (0.035–0.037)	0.87	0.86
	4-year-olds	5528.15 (1456)	0.039 (0.038–0.040)	0.86	0.85
	5-year-olds	5379.22 (1456)	0.041 (0.040–0.042)	0.89	0.89
Samoa	Full sample	89,517.11 (1674)	0.066 (0.065–0.066)	0.94	0.93
	3-year-olds	31,146.12 (1616)	0.065 (0.064–0.065)	0.93	0.93
	4-year-olds	34,712.58 (1674)	0.065 (0.065–0.066)	0.94	0.93
	5-year-olds	14,613.05 (1674)	0.062 (0.061–0.062)	0.94	0.94
Tonga	Full sample	23,936.63 (1793)	0.045 (0.044–0.045)	0.93	0.93
	3-year-olds	6172.07 (1793)	0.039 (0.038–0.040)	0.91	0.91
	4-year-olds	7724.07 (1793)	0.040 (0.039–0.041)	0.92	0.92
	5-year-olds	7839.75 (1793)	0.041 (0.040–0.042)	0.90	0.90
Tuvalu	Full sample	4144.74 (2043)	0.043 (0.041–0.045)	0.95	0.94
	3-year-olds	2540.33 (2043)	0.039 (0.033–0.043)	0.95	0.95
	'4-year-olds	2483.74 (1854)	0.043 (0.039–0.048)	0.89	0.88
	5-year-olds	2600.80 (2043)	0.037 (0.033–0.042)	0.89	0.89

Note. $p < .001$ for all models, *df* degrees of freedom, *CI* Confidence interval. Differing *df* across CFAs within a country highlight instances where items were dropped from the model to enable estimation of model parameters. In Brazil, read6 and write6 were dropped from the 3yo CFA, and write1 was dropped from the 5yo CFA. In Kiribati, all phys items were dropped from the 4yo CFA. In Samoa, phys2 was dropped from the 3yo CFA. In Tuvalu, phys3, phys4, and phys5 were dropped from the 4yo CFA

skills have been demonstrated to be important predictors of outcomes throughout the life course [39, 40], and thus results are consistently strong across countries. Such skills are also arguably more easily observable (i.e. it is likely that a caregiver or teacher knows if a child can read or count, as opposed to whether they know if a child is always wanting to learn new things as measured in the Approaches to Learning domain, or if a child knows good from bad foods as captured in the Physical Health domain), which may also have had an influence on results.

In contrast, results across Physical Health, Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, and Approaches to Learning domains tended to demonstrate more variation across country samples. This is unsurprising considering the nature of the skills measured by these domains, including health and hygiene practices, verbal communication abilities, knowledge of culture and culturally acceptable behaviours, social interactions and emotional regulation, as well as how tasks are approached and the ability to complete them, which are aspects of development that would be considered to be more contextually and culturally specific. To illustrate, although Social and Emotional Skills was demonstrated to be one of the distinct domains that the eHCI captures within each country, variation in the strength of factor loadings of items that form this domain between Brazil and Lao PDR might be, in part, explained by cultural differences in social interactions and the expression of emotions between the two countries. Item factor loadings were lower in Lao PDR (on average around 0.65–0.70), an individualistic culture in which emotion is perceived to be experienced internally within an individual, whereas the same item loadings were higher in Brazil (on average around 0.80–0.85), a collectivist culture whereby emotions are thought to occur between people and thus are expressed openly [41]. These results could also be attributed to variation in methodological bias across countries [42]. For instance, acquiescence, the tendency to agree with statements, has been demonstrated to be more common in collectivist cultures [43]. Variation of results across countries in this way highlights the important influence of culture on the measurement of children's development, and the need for tools to capture not only the aspects of children's development that are important predictors of later outcomes, but to also be aligned with local culture in order to produce information that both accurately reflects children's abilities and is relevant to local policy and practice [1, 8].

Reverse-scored items in the Physical Health, Social and Emotional, and Perseverance domains had weak factor loadings in all countries but Brazil, indicating that, compared to other items, they are poorer measures of the underlying constructs being measured by these domains. These results reflect initial findings from analyses conducted in Tonga following the development of the tool, with Rasch analyses indicating that a number of reverse-scored items did not fit the model well relative to other items [10]. It is possible that enumerators and/or respondents had difficulty in understanding and/or responding to these negatively worded items. Previous research has shown that reverse-scored items tend to load onto a factor separate to the construct they are intended to measure, that instead reflects aspects of item method [44, 45]. This was not observed in Brazil however, and this could be due to sample differences in this context relative to other countries. Specifically, Brazil was the only sample for which only preschool teachers completed the eHCI, as opposed to a

Table 5 Factor loadings from confirmatory factor analysis in Tonga

	F1	F2	F3	F4	F5	F6	F7	F8	F9
Phys 1	0.10								
Phys 2	-0.73								
Phys 3	-0.69								
Phys 4	-0.82								
Comm 1		0.81							
Comm 2		0.91							
Comm 3		0.96							
Comm 4		0.85							
Comm 5		0.65							
Cult 1			0.81						
Cult 2			0.88						
Cult 3			0.88						
Cult 4			0.88						
Cult 5			0.88						
Cult 6			0.88						
Cult 7			0.66						
Cult 8			0.73						
Soc 1				0.65					
Soc 2				0.70					
Soc 3				0.88					
Soc 4				0.88					
Soc 5				0.61					
Soc 6				-0.10					
Soc 7				-0.73					
Soc 8				0.69					
Soc 9				0.65					
Soc 10				0.05					
Soc 11				-0.14					
Soc 12				0.81					
Soc 13				0.82					
Persev 1					0.76				
Persev 2					0.77				
Persev 3					-0.11				
Persev 4					-0.46				
Appr 1						0.82			
Appr 2						0.90			
Appr 3						0.90			
Appr 4						0.68			
Appr 5						0.93			
Num 1							0.86		
Num 2							0.85		
Num 3							0.86		
Num 4							0.86		
Num 5							0.79		
Num 6							0.79		
Num 7							0.76		
Num 8							0.84		
Num 9							0.85		
Num 10							0.85		
Num 11							0.84		
Num 12							0.87		
Read 1								0.77	
Read 2								0.93	
Read 3								0.92	
Read 4								0.78	
Read 5								0.89	
Read 6								0.75	
Writ 1									0.86
Writ 2									0.94
Writ 3									0.93
Writ 4									0.96
Writ 5									0.94

Table 5 Factor loadings from confirmatory factor analysis in Tonga *(Continued)*

Note. *Phys* Physical Health, *Comm* Verbal Communication, *Cult* Cultural Knowledge, *Soc* Social and Emotional, *Persev* Perseverance, *Appr* Approaches to Learning, *Num* Numeracy, *Read* Reading, and *Writ* Writing

combination of children's caregivers and teachers, or caregivers only as in other samples. It might be that children's caregivers and teachers do not respond to the eHCI in the same way, or that a minimum level of education or literacy is required to understand and respond to items. Indeed, results demonstrated better fit of data to the theoretical structure of the eHCI in Lao PDR amongst more educated caregivers who responded to the tool, compared to those less-educated. However, weak factor loadings for reverse-scored items were maintained when analyses were run separately for caregivers who had low versus high education. Together, results raise important questions regarding respondent reliability that need to be explored by future research.

When considering children's age in determining how the instrument is operating, results were inconsistent

Table 6 Model fit indices from confirmatory factor analyses in Lao PDR grouped by maternal education

	X^2 (df)	RMSEA (90% CI)	CFI	TLI
Low maternal education	19188.69 (1456)	0.043 (0.043–0.044)	0.90	0.89
High maternal education	3222.03 (1456)	0.035 (0.033–0.037)	0.94	0.93

Note. $p < .001$ for all models, *df* degrees of freedom, *CI* Confidence interval

Table 7 Internal reliability of eHCI domains

	Brazil		China		Kiribati		Lao PDR		Samoa		Tonga		Tuvalu	
	Cα	Oα	Cα	Oα	Cα	Oα	Cα	Oα	Cα	Oα	Cα	Oα	Cα	Oα
Physical Health	.20	.52	.25	.43	.61	.76	–	–	.34	.48	.47	.56	.61	.77
Verbal Communication	.67	.93	.67	.87	.78	.91	.70	.88	.78	.90	.69	.92	.73	.93
Cultural Knowledge	.52	.89	.80	.90	.84	.93	.53	.79	.89	.95	.79	.89	.82	.92
Social and Emotional	.85	.95	.71	.84	.83	.91	.73	.84	.85	.90	.50	.70	.80	.91
Perseverance	.75	.89	.48	.62	.26	.33	.15	.17	.38	.45	.23	.31	.60	.75
Approaches to Learning	.71	.91	.70	.86	.67	.86	.74	.88	.85	.94	.80	.91	.73	.90
Numeracy	.84	.94	.87	.95	.86	.94	.80	.91	.92	.97	.89	.96	.90	.97
Reading	.61	.87	.79	.90	.79	.91	.72	.91	.87	.95	.82	.92	.82	.93
Writing	.78	.87	.59	.78	.83	.94	.65	.93	.89	.96	.83	.96	.83	.94

Note. *Cα* Cronbach's alpha, *Oα* Ordinal alpha

across countries (i.e. in some countries best fit was observed for 3-year-olds and in others for 5-year-olds), and with the exception of Brazil, model fit indices did not vary in magnitude greatly, indicating that the eHCI appears to work relatively consistently across the age range of 3–5 years. Although these results provide some insight into how effective the eHCI is in measuring development across children of different ages, analyses focused on the discriminant validity of the tool, including if items capture a continuum of development for children of different ages, are needed to further explore this question.

Finally, internal reliability results for Physical Health and Perseverance domains indicated that items that form these domains, collectively, are not a good measure of these underlying constructs. It is possible that this is a result of reverse-scored items (which could be measuring constructs separate to that intended) making up a large proportion of these domains (i.e. one in two items in the Perseverance domain are reverse-scored across countries, while for the Physical Health domain this is the case for one in two items in Brazil, one in three items in Samoa and China, one in four items in Tonga, and one in five items in Tuvalu and Kiribati). Nevertheless, local adaptation of the tool in each country deemed all items important to children's early development in their contexts, and thus it was not the intention of the current research to exclude items on the basis of psychometric results. An example is the first item in the Physical Health domain regarding children being frequently sick. Although we would not naturally assume that this item measures a child's skills or capabilities (and subsequently it does not work well in the model), in the contexts of countries of focus whereby illness is common, it was deemed important for the eHCI to provide information regarding children's experience of illness as one aspect of their holistic development.

Study limitations

When interpreting results of the study, it is important to be cognizant of three limitations. First, although the majority of countries studied utilized a census approach to data collection and thus are considered nationally representative, sampling strategies employed in Brazil, China, and Lao PDR (see Table 2) posit that results may differ if eHCI data were to be collected on nationally representative samples in these countries. Next, results indicate that reverse-scored items may not be operating as intended. Beyond analyses presented in this study, however, the information required to be able to explore this further (for example, insight into respondents' understanding of reverse-scored items) is not currently available. Finally, it is important to reiterate that demonstrating consistent internal factor structure and reliability, as has been done in this study, is not complete evidence of a valid tool and must be considered together with results from additional psychometric analyses.

Study implications

Relative to other measures of early childhood development currently utilized, the eHCI requires minimal resources to be implemented. Initial psychometric results suggest that this has not come at the cost of the validity and reliability of the instrument. Demonstrating a consistent internal factor structure and reliability is one important aspect of the comprehensive evaluation of an instrument's validity and reliability. Although not within the scope of the current study, additional work is underway to explore the extent to which eHCI domains can discriminate amongst children's abilities by a range of demographic and contextual variables, are associated with scores on other measures of child development, show reliability amongst respondents, and are able to predict children's future outcomes. A low-burden instrument that is both easily adaptable and psychometrically robust within multiple contexts in this way has potential utility internationally. Indeed, such a tool might better enable population monitoring of children's development, as is required for the tracking of progress towards the Sustainable Development Agenda, particularly in low and middle income countries.

In terms of the reporting of and utilization of data produced by the eHCI, results suggest that eHCI data should continue to be reported across the instrument's 9 theoretically-based developmental domains, or 8 domains in the case of Lao PDR. Reporting of children's development across different areas of development in this way enables the identification of areas of both strength and need, and as a result can help to shape more targeted approaches to intervention or policy development. SDG 4.2 however, calls for the monitoring of children who are "developmentally on track", a concept that, as with children's development more broadly, is likely to vary across contexts. As such, if the eHCI is to be recommended to track progress against SDG 4.2 in future, research needs to not only work to further validate the instrument, but also determine how "developmentally on track" might be classified utilizing eHCI data.

Conclusion

Initial psychometric results demonstrate that scores produced by the eHCI, after processes of local adaptation, translation and implementation, maintained a similar factor structure of 9 theoretically-based developmental domains (or 8 domains in the case of Lao PDR) within a range of low and middle income countries. Future research is needed to build on these results and help to determine if the eHCI is able to fulfil its purpose of being a reliable, valid, and feasible tool which can help to facilitate the evaluation of early childhood policies and programs as well as measurement and monitoring of children's development in the early years, particularly in low and middle income countries.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12887-019-1852-5.

Additional file 1: Table S1. Brazil eHCI items and n (%) children for whom their caregiver/ teacher reported yes/able.

Additional file 2: Table S2. China eHCI items and n (%) children for whom their caregiver/ teacher reported yes/able.

Additional file 3: Table S3. Kiribati eHCI items and n (%) children for whom their caregiver/ teacher reported yes/able.

Additional file 4: Table S4. Lao PDR eHCI items and n (%) children for whom their caregiver/ teacher reported yes/able.

Additional file 5: Table S5. Samoa eHCI items and n (%) children for whom their caregiver/ teacher reported yes/able.

Additional file 6: Table S6. Tuvalu eHCI items and n (%) children for whom their caregiver/ teacher reported yes/able.

Additional file 7: Table S7. Factor loadings from confirmatory factor analysis in Brazil.

Additional file 8: Table S8. Factor loadings from confirmatory factor analysis in China.

Additional file 9: Table S9. Factor loadings from confirmatory factor analysis in Kiribati.

Additional file 10: Table S10. Factor loadings from confirmatory factor analysis in Lao PDR.

Additional file 11: Table S11. Factor loadings from confirmatory factor analysis in Samoa.

Additional file 12: Table S12. Factor loadings from confirmatory factor analysis in Tuvalu.

Additional file 13: Table S13. Factor loadings from confirmatory factor analysis in Lao PDR – low maternal education.

Additional file 14: Table S14. Factor loadings from confirmatory factor analysis in Lao PDR – high maternal education.

Abbreviations

CFA: Confirmatory Factor Analysis; CFI: Comparative Fit Index; CREDI: Caregiver Reported Early Development Instrument; ECDI: Early Childhood Development Instrument; EDI: Early Development Instrument; EHCI: early Human Capability Index; IDELA: International Development and Early Learning Assessment; MDAT: Malawi Developmental Assessment Tool; MODEL: Measurement of Development and Early Learning; PDR: People's Democratic Republic; PRIDI: Regional Project on Child Development Indicators; RMSEA: Root Mean Square Error of Approximation; SDG: Sustainable Development Goal; TLI: Tucker-Lewis Fit Index; WLSMV: Weighted Least Squares Mean and Variance

Acknowledgements

We wish to acknowledge the agencies who supported the collection of eHCI data which were utilized in the current study. Data collection was funded by the Fundação Maria Cecilia Souto Vidigal in Brazil, the China Development Research Foundation in China, the World Bank Group through the Early Childhood Education Project in Lao PDR, and the Global Partnership for Education through the Pacific Early Age Readiness and Learning (PEARL) Program in Kiribati, Samoa, Tonga, and Tuvalu.

Authors' contributions

SAB developed the instrument and led adaptations within each country. CZ and DDS contributed to the acquisition of data. AS led the analysis and writing of the manuscript. SAB and TG contributed to the analysis plan and supported development of the manuscript. AS, TG, CZ, DDS, JL, and SAB contributed to the interpretation of results, revised the manuscript, and approved the final version of the manuscript for submission.

Funding

Not applicable.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author (AS) on reasonable request.

Ethics approval and consent to participate

The current study is a secondary analysis of pre-existing, de-identified data and thus was deemed exempt from requiring ethical review by the University of Adelaide Human Research Ethics Committee. Consent to participate was obtained from respondents verbally prior to data collection. Verbal consent was sought as this was the method through which all data were collected from respondents (i.e. enumerators asked respondents a series of questions verbally), and this was aligned with the various ethical and research approvals in each country.

Consent for publication

Not applicable.

Competing interests
AS and SAB are employed by the World Bank Group outside of the submitted research. The authors declare no other no other competing interests.

Author details
[1]School of Public Health, University of Adelaide, Level 9, Adelaide Health and Medical Science Building, 57 North Terrace, Adelaide, South Australia 5005, Australia. [2]Telethon Kids Institute, University of Western Australia, Level 15, 31 Flinders St, Adelaide, South Australia 5000, Australia. [3]Department of Developmental Psychology and Personality, Universidad Federal do Rio Grande do Sul, Av. Paulo Gama, 110, Bairro Farroupilha, Porto Alegre, Rio Grande do Sul 90040-060, Brazil. [4]Department of Economy, University of São Paulo, Av. Prof. Luciano Gualberto, 908, Butantã, São Paulo 05508-010, Brazil. [5]Population Health Sciences, Bristol Medical School, University of Bristol, First Floor, 5 Tyndall Avenue, Bristol BS8 1UD, UK.

Received: 30 September 2019 Accepted: 22 November 2019
Published online: 03 December 2019

References

1. Raikes A, Yoshikawa H, Britto PR, Iruka I. Children, youth and developmental science in the 2015–2030 global sustainable development goals. Soc Res Child Dev Soc Policy Rep. 2017;30(3):1–23.
2. Britto PR, Lye SJ, Proulx K, Yousafzai AK, Matthews SG, Vaivada T, et al. Nurturing care: promoting early childhood development. Lancet. 2017;389:91–102.
3. Fernald L, Prado E, Kariger P, Raikes A. A toolkit for measuring early childhood development in low- and middle-income countries. Washington DC: World Bank Group; 2017.
4. United Nations General Assembly. Report of the open working Group of the General Assembly on sustainable development goals. New York City: UNGA; 2014. a/68/970.
5. World Bank. Atlas of sustainable development goals 2018: world development indicators. Washington, DC: World Bank; 2018.
6. Britto PR. Early moments matter for every child. New York: UNICEF; 2017.
7. Hambleton RK, Merenda PF, Spielberger CD. Adapting educational and psychological tests for cross cultural assessment. Mahwah: Lawrence Erlbaum Associates Inc.; 2005.
8. Keller H, Kärtner J. The cultural solution of universal developmental tasks. Gelfand M, Chiu C, Hong Y, editors. New York: Oxford University Press; 2013.
9. Rao N, Sun J, Becher Y. Assessing early development and learning across cultures: the East Asia Pacific-early child development scales. Assess Dev Matters. 2015;7(2):21–5.
10. Brinkman S, Thanh VB. Early Child Development in Tonga. Washington, DC: World Bank; 2017. Available from: http://documents.worldbank.org/curated/en/898671481187892018/pdf/110821-PUB-PUBLIC-pubdate-12-7-2016.pdf
11. MacDonald K, Brinkman S, Jarvie W, Machuca-Sierra M, McDonall K, Messaoud-Galusi S, et al. Pedagogy versus School Readiness: The Impact of a Randomized Reading Instruction Intervention and Community-Based Playgroup Intervention on Early Grade Reading Outcomes in Tonga. WPS 7944. Washington, D.C.: World Bank Group; 2017.
12. Brinkman S, Sincovich A, Danchev P. The Status of Early Childhood Health and Development in Northern Lao PDR Baseline Results from the Early Childhood Education Project. Washington, DC: World Bank Group; 2016. Available from: http://documents.worldbank.org/curated/en/228221488446175113/The-status-of-early-childhood-health-and-development-in-northern-Lao-PDR
13. Brinkman S, Sincovich A, Than VB. Early Childhood Development in Samoa: Baseline results from the Samoan Early Human Capability Index. Washington, D.C: World Bank Group; 2017. Available from: http://documents.worldbank.org/curated/en/363761563248716398/pdf/Main-Report.pdf
14. Brinkman S, Sincovich A, Thanh VB. The Status of Early Childhood Health and Development in Kiribati: Results from a population wide census. Washington, D.C: World Bank Group; 2017. Available from: http://documents.worldbank.org/curated/en/853791563251577440/pdf/The-Status-of-Early-Childhood-Health-and-Development-in-Kiribati-Results-from-a-Population-Wide-Census.pdf
15. Brinkman S, Sincovich A, Thanh VB. How are Tuvalu's children developing? Evidence-based policy recommendations for better early childhood development in Tuvalu. Washington, D.C: World Bank Group; 2017. Available from: http://documents.worldbank.org/curated/en/373411563250298102/How-are-Tuvalus-children-developing-Evidence-based-policy-recommendations-for-better-early-childhood-development-of-Tuvaluan-children
16. Santos DD, Zanon C. The Brazilian experience with the eHCI: the adaptation process, reduction proceedings, and validation study. Laboratório de Estudos e Pesquisas em Economia Social; 2017.
17. Zhao J, Brinkman SA, Song Y, Lu C, Young ME, Zhang Y, et al. Measuring Early Childhood Development with the Early Human Capability Index (eHCI): A Reliability and Validity Study in China. J Pediatr. under review.
18. Janus M, Offord D. Development and psychometric properties of the early development instrument (EDI): a measure of children's school readiness. Can J Behav Sci. 2007;39:1–22.
19. Rao N, Sun J, Richards B, Weber AM, Sincovich A, Darmstadt GL, et al. Assessing Diversity in Early Childhood Development in the East Asia-Pacific. Child Indic Res. 2018;12(1):235–54.
20. Pisani L, Borisova I, Dowd AJ. Developing and validating the international development and early learning assessment (IDELA). Int J Educ Res. 2018;91:1–15.
21. International Test Commission. The ITC Guidelines for Translating and Adapting Tests. 2nd ed; 2017.
22. Downing SM, Haladyna TM. Handbook of test development. Mahwah: Lawrence Erlbaum Associates; 2006.
23. World Bank. World Bank open data. 2019 Available from: https://data.worldbank.org/.
24. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model Multidiscip J. 1999;6(1):1–55.
25. Yu C. Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. Los Angeles: University of California; 2002.
26. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: A review.The Journal of Educational Research. 2006;99:323–338.
27. Costello AB, Osborne JW. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Pract Assess Res Eval. 2005;10(7):1–9.
28. Muthen LK, Muthen BO. Mplus Version 7.31. Los Angeles: Muthen & Muthen; 2012.
29. Flora DB, LaBrish C, Chalmers RP. Old and new ideas for data screening and assumption testing for exploratory and confirmatory factor analysis. Front Psychol. 2012;3(55):1–21.
30. Holgado-Tello FP, Chacón-Moscoso S, Barbero-García I, Vila-Abad E. Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. Qual Quant. 2010;44(1):153–66.
31. Brown TA. Confirmatory factor analysis for applied research. New York: Guildford Press; 2006.
32. Pett MA, Lackey NR, Sullivan JJ. Making sense of factor analysis: the use of factor analysis for instrument development in health care research. Thousand Oaks, California: Sage Publications; 2003.
33. Gadermann AM, Guhn M, Zumbo BD. Estimating ordinal reliability for Likert-type and ordinal item response data: a conceptual, empirical, and practical guide. Pract Assess Res Eval. 2012;17(3):1–13.
34. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. BMJ (Clin Res ed). 1997;314(7080):572.
35. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.
36. Byrne BM. Structural equation modeling with Mplus: basic concepts, applications, and programming. New York: Taylor and Francis Group; 2012.
37. Yapa HM, Bärnighausen T. Implementation science in resource-poor countries and communities. Implement Sci. 2018;13(154):1–13.
38. Janus M, Brinkman S, Duku E. Validity and psychometric properties of the early development instrument in Canada, Australia, United States and Jamaica. Soc Indic Res. 2011;103(2):283–97.
39. Brinkman S, Gregory T, Harris J, Hart B, Blackmore S, Janus M. Associations between the early development instrument at age 5 and Reading and numeracy skills at ages 8, 10 and 12: a prospective linked data study. Child Indic Res. 2013;6(4):695–708.
40. Duncan G, Claessens A, Huston A, Pagani L, Engel M, Sexton H, et al. School readiness and later achievement. Dev Psychol. 2007;43(6):1428–46.
41. Mesquita B. Emotions in collectivist and individualist contexts. J Pers Soc Psychol. 2001;80(1):68–74.

42. He J, van de Vijver F. Bias and Equivalence in Cross-Cultural Research. Online Readings Psychol Cult. 2012;2(2):1–19.
43. Harzing AW. Response styles in cross-national survey research: a 26-country study. Int J Cross-cult Manag. 2006;6:243–66.
44. Swain SD, Weathers D, Niedrich RW. Assessing three sources of Misresponse to reversed Likert items. J Mark Res. 2008;45(1):116–31.
45. Netemeyer RG, Bearden WO, Sharma S. Scaling procedures: Issues and applications. Newbury Park: Sage Publications, Inc; 2003.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix C: Supplementary materials from published journal article 1

Supplementary Table 3.1. Brazil eHCI items and n (%) children for whom their caregiver/teacher reported yes/able

Domain	Item	Yes/Able	Missing
Physical Health	Is this child often sick?*	186 (10.3)	13 (0.7)
	Does this child have basic knowledge of hygiene (e.g., comes to school clean, washes hands, has a clean backpack and toothbrush)?	1,737 (96.0)	10 (0.6)
Verbal Communication	Can this child communicate what he/she wants with gestures (e.g., crying or pointing)?	1,776 (98.1)	22 (1.2)
	Does this child understand the language spoken?	1,797 (99.3)	3 (0.2)
	Can this child use words to get what he/she wants?	1,731 (95.6)	16 (0.3)
	Can this child use a word or simple phrase to tell about his/her day?	1,663 (91.9)	17 (0.9)
	Can this child use multiple phrases to tell about his/her day?	1,321 (73.0)	17 (0.9)
	Can this child talk and listen to another in a conversation (maintain a dialogue)?	1,537 (84.9)	24 (1.3)
Cultural Knowledge	Can this child identify two animals?	1,781 (98.4)	15 (0.8)
	Can this child identify two important types of food?	1,660 (91.7)	15 (0.8)
	Can this child identify two plants that provide food/fruit?	1,020 (56.4)	153 (8.5)
	Can this child sing familiar children’s songs (e.g., Happy Birthday)?	1,733 (95.7)	19 (1.0)
	Can this child sing a holiday song (e.g., Christmas, or other dates or cultural events)?	1,626 (89.8)	25 (1.4)
Social and Emotional	Does this child share his/her toys and belongings?	1,611 (89.0)	20 (1.1)
	Does this child take care of his/her own belongings?	1,656 (91.5)	16 (0.9)
	Does this child show respect for adults?	1,716 (94.8)	20 (1.1)
	Does this child show respect for other children?	1,635 (90.3)	28 (1.5)
	Does this child accept responsibility for his/her actions?	1,462 (80.8)	41 (2.3)
	Does this child have regard for the feelings of others?	1,604 (88.6)	46 (2.5)
	Is this child collaborative?	1,633 (90.2)	23 (1.3)
	Is this child friendly with other children?	1,647 (91.0)	42 (2.3)
	Does this child kick, bite or hit adults or other children?*	208 (11.5)	47 (2.6)
	Does this child have difficulty waiting for his/her turn?*	554 (30.6)	22 (1.2)
	Does this child understand the difference between right and wrong?	1,551 (85.7)	53 (2.9)
	Can this child follow simple instructions on how to do something?	1,654 (90.9)	31 (1.7)
Perseverance	Does this child perform tasks autonomously?	1,485 (82.0)	35 (1.9)
	Does this child stick with a task until completion?	1,475 (81.5)	26 (1.4)
	Does this child need to be constantly reminded to finish something?*	485 (26.8)	25 (1.4)
	Is this child easily distracted in a task?*	643 (35.5)	26 (1.4)

Approaches to Learning	Does this child show more curiosity with something new than with something familiar?	1,633 (90.2)	18 (1.0)
	Does this child investigate/explore a new toy, game, puzzle or object?	1,721 (95.1)	5 (0.3)
	Does this child use objects in fantasy play/in using his/her own imagination?	1,713 (94.6)	14 (0.8)
	Is this child interested in sports and games?	1,740 (96.1)	17 (0.9)
	Does this child feel free to explore the school environment even without the presence of the teacher?	1,507 (83.3)	34 (1.9)
	Does this child show interest or curiosity when dealing with a new task or activity?	1,613 (89.1)	23 (1.3)
Numeracy	Does this child recognize geometric shapes (e.g., triangle, circle, square)?	1,017 (56.2)	103 (5.7)
	Can this child name and identify at least 3 colors?	1,538 (85.0)	63 (3.5)
	Can this child sort and classify objects with common characteristics (e.g., shape, color, size)?	1,249 (69.0)	83 (4.6)
	Can this child name and recognize the symbols of all the numbers from 1 to 10?	489 (27.0)	121 (6.7)
	Can this child count to 20?	439 (24.3)	122 (6.7)
	Can this child identify differences in height/size (e.g., a horse is taller than a dog)?	1,286 (71.0)	123 (6.8)
	Does this child know the sequence of events in a day (e.g., breakfast, lunch, dinner and bedtime)?	1,161 (64.1)	128 (7.1)
	Does this child understand the concepts of yesterday, today and tomorrow?	523 (28.9)	143 (7.9)
	Does this child have notions of weight (e.g., an elephant weighs more than a mouse)?	1,127 (62.3)	146 (8.1)
	Can this child compare amounts (e.g., the number 8 is larger than the number 2)?	401 (22.2)	127 (7.0)
Reading	Can this child follow directional reading (e.g., from left to right, top to bottom)?	664 (36.7)	119 (6.6)
	Can this child identify at least 3 letters of the alphabet?	1,383 (76.4)	60 (3.3)
	Can this child identify at least 10 letters of the alphabet?	831 (45.9)	103 (5.7)
	Can this child recognize/identify at least 4 simple and usual words?	290 (16.0)	77 (4.3)
	Can this child identify or read complex words?	21 (1.2)	64 (3.5)
	Can this child read simple sentences?	10 (0.6)	53 (2.9)
Writing	Can this child scribble on paper using a pen / pencil / crayon?	1,787 (98.7)	21 (1.2)
	Can this child draw something identifiable (e.g., a stick figure)?	1,289 (71.2)	42 (2.3)
	Can this child write at least 3 letters (e.g., A, B, C)?	1,056 (58.3)	53 (2.9)
	Can this child write his/her name?	905 (50.0)	55 (3.0)
	Can this child write (or copy) simple words?	586 (32.4)	67 (3.7)
	Can this child write simple sentences?	37 (2.0)	62 (3.4)

Note. * = reverse scored items.

Supplementary Table 3.2. China eHCI items and n (%) children for whom their caregiver/teacher reported yes/able

Domain	Item	Yes/Able	Missing
Physical Health	Is this child frequently sickly?*	2,343 (20.5)	112 (1.0)
	Does this child have good hygiene i.e. always wash their hands after toileting?	7,291 (63.8)	128 (1.1)
	Does this child have a regular diet?	9,875 (86.4)	139 (1.2)
Verbal Communication	Can this child communicate their needs by crying or pointing?	11,055 (96.7)	39 (0.3)
	Can this child understand local language?	10,101 (88.4)	70 (0.6)
	Can this child use words to get their needs met?	9,473 (82.9)	104 (0.9)
	Can this child tell you about their day using a single word or simple sentence?	8,035 (70.3)	82 (0.7)
	Can this child tell you about their day using multiple sentences?	10,048 (87.9)	107 (0.9)
	Can this child take turns speaking in a conversation?	6,380 (55.8)	140 (1.2)
Cultural Knowledge	Can this child show sympathy or compassion for others?	7,324 (64.1)	197 (1.7)
	Can this child tolerate mistakes of others?	6,692 (58.5)	163 (1.4)
	Can this child identify two culturally important food/fruits?	8,716 (76.2)	102 (0.9)
	Does this child talk politely?	9,310 (81.4)	127 (1.1)
	Does this child also treat the people well if those people treated him/he well?	8,557 (74.9)	147 (1.3)
	Does this child demonstrate respect for adults?	8,226 (72.0)	163 (1.4)
	Is this child good to his or her parents?	8,174 (71.5)	173 (1.5)
Social and Emotional	Is the child happy to share their toys and belongings?	9,130 (79.9)	105 (0.9)
	Does this child take care of their own things?	9,395 (82.2)	111 (1.0)
	Does this child demonstrate respect for other children?	7,389 (64.6)	154 (1.3)
	Does this child accept responsibility for their actions?	6,321 (55.3)	150 (1.3)
	Is this child considerate of other people's feelings?	4,489 (39.3)	151 (1.3)
	Is this child helpful?	7,592 (66.4)	139 (1.2)
	Is this child friendly to other children?	9,533 (83.4)	142 (1.2)
	Does this child kick, bite or hit adults or other children?*	3,673 (32.1)	102 (0.9)
	Is this child impatient?*	7,050 (61.7)	163 (1.4)
	Does this child understand the difference between right and wrong?	5,388 (47.1)	148 (1.3)
Perseverance	Does this child perform tasks independently?	5,893 (51.6)	168 (1.5)
	Does this child always keep at a task until they are finished?	5,450 (47.7)	155 (1.4)
	Does this child need constant reminding to finish something off?*	7,962 (69.7)	157 (1.4)
	Does this child get easily distracted from a task?*	7,578 (66.3)	176 (1.5)
Approaches to Learning	Does this child show more curiosity about something new in comparison to something familiar?	10,478 (91.7)	103 (0.9)
	Does this child investigate/explore the function of a new toy/game/puzzle or object?	9,746 (85.3)	110 (1.0)
	Does this child always want to learn something new?	9,264 (81.0)	147 (1.3)

	When in an unfamiliar environment with a familiar person present, does this child feel free to explore?	9,087 (79.5)	157 (1.4)
	Is this child diligent in their approach to a new job or task?	8,842 (77.4)	178 (1.6)
	Does this child will to find out answer if she/he does not understand something?	6,898 (60.3)	224 (2.0)
Numeracy	Can this child recognise geometric shapes (e.g. triangle, circle, square)?	8,468 (74.1)	142 (1.2)
	Can this child name and identify at least 3 colours?	9,231 (80.8)	172 (1.5)
	Can this child sort and classify objects by common characteristics (e.g. shape, colour, size)?	7,874 (68.9)	170 (1.5)
	Can this child name and recognise the symbol of all numbers from 1 to 10?	8,281 (72.4)	139 (1.2)
	Can this child count to 10 without any help?	8,783 (76.8)	148 (1.3)
	Can this child count to 20 without any help?	6,428 (56.2)	135 (1.2)
	Can this child count to 100 without any help?	2,678 (23.4)	169 (1.5)
	Does this child know that a cow is taller than a dog?	9,665 (84.6)	185 (1.6)
	Does this child know the order of the day (e.g., breakfast then lunch then dinner then sleep?)	8,325 (72.8)	187 (1.6)
	Does this child understand the concepts of yesterday, today and tomorrow?	4,917 (43.0)	185 (1.6)
	Does this child know that an elephant weighs more than a mouse?	8,527 (74.6)	166 (1.5)
	Does the child know that the number 8 is bigger than the number 2?	6,761 (59.1)	210 (1.8)
Reading	Can this child recognise 10 Chinese characters?	4,675 (40.9)	185 (1.6)
	Can this child recognise 20 Chinese characters?	2,708 (23.7)	177 (1.5)
	Can this child recognise 100 Chinese characters?	820 (7.2)	190 (1.7)
	Can this child hold one book in right way?	8,306 (72.7)	188 (1.6)
	Can this child follow reading directions? (i.e. left to right, top to bottom)	6,406 (56.0)	174 (1.5)
	Can this child read a book and turn pages by himself?	7,649 (66.9)	188 (1.6)
	Can this child read simple sentences?	4,065 (35.6)	189 (1.7)
	Can this child read complex sentences?	2,237 (19.6)	205 (1.8)
Writing	Can this child scribble on a page using a pen/pencil/crayon?	4,515 (39.5)	208 (1.8)
	Can this child write at least 3 characters?	4,342 (38.0)	199 (1.7)
	Can this child write simple sentences?	1,677 (14.7)	211 (1.8)

Note. * = reverse scored items.

Supplementary Table 3.3. Kiribati eHCI items and n (%) children for whom their caregiver/teacher reported yes/able

Domain	Item	Yes/Able	Missing
Physical Health	Does the child get sick often?*	5,883 (70.5)	419 (5.0)
	Is the child practicing cleanliness and healthy living (cleans hands after toileting)?	6,330 (75.9)	285 (3.4)
	Is the child personally practicing cleanliness and healthy living on their own?	5,252 (63.0)	170 (2.0)
	Is the child careful from being hurt (e.g. burns, drowning, falls)?	6,670 (80.0)	192 (2.3)
	Does the child know the difference between good and bad food?	5,991 (71.8)	449 (5.4)
Verbal Communication	The child is able to use a sequence of words.	5,620 (67.4)	161 (1.9)
	The child is able to use a simple sentence.	5,015 (60.1)	167 (2.0)
	The child is able to wait for the other person to finish speaking in a conversation, before they speak.	4,893 (58.7)	128 (1.5)
	The child is able to explain things in Kiribati.	5,404 (64.8)	144 (1.7)
	The child is able to communicate as a mature person (talkative, enquiring).	6,345 (76.1)	118 (1.4)
	The child knows their name.	7,806 (93.6)	102 (1.2)
	The child knows the name of one of their parents/guardians.	7,708 (92.4)	92 (1.1)
Cultural Knowledge	The child is able to exhibit behaviours of affection, understanding and patience to others.	6,515 (78.1)	101 (1.2)
	The child is able to identify two valuable foods in Kiribati.	5,488 (65.8)	121 (1.5)
	The child is able to identify two edible plants in Kiribati.	5,743 (68.8)	160 (1.9)
	The child is able to express Kiribati behaviours and traditions (giving respect to others, being humble).	5,106 (61.2)	128 (1.5)
	The child is able to exhibit behaviours of trustworthiness and commitment to do something.	5,313 (63.7)	158 (1.9)
	The child is able to make good friendships.	6,515 (78.1)	140 (1.7)
	The child is able to join cultural and traditional way of Kiribati life (Kiribati local dance).	5,101 (61.1)	162 (1.9)
	The child is able to say a short prayer.	6,631 (79.5)	145 (1.7)
Social and Emotional	The child is willing to share his toys and belongings with others.	6,501 (77.9)	99 (1.2)
	The child is able to keep his belongings very well.	6,179 (74.1)	112 (1.3)
	The child knows how to respect older people.	4,883 (58.5)	153 (1.8)
	The child knows how to respect other children.	5,346 (64.1)	157 (1.9)
	The child accepts his/her responsibilities when he/she is being instructed to carry them out.	6,725 (80.6)	133 (1.6)
	The child welcomes the opinions of others.	6,457 (77.4)	164 (2.0)
	The child does what he/she is supposed to do, or not to do.	5,446 (65.3)	157 (1.9)
	The child is willing to help others.	6,089 (73.0)	153 (1.8)
	The child communicates easily with other children.	6,786 (81.3)	101 (1.2)
	The child frequently kicks, bites, or hits older people or children.*	3,511 (42.1)	141 (1.7)
	The child can be patient long enough before receiving his/her needs.	5,760 (69.0)	130 (1.6)
	The child always knows the difference between good and bad.	6,147 (73.7)	144 (1.7)
	The child can follow simple instructions.	7,055 (84.6)	143 (1.7)

Perseverance	The child can mostly do his/her work on his own.	5,512 (66.1)	88 (1.1)
	The child always completes his/her work.	4,894 (58.7)	106 (1.3)
	The child always needs to be reminded about completing what he/she was doing.*	6,369 (76.7)	137 (1.6)
	The child gets bored quickly when he/she was doing his/her job/task.*	5,671 (68.0)	159 (1.9)
Approaches to Learning	The child prefers learning new ideas to familiar concepts.	7,293 (87.4)	118 (1.4)
	The child examines how a new toy works.	7,157 (85.8)	124 (1.5)
	The child always desires learning of new concepts.	7,104 (85.2)	150 (1.8)
	When the child is placed in an unfamiliar setting with a person they know, they are delighted to learn.	3,258 (39.1)	139 (1.7)
	The child is keen to learn new activities.	5,807 (69.6)	154 (1.8)
Numeracy	The child is able to see shapes such as a triangle, a circle, and a square.	6,360 (76.2)	127 (1.5)
	The child is able to name and identify 3 colours or more.	5,412 (64.9)	136 (1.6)
	The child is able to sort and classify objects (such as shapes, colours and sizes).	4,947 (59.3)	198 (2.4)
	The child is able to pronounce and recognise numbers from 1 to 10.	4,955 (59.4)	157 (1.9)
	The child is able to count up to 10.	5,851 (70.1)	166 (2.0)
	The child is able to count up to 20.	1,512 (18.1)	174 (2.1)
	The child is able to count up to 100.	398 (4.8)	190 (2.3)
	The child is aware that a dog is taller than a rat.	6,246 (74.9)	178 (2.1)
	The child is aware of the order of time in a day (morning, then afternoon then evening).	3,007 (36.0)	182 (2.2)
	The child is aware of yesterday, today and tomorrow.	2,273 (27.2)	198 (2.4)
	The child is aware that a chair is heavier than a pencil.	6,203 (74.4)	174 (2.1)
	The child is aware that number 8 is larger than number 2.	3,502 (42.0)	224 (2.7)
Reading	The child knows the pronunciation of three letters in the sequence of A E I?	6,661 (79.8)	169 (2.0)
	The child is able to identify 3 letters or more in the sequence of A E I?	5,062 (60.7)	163 (2.0)
	The child is able to identify 10 letters or more in the sequence of A E I . . . ?	2,935 (35.2)	331 (4.0)
	The child is able to properly hold the book and appropriately turn its pages in the right order?	4,201 (50.4)	194 (2.3)
	The child is able to follow the right way of reading (from left to right, from top to bottom).	2,157 (25.9)	198 (2.4)
	The child is able to read 4 or more familiar words.	3,304 (39.6)	206 (2.5)
Writing	The child is able to draw a picture that could be recognised (persons image).	5,692 (68.2)	141 (1.7)
	The child is able to copy or trace the outline of a letter over an already written letter.	5,349 (64.1)	164 (2.0)
	The child is able to write 3 letters or more (A E I).	4,288 (51.4)	165 (2.0)
	The child is able to write his name.	2,706 (32.4)	177 (2.1)
	The child is able to write simple words.	2,480 (29.7)	174 (2.1)

Note. * = reverse scored items.

Supplementary Table 3.4. Lao PDR eHCI items and n (%) children for whom their caregiver/teacher reported yes/able

Domain	Item	Yes/Able	Missing
Verbal Communication	Can child tell you what he/she wants?	7154 (95.5)	0 (0.0)
	Can child speak a few simple words or sentences to explain what happened to him/her?	6713 (89.6)	0 (0.0)
	Can child speak many words or sentences to explain what happened to him/her?	5671 (75.7)	0 (0.0)
	Can child communicate well with you on any topics?	5708 (76.2)	0 (0.0)
Cultural Knowledge	Can child tell a type of animal, at least two animals in the area?	7184 (95.9)	0 (0.0)
	Can child tell a food name, at least two dishes that are available in the area?	6695 (89.4)	0 (0.0)
	Can child tell a name of a plant/vegetable/fruit, at least two types in the area?	6760 (90.2)	0 (0.0)
	Can child sing?	3976 (53.1)	0 (0.0)
	Can child participate in traditional events such as giving alms and seeing monks?	1937 (25.9)	0 (0.0)
Social and Emotional	Is child happy to share his/her toys with others?	5874 (78.4)	0 (0.0)
	Does child know how to take care of his/her belongings?	4950 (66.1)	0 (0.0)
	Has child shown respect to elders?	1907 (25.5)	0 (0.0)
	Does child respect other kids?	2800 (37.4)	0 (0.0)
	Is child responsible for his/her own behaviour?	1844 (24.6)	0 (0.0)
	Does child consider other people’s feelings?	1560 (20.8)	0 (0.0)
	Can child help other people?	3104 (41.4)	0 (0.0)
	Is child friendly to other kids?	6375 (85.1)	0 (0.0)
	Is child hot tempered?*	4141 (55.3)	0 (0.0)
	Can child understand the difference between right and wrong?	2291 (30.6)	0 (0.0)
	Can child respond to a simple instruction?	6376 (85.1)	0 (0.0)
	Is your child very clingy (i.e. does not want to leave their parent’s side)?*	6502 (86.8)	0 (0.0)
	Does your child understand his/her feelings and is able to describe his/her feelings, for example by saying “I’m happy…: or “ I’m sad…”?	2935 (39.2)	0 (0.0)
	Does your child stop an activity when told to do so straight away?	6243 (83.3)	0 (0.0)
Perseverance	Can child do something on his/her own?	4791 (63.9)	0 (0.0)
	When child is doing something, does he/she finish it?	3692 (49.3)	0 (0.0)
	Does child have to be told several times so that then he/she can finish what he/she is doing?*	4258 (56.8)	0 (0.0)
	When child is doing something, does he/she lose focus easily?*	3912 (52.5)	0 (0.0)
Approaches to Learning	Does child show any sign of interest to learn new things?	6550 (87.4)	0 (0.0)
	Does child try to learn how to play with new toys?	6230 (83.1)	0 (0.0)
	Has child ever used any object to role play (using imagination) e.g. banana stem horse, sword fighting etc?	4832 (64.5)	0 (0.0)

	Does child show interest in playing games such as collecting pebbles, jumping with rubber bands, hide and seek, ball throwing etc?	5562 (74.2)	0 (0.0)
	When the kids go somewhere with you or another household member, are they brave to survey or ask some questions with another person?	4874 (65.0)	0 (0.0)
	When child is doing a particular activity, they will pay intense attention to the activity.	4862 (62.5)	0 (0.0)
	Does child think things out before starting a task?	2613 (34.9)	0 (0.0)
Numeracy	Can child distinguish between a triangle, circle and rectangle?	1203 (16.1)	0 (0.0)
	Can child tell at least three different colours?	4345 (58.0)	0 (0.0)
	Can child distinguish objects based on shape, colour, and size?	2050 (27.4)	0 (0.0)
	Can child count from 1 to 10?	2900 (38.7)	0 (0.0)
	Can child count from 1 to 20?	1320 (17.6)	0 (0.0)
	Does child know that a tiger is taller than a cat?'	2527 (33.7)	0 (0.0)
	Does child know morning, afternoon and evening?	5623 (75.0)	0 (0.0)
	Can child tell you if it is yesterday, today or tomorrow?	4813 (64.2)	0 (0.0)
	Does child know that an elephant is heavier than a pig?	1919 (25.6)	0 (0.0)
	Does child know that 8 is more than 2?	1043 (13.9)	0 (0.0)
Reading	Can child read in the correct direction from left to right and from top to bottom (even if they can't read)?	1288 (17.2)	0 (0.0)
	Can child tell at least 3 letters of the alphabet?	2070 (27.6)	0 (0.0)
	Can child tell at least 10 letters of the alphabet?	963 (12.9)	0 (0.0)
	Can child read at least 4 simple words?	376 (5.0)	0 (0.0)
	Can child read difficult words such as axe, buffalo?	323 (4.3)	0 (0.0)
	Can child read simple sentences?	314 (4.2)	0 (0.0)
Writing	Can child write or draw with pencil, coloured pencils or pen?	4632 (61.8)	0 (0.0)
	Can child draw a picture that you can tell what it is?	993 (13.3)	0 (0.0)
	Can child write at least three letters such as A, B, C?	1236 (16.5)	0 (0.0)
	Can child write his/her own name?	310 (4.1)	0 (0.0)
	Can child write a simple word?	239 (3.2)	0 (0.0)
	Can child write a simple sentence?	135 (1.8)	0 (0.0)

Note. * = reverse scored items.

Supplementary Table 3.5. Samoa eHCI items and n (%) children for whom their caregiver/teacher reported yes/able

Domain	Item	Yes/Able	Missing
Physical Health	Is the child frequently sick?*	1076 (8.8)	0 (0.0)
	Does the child have good hygiene (e.g. always washes hands after toileting)?	9503 (78.0)	0 (0.0)
	Does the child know the difference between healthy and unhealthy food?	4757 (39.0)	0 (0.0)
Verbal Communication	Does the child use a group of words in conversation?	6955 (57.1)	0 (0.0)
	Can the child talk about something that they have done?	7895 (64.8)	0 (0.0)
	Can the child use respectful language?	4618 (37.9)	0 (0.0)
	Does the child always ask questions?	7980 (65.5)	0 (0.0)
Cultural Knowledge	Does the child show respect and compassion to others?	4597 (37.7)	0 (0.0)
	Can the child name two Samoan foods?	4960 (40.7)	0 (0.0)
	Can the child give the names of two plants?	4809 (39.4)	0 (0.0)
	Can the child show the Samoan cultural values of humility?	5311 (43.6)	0 (0.0)
	Do you believe and trust that this child can do anything that you ask them to do?	4304 (35.3)	0 (0.0)
	Does the child do sharing in relationships and work co-operatively with others?	4920 (40.4)	0 (0.0)
	Does the child participate in community cultural routines (e.g. sports and entertainment)?	5118 (42.0)	0 (0.0)
	Can the child say a short prayer?	7090 (58.2)	0 (0.0)
Social and Emotional	Is the child happy to share toys and belongings?	7189 (59.0)	0 (0.0)
	Does the child take care of their own belongings?	6348 (52.1)	0 (0.0)
	Does the child show respect to adults?	4762 (39.1)	0 (0.0)
	Does the child show respect to children?	4724 (38.7)	0 (0.0)
	Does the child accept decision making for their actions?	5344 (43.8)	0 (0.0)
	If you discipline the child on appropriate behaviour, do they learn or not?	5257 (43.1)	0 (0.0)
	Is the child considerate of other people’s feelings?	4620 (37.9)	0 (0.0)
	Is the child always helpful?	7449 (61.1)	0 (0.0)
	Does the child have friends?	10113 (83.0)	0 (0.0)
	Is the child playful and undisciplined?*	9602 (78.8)	0 (0.0)
	Is the child impatient?*	6111 (50.1)	0 (0.0)
	Does the child understand the difference between right and wrong?	5514 (45.2)	0 (0.0)
	Does the child follow simple commands?	8396 (68.9)	0 (0.0)
Perseverance	Can the child perform tasks independently?	4295 (35.2)	0 (0.0)
	Does the child always do tasks completely?	4216 (34.6)	0 (0.0)
	Is the child usually ordered to complete tasks?*	7221 (59.2)	0 (0.0)
	Does the child lose interest easily and give up?*	8228 (67.5)	0 (0.0)
	Does the child show more curiosity about something new in comparison to something familiar?	8382 (68.8)	0 (0.0)

Approaches to Learning	Is the child happy to explore and investigate the function of his/her toys?	8443 (69.3)	0 (0.0)
	Does the child always want to learn new things?	8171 (67.0)	0 (0.0)
	When in an unfamiliar environment with a familiar person present, does the child feel free to explore?	6839 (56.1)	0 (0.0)
	Is the child diligent in the tasks that you ask him/her to do?	5022 (41.2)	0 (0.0)
Numeracy	Can the child recognise geometric shapes (e.g. triangle, circle, square)?	4259 (34.9)	0 (0.0)
	Can the child observe and identify three colours?	4580 (37.6)	0 (0.0)
	Can the child sort and classify objects in common characteristics of colours, shapes and sizes?	3466 (28.4)	0 (0.0)
	Can the child name and recognise the symbols of all numbers from 1-10?	5088 (41.7)	0 (0.0)
	Can the child count to 10?	8255 (67.7)	0 (0.0)
	Can the child count to 20?	2630 (21.6)	0 (0.0)
	Can the child count to 100?	679 (5.6)	0 (0.0)
	Does the child know that a horse is taller than a dog?	5141 (42.2)	0 (0.0)
	Does the child know the order of the day (morning, afternoon, then evening)?	3022 (24.8)	0 (0.0)
	Does the child understand the concepts of yesterday, today and tomorrow?	3102 (25.4)	0 (0.0)
	Does the child know that a bus weighs more than a bicycle?	4322 (35.5)	0 (0.0)
	Does the child know that number 8 is bigger than number 2?	3900 (32.0)	0 (0.0)
Reading	Does the child know the sounds of 3 letters of the Samoan alphabet?	6563 (53.8)	0 (0.0)
	Can the child identify at least 3 letters of the Samoan alphabet?	6119 (50.2)	0 (0.0)
	Can the child identify at least 10 letters of the Samoan alphabet?	4294 (35.2)	0 (0.0)
	Can the child hold a book and turn the pages in the right way?	4592 (37.7)	0 (0.0)
	Can the child follow the right reading direction (e.g. left to right, top to bottom)?	3148 (25.8)	0 (0.0)
	Can the child read simple and popular words?	3296 (27.0)	0 (0.0)
Writing	Can the child draw something that is identifiable (e.g. a stick person)?	5467 (44.8)	0 (0.0)
	Can the child copy or trace the shape of a letter? (e.g. A, E, I)	5178 (42.5)	0 (0.0)
	Can the child write 3 letters? (e.g. A, E, I)	5028 (41.2)	0 (0.0)
	Can the child write their own name?	3228 (26.5)	0 (0.0)
	Can the child write simple words?	3334 (27.3)	0 (0.0)

Note. * = reverse scored items.

Supplementary Table 3.6. Tuvalu eHCI items and n (%) children for whom their caregiver/teacher reported yes/able

Domain	Item	Yes/Able	Missing
Physical Health	Does this child get sick often?*	219 (39.9)	3 (0.5)
	Is the child practicing cleanliness and healthy living (clean hands each time he/she uses the toilet)?	440 (80.1)	0 (0.0)
	Is the child personally practicing cleanliness and healthy living on his own?	337 (61.4)	0 (0.0)
	Is the child being careful from being hurt (burnt, drown, fall, stumble)?	427 (77.8)	3 (0.5)
	Does the child know the difference between good and bad food?	433 (78.9)	0 (0.0)
Verbal Communication	The child is able to use a group of words.	488 (88.9)	0 (0.0)
	The child is able to use a complete sentence.	346 (63.0)	0 (0.0)
	The child is able to wait for the other person to finish speaking, in a conversation, before he/she could speak.	263 (47.9)	0 (0.0)
	The child is able to interpret things in Tuvaluan.	351 (63.9)	0 (0.0)
	The child is able to communicate as a mature person (talkative, enquiring).	378 (68.9)	0 (0.0)
	The child knows his/her name.	534 (97.3)	1 (0.2)
	The child knows the name of one of his/her parents/guardians.	527 (96.0)	0 (0.0)
Cultural Knowledge	The child is able to exhibit behaviours of affection, understanding and patience to others.	364 (66.3)	0 (0.0)
	The child is able to identify two valuable foods in Tuvalu.	381 (69.4)	0 (0.0)
	The child is able to identify two edible plants in Tuvalu.	398 (72.5)	1 (0.2)
	The child is able to express Tuvaluan behaviours and traditions (giving respect to others, being humble).	277 (50.5)	1 (0.2)
	The child is able to exhibit behaviours of loyalty and commitment to do something.	312 (56.8)	0 (0.0)
	The child is able to demonstrate qualities of good friendship.	448 (81.6)	0 (0.0)
	The child is able to join cultural and traditional way of Tuvalu life (Tuvaluan local dance).	408 (74.3)	0 (0.0)
	The child is able to say a usual short prayer.	474 (86.3)	0 (0.0)
	The child is able to say a short prayer using own words.	232 (42.3)	0 (0.0)
Social and Emotional	The child is willing to share his toys and belongings with others.	432 (78.7)	0 (0.0)
	The child is able to keep his belongings very well.	377 (68.7)	0 (0.0)
	The child knows how to respect older people.	357 (65.0)	0 (0.0)
	The child knows how to respect other children.	334 (60.8)	0 (0.0)
	The child accepts his/her responsibilities when he/she is being instructed to carry them out.	435 (79.2)	0 (0.0)
	The child welcomes the opinions of others.	341 (62.1)	0 (0.0)
	The child continually does whatever was told of him/her not to do.*	343 (62.5)	0 (0.0)
	The child is willing to help others.	485 (88.3)	1 (0.2)
	The child gets along easily with other children.	460 (83.8)	0 (0.0)
	The child frequently kicks, bites, or hits older children or people.*	207 (37.7)	1 (0.2)
	The child can be patient long enough before receiving his/her needs.	333 (60.7)	1 (0.2)

	The child always knows the difference between good and bad.	357 (65.0)	0 (0.0)
	The child can follow simple instructions.	519 (94.5)	0 (0.0)
Perseverance	The child can always do his/her work on his own.	411 (74.9)	1 (0.2)
	The child always completes his/her work.	318 (57.9)	1 (0.2)
	The child always needs to be reminded about completing what he/she was doing.*	402 (73.2)	1 (0.2)
	The child gets bored quickly when he/she was doing his/her job task.*	368 (67.0)	1 (0.2)
Approaches to Learning	The child prefers learning new ideas to familiar concepts.	499 (90.9)	1 (0.2)
	The child examines how a new toy works.	495 (90.2)	1 (0.2)
	The child always desires learning of new concepts.	495 (90.2)	1 (0.2)
	When the child is placed in an unfamiliar setting with a person he/she knows, would he be delighted to learn?	424 (77.2)	1 (0.2)
	The child always considers a school activity carefully and works on it wholeheartedly.	368 (67.0)	1 (0.2)
Numeracy	The child is able to see shapes such as a triangle, a circle, and a square.	338 (61.6)	1 (0.2)
	The child is able to pronounce names, and divide 3 colours of more.	336 (61.2)	1 (0.2)
	The child is able to divide and arrange these items to their own parts (such as shapes, colours and sizes).	332 (60.5)	1 (0.2)
	The child is able to pronounce and recognise numbers from 1 to 10.	316 (57.6)	1 (0.2)
	The child is able to count up to 10.	498 (90.7)	1 (0.2)
	The child is able to count up to 20.	224 (40.8)	1 (0.2)
	The child is able to count up to 100.	62 (11.3)	2 (0.4)
	The child is aware that the dog is taller than the mouse.	410 (74.7)	1 (0.2)
	The child is aware of the order of time in a day (morning, then afternoon then evening).	253 (46.1)	1 (0.2)
	The child is aware of yesterday, today and tomorrow.	189 (34.4)	1 (0.2)
	The child is aware that the chair is heavier than a pencil.	383 (69.8)	0 (0.0)
	The child is aware that the number 8 is larger than the number 2.	239 (43.5)	2 (0.4)
Reading	The child knows the pronunciation of three letters in the sequence A E I.	399 (72.7)	1 (0.2)
	The child is able to identify 3 letters or more in the sequence A E I.	347 (63.2)	1 (0.2)
	The child is able to identify 10 letters or more in the sequence A E I.	210 (38.3)	1 (0.2)
	The child is able to properly hold the book and appropriately turn its pages in the right order.	308 (56.1)	1 (0.2)
	The child is able to follow the right way or reading (from left to right, from top to bottom).	232 (42.3)	1 (0.2)
	The child is able to read 4 or more familiar words.	160 (29.1)	1 (0.2)
Writing	The child is able to draw a picture that could be recognised (person's image).	439 (80.0)	1 (0.2)
	Can this child copy or trace the outline of a letter over an already written letter?	364 (66.3)	1 (0.2)
	The child is able to write 3 letters or more (A, E, I).	282 (51.4)	1 (0.2)
	The child is able to write his name.	237 (43.2)	1 (0.2)
	The child is able to write simple words.	184 (33.5)	4 (0.7)

Note. * = reverse scored items.

Supplementary Table 3.7. Factor loadings from confirmatory factor analysis in Brazil

	F1	F2	F3	F4	F5	F6	F7	F8	F9
Phys 1	0.49								
Phys 2	0.72								
Comm 1		0.72							
Comm 2		0.66							
Comm 3		0.84							
Comm 4		0.94							
Comm 5		0.90							
Comm 6		0.88							
Cult 1			0.93						
Cult 2			0.76						
Cult 3			0.88						
Cult 4			0.99						
Cult 5			0.88						
Soc 1				0.77					
Soc 2				0.73					
Soc 3				0.92					
Soc 4				0.94					
Soc 5				0.90					
Soc 6				0.88					
Soc 7				0.85					
Soc 8				0.89					
Soc 9				0.64					
Soc 10				0.55					
Soc 11				0.81					
Soc 12				0.85					
Persev 1					0.87				
Persev 2					0.95				
Persev 3					0.83				
Persev 4					0.83				
Appr 1						0.70			
Appr 2						0.82			
Appr 3						0.83			
Appr 4						0.85			
Appr 5						0.64			
Appr 6						1.00			
Num 1							0.78		
Num 2							0.88		
Num 3							0.83		
Num 4							0.80		
Num 5							0.89		
Num 6							0.85		
Num 7							0.70		
Num 8							0.72		
Num 9							0.80		
Num 10							0.87		
Read 1								0.67	
Read 2								0.93	
Read 3								0.93	
Read 4								0.74	
Read 5								0.43	
Read 6								0.36	

Writ 1	0.21
Writ 2	0.88
Writ 3	0.97
Writ 4	0.98
Writ 5	0.97
Writ 6	0.65

Note. Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

Supplementary Table 3.8. Factor loadings from confirmatory factor analysis in China

	F10	F2	F3	F4	F5	F6	F7	F8	F9
Phys 1	0.12								
Phys 2	0.71								
Phys 3	0.64								
Comm 1		0.52							
Comm 2		0.86							
Comm 3		0.90							
Comm 4		0.74							
Comm 5		0.70							
Comm 6		0.78							
Cult 1			0.71						
Cult 2			0.79						
Cult 3			0.62						
Cult 4			0.79						
Cult 5			0.84						
Cult 6			0.84						
Cult 7			0.80						
Soc 1				0.67					
Soc 2				0.67					
Soc 3				0.79					
Soc 4				0.80					
Soc 5				0.74					
Soc 6				0.77					
Soc 7				0.59					
Soc 8				0.14					
Soc 9				0.04					
Soc 10				0.68					
Persev 1					0.82				
Persev 2					0.83				
Persev 3					0.05				
Persev 4					0.02				
Appr 1						0.65			
Appr 2						0.70			
Appr 3						0.73			
Appr 4						0.59			
Appr 5						0.80			
Appr 6						0.81			
Num 1							0.76		
Num 2							0.79		
Num 3							0.78		
Num 4							0.89		
Num 5							0.89		
Num 6							0.90		
Num 7							0.77		
Num 8							0.69		
Num 9							0.74		
Num 10							0.76		
Num 11							0.73		
Num 12							0.86		
Read 1								0.94	
Read 2								0.88	
Read 3								0.47	

Read 4	0.77	
Read 5	0.84	
Read 6	0.75	
Read 7	0.81	
Read 8	0.73	
Writ 1		0.66
Writ 2		0.89
Writ 3		0.69

Note. Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

Supplementary Table 3.9. Factor loadings from confirmatory factor analysis in Kiribati

	F1	F2	F3	F4	F5	F6	F7	F8	F9
Phys 1	0.09								
Phys 2	0.73								
Phys 3	0.79								
Phys 4	0.82								
Phys 5	0.88								
Comm 1		0.84							
Comm 2		0.84							
Comm 3		0.68							
Comm 4		0.90							
Comm 5		0.87							
Comm 6		0.81							
Comm 7		0.84							
Cult 1			0.75						
Cult 2			0.89						
Cult 3			0.91						
Cult 4			0.82						
Cult 5			0.81						
Cult 6			0.81						
Cult 7			0.69						
Cult 8			0.81						
Soc 1				0.67					
Soc 2				0.72					
Soc 3				0.82					
Soc 4				0.78					
Soc 5				0.77					
Soc 6				0.72					
Soc 7				0.65					
Soc 8				0.83					
Soc 9				0.70					
Soc 10				-0.06					
Soc 11				0.61					
Soc 12				0.80					
Soc 13				0.79					
Persev 1					0.69				
Persev 2					0.92				
Persev 3					-0.42				
Persev 4					-0.24				
Appr 1						0.84			
Appr 2						0.84			
Appr 3						0.82			
Appr 4						0.48			
Appr 5						0.91			
Num 1							0.88		
Num 2							0.88		
Num 3							0.83		
Num 4							0.82		
Num 5							0.79		
Num 6							0.59		
Num 7							0.44		
Num 8							0.85		
Num 9							0.74		

Num 10	0.76		
Num 11	0.84		
Num 12	0.80		
Read 1		0.84	
Read 2		0.92	
Read 3		0.84	
Read 4		0.76	
Read 5		0.80	
Read 6		0.75	
Writ 1			0.90
Writ 2			0.88
Writ 3			0.94
Writ 4			0.89
Writ 5			0.87

Note. Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

suppl

Supplementary Table 3.10. Factor loadings from confirmatory factor analysis in Lao PDR

	F1	F2	F3	F4	F5	F6	F7	F8
Comm 1	0.73							
Comm 2	0.91							
Comm 3	0.94							
Comm 4	0.80							
Cult 1		0.58						
Cult 2		0.80						
Cult 3		0.83						
Cult 4		0.81						
Cult 5		0.59						
Soc 1			0.60					
Soc 2			0.67					
Soc 3			0.74					
Soc 4			0.67					
Soc 5			0.72					
Soc 6			0.75					
Soc 7			0.76					
Soc 8			0.58					
Soc 9			0.03					
Soc 10			0.72					
Soc 11			0.59					
Soc 12			-0.25					
Soc 13			0.63					
Soc 14			0.43					
Persev 1				0.97				
Persev 2				0.88				
Persev 3				-0.25				
Persev 4				-0.31				
Appr 1					0.81			
Appr 2					0.85			
Appr 3					0.70			
Appr 4					0.79			
Appr 5					0.67			
Appr 6					0.70			
Appr 7					0.63			
Num 1						0.68		
Num 2						0.72		
Num 3						0.67		
Num 4						0.77		
Num 5						0.76		
Num 6						0.64		
Num 7						0.84		
Num 8						0.84		
Num 9						0.70		
Num 10						0.83		
Read 1							0.79	
Read 2							0.93	
Read 3							0.94	
Read 4							0.87	
Read 5							0.63	
Read 6							0.83	
Writ 1								0.71

Writ 2	0.84
Writ 3	0.93
Writ 4	0.92
Writ 5	0.92
Writ 6	0.85

Note. Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

Supplementary Table 3.11. Factor loadings from confirmatory factor analysis in Samoa

	F1	F2	F3	F4	F5	F6	F7	F8	F9
Phys 1	0.04								
Phys 2	0.61								
Phys 3	1.04								
Comm 1		0.87							
Comm 2		0.91							
Comm 3		1.04							
Comm 4		0.62							
Cult 1			0.80						
Cult 2			0.94						
Cult 3			0.96						
Cult 4			0.91						
Cult 5			0.96						
Cult 6			0.97						
Cult 7			0.70						
Cult 8			0.77						
Soc 1				0.85					
Soc 2				0.87					
Soc 3				0.94					
Soc 4				0.95					
Soc 5				0.94					
Soc 6				0.95					
Soc 7				0.93					
Soc 8				0.86					
Soc 9				0.63					
Soc 10				-0.32					
Soc 11				-0.59					
Soc 12				0.93					
Soc 13				0.82					
Persev 1					0.74				
Persev 2					1.09				
Persev 3					-0.38				
Persev 4					-0.27				
Appr 1						0.84			
Appr 2						0.82			
Appr 3						0.92			
Appr 4						0.89			
Appr 5						1.04			
Num 1							0.81		
Num 2							0.92		
Num 3							0.92		
Num 4							0.91		
Num 5							0.82		
Num 6							0.85		
Num 7							0.80		
Num 8							0.85		
Num 9							0.92		
Num 10							0.92		
Num 11							0.92		
Num 12							0.96		
Read 1								0.90	
Read 2								0.92	

Read 3	0.90	
Read 4	0.90	
Read 5	0.87	
Read 6	0.92	
Writ 1		0.90
Writ 2		0.95
Writ 3		0.95
Writ 4		0.97
Writ 5		0.93

Note. Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

Supplementary Table 3.12. Factor loadings from confirmatory factor analysis in Tuvalu

	F1	F2	F3	F4	F5	F6	F7	F8	F9
Phys 1	0.09								
Phys 2	0.64								
Phys 3	0.79								
Phys 4	1.00								
Phys 5	0.95								
Comm 1		0.76							
Comm 2		0.79							
Comm 3		0.76							
Comm 4		0.90							
Comm 5		0.79							
Comm 6		0.81							
Comm 7		0.95							
Cult 1			0.69						
Cult 2			0.94						
Cult 3			0.94						
Cult 4			0.81						
Cult 5			0.83						
Cult 6			0.72						
Cult 7			0.77						
Cult 8			0.80						
Cult 9			0.75						
Soc 1				0.72					
Soc 2				0.83					
Soc 3				0.86					
Soc 4				0.79					
Soc 5				0.77					
Soc 6				0.79					
Soc 7				-0.21					
Soc 8				0.80					
Soc 9				0.65					
Soc 10				0.14					
Soc 11				0.70					
Soc 12				0.88					
Soc 13				0.76					
Persev 1					0.87				
Persev 2					0.91				
Persev 3					0.04				
Persev 4					0.40				
Appr 1						0.81			
Appr 2						0.96			
Appr 3						0.97			
Appr 4						0.81			
Appr 5						1.13			
Num 1							0.90		
Num 2							0.93		
Num 3							0.90		
Num 4							0.83		
Num 5							0.87		
Num 6							0.80		
Num 7							0.79		
Num 8							0.91		

Num 9	0.92		
Num 10	0.84		
Num 11	0.86		
Num 12	0.91		
Read 1		0.80	
Read 2		0.93	
Read 3		0.87	
Read 4		0.81	
Read 5		0.88	
Read 6		0.80	
Writ 1			0.74
Writ 2			0.93
Writ 3			0.99
Writ 4			0.95
Writ 5			0.81

Note. Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

Supplementary Table 3.13. Factor loadings from confirmatory factor analysis in Lao PDR – low maternal education

	F1	F2	F3	F4	F5	F6	F7	F8
Comm 1	0.75							
Comm 2	0.90							
Comm 3	0.95							
Comm 4	0.80							
Cult 1		0.59						
Cult 2		0.81						
Cult 3		0.84						
Cult 4		0.80						
Cult 5		0.58						
Soc 1			0.59					
Soc 2			0.67					
Soc 3			0.74					
Soc 4			0.68					
Soc 5			0.71					
Soc 6			0.74					
Soc 7			0.76					
Soc 8			0.58					
Soc 9			0.04					
Soc 10			0.72					
Soc 11			0.59					
Soc 12			-0.26					
Soc 13			0.63					
Soc 14			0.49					
Persev 1				0.97				
Persev 2				0.88				
Persev 3				-0.27				
Persev 4				-0.32				
Appr 1					0.82			
Appr 2					0.85			
Appr 3					0.72			
Appr 4					0.79			
Appr 5					0.69			
Appr 6					0.69			
Appr 7					0.64			
Num 1						0.66		
Num 2						0.71		
Num 3						0.66		
Num 4						0.76		
Num 5						0.75		
Num 6						0.63		
Num 7						0.85		
Num 8						0.85		
Num 9						0.69		
Num 10						0.82		
Read 1							0.78	
Read 2							0.92	
Read 3							0.93	
Read 4							0.86	
Read 5							0.65	

Read 6	0.84	
Writ 1		0.71
Writ 2		0.82
Writ 3		0.92
Writ 4		0.91
Writ 5		0.92
Writ 6		0.84

Note. Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

Supplementary Table 3.14. Factor loadings from confirmatory factor analysis in Lao PDR – high maternal education

	F1	F2	F3	F4	F5	F6	F7	F8
Comm 1	0.54							
Comm 2	1.00							
Comm 3	0.87							
Comm 4	0.81							
Cult 1		0.57						
Cult 2		0.76						
Cult 3		0.80						
Cult 4		0.86						
Cult 5		0.64						
Soc 1			0.57					
Soc 2			0.64					
Soc 3			0.69					
Soc 4			0.57					
Soc 5			0.74					
Soc 6			0.80					
Soc 7			0.75					
Soc 8			0.56					
Soc 9			0.08					
Soc 10			0.73					
Soc 11			0.60					
Soc 12			-0.13					
Soc 13			0.61					
Soc 14			0.25					
Persev 1				0.96				
Persev 2				0.87				
Persev 3				-0.10				
Persev 4				-0.29				
Appr 1					0.65			
Appr 2					0.76			
Appr 3					0.51			
Appr 4					0.65			
Appr 5					0.46			
Appr 6					0.73			
Appr 7					0.59			
Num 1						0.71		
Num 2						0.77		
Num 3						0.67		
Num 4						0.81		
Num 5						0.76		
Num 6						0.70		
Num 7						0.81		
Num 8						0.79		
Num 9						0.76		
Num 10						0.84		
Read 1							0.79	
Read 2							0.95	
Read 3							0.94	
Read 4							0.89	

Read 5	0.57	
Read 6	0.80	
Writ 1		0.67
Writ 2		0.89
Writ 3		0.94
Writ 4		0.94
Writ 5		0.89
Writ 6		0.86

Note. Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

Appendix D: Published journal article 2

SSM - Population Health 11 (2020) 100613

Contents lists available at ScienceDirect

SSM - Population Health

journal homepage: http://www.elsevier.com/locate/ssmph

Measuring early child development in low and middle income countries: Investigating the validity of the early Human Capability Index

Alanna Sincovich [a,b,*], Tess Gregory [a,b], Cristian Zanon [c], Daniel D. Santos [d], John Lynch [a,e], Sally A. Brinkman [a,b]

[a] School of Public Health, University of Adelaide, Level 9, Adelaide Health and Medical Science Building, 57 North Terrace, Adelaide, South Australia, 5005, Australia
[b] Telethon Kids Institute, University of Western Australia, Level 15, 31 Flinders St., Adelaide, South Australia, 5000, Australia
[c] Department of Developmental Psychology and Personality, Universidade Federal Do Rio Grande Do Sul, Ramiro Barcelos, 2600, Porto Alegre, RS, CEP 90035-003, Brazil
[d] Department of Economy, University of São Paulo, Av. Prof. Luciano Gualberto, 908, Butantã, São Paulo, 05508-010, Brazil
[e] Population Health Sciences, Bristol Medical School, University of Bristol, First Floor, 5 Tyndall Avenue, Bristol, BS8 1UD, United Kingdom

A R T I C L E I N F O

Keywords:
Early child development
Early human capability index
Low and middle income countries
Population monitoring
Program evaluation
Sustainable development goals

A B S T R A C T

Inclusion of early child development in the United Nations Sustainable Development Agenda raises issues of how this goal should be monitored, particularly in low resource settings. The aim of this paper was to explore the validity of the early Human Capability Index (eHCI); a population measure designed to capture the holistic development of children aged 3–5 years. Convergent, divergent, discriminant and concurrent validity were examined by exploring the associations between eHCI domains and child (sex, age, stunting status, preschool attendance) and family (maternal education, home learning environment) characteristics. Analyses were repeated using data from seven low and middle income countries: Brazil (n = 1810), China (n = 11421), Kiribati (n = 8339), Lao PDR (n = 7493), Samoa (n = 12191), Tonga (n = 6214), and Tuvalu (n = 549). Correlations and linear regressions provide evidence that within these country samples, the tool is capturing the aspects of early child development that it was designed to measure. Although the tool was intended to measure development of children aged 3–5 years, results suggest it can be validly applied to children aged 2–6 years. The eHCI is free, requires minimal implementation resources, captures development across domains and abilities, and is designed to allow cultural and contextual concepts to be included. The eHCI appears psychometrically robust in diverse country contexts and could enable evaluation of early years policies and programs, as well as monitoring of children's development to track progress towards the Sustainable Development Agenda.

1. Introduction

Monitoring children's outcomes is key to improving understanding of the early determinants of health and development because it helps identify the supports required to enable children to reach their developmental potential (Young, 2007). Tracking child health and development in low and middle income countries (LMICs) is a challenge due to lack of appropriate tools and capacity to implement measurement. The early Human Capability Index (eHCI) was designed to measure holistic development in children aged 3–5 years, be feasible for use in low resource settings, and capture locally-relevant early child development (ECD) (Sincovich et al., 2019). This paper explores the convergent, divergent, discriminant, and concurrent validity (Fig. 1) of the eHCI in several LMICs, highlighting how the tool could enable ECD

Abbreviations: AEDC, Australian Early Development Census; ASQ, Ages and Stages Questionnaire; CI, confidence interval; CREDI, Caregiver Reported Early Development Instrument; EAP-ECDS, East Asia Pacific Early Child Development Scales; ECD, early child development; EDI, Early Development Instrument; eHCI, early Human Capability Index; IDELA, International Development and Early Learning Assessment; LMICs, low and middle income countries; MICS-ECDI, Multiple Indicator Cluster Survey Early Childhood Development Index; MODEL, Measure of Development and Early Learning; PDR, People's Democratic Republic; SDG, Sustainable Development Goal.

* Corresponding author. School of Public Health, University of Adelaide, Level 9, Adelaide Health and Medical Science Building, 57 North Terrace, Adelaide, South Australia, 5005, Australia.

E-mail addresses: alanna.sincovich@adelaide.edu.au (A. Sincovich), tess.gregory@adelaide.edu.au (T. Gregory), crstn.zan@gmail.com (C. Zanon), ddsantos@fearp.usp.br (D.D. Santos), john.lynch@adelaide.edu.au (J. Lynch), sally.brinkman@adelaide.edu.au (S.A. Brinkman).

https://doi.org/10.1016/j.ssmph.2020.100613
Received 23 February 2020; Received in revised form 18 April 2020; Accepted 4 June 2020
Available online 7 June 2020

measurement in these contexts.

1.1. *The SDG challenge: characteristics of a useful measure of ECD in LMICs*

Sustainable Development Goal (SDG) 4.2 is focused on ECD. To track progress against this target, countries are seeking population measures to better understand their children's early health and development. In addition to being psychometrically robust, such a measure needs to be cost-effective, therefore fees to use the tool, enumerator training required, and administration time need to be minimal (Fernald, Prado, Kariger, & Raikes, 2017). It should cover a range of development as well as levels of ability, and importantly, be sensitive enough to detect changes in children's capabilities (Mustard, 2007). Further, such a tool should be adaptable across diverse cultures and contexts so that it not only accurately reflects children's abilities, but also captures locally-relevant and culturally-influenced aspects of development to inform local policy and practice (Gove & Black, 2016).

The selection of a tool for the purposes of measuring ECD requires compromise among different priorities and measurement ideals determined by the aims of assessment, the age range of children, and any financial and logistical constraints. For instance, although the direct assessment of ECD (Pisani, Borisova, & Dowd, 2018) is often argued to produce scores with less bias than those through a measure of adult report (Bennetts, Mensah, Westrupp, Hackworth, & Reilly, 2016), this method of assessment requires highly trained enumerators, is more time consuming to conduct, more costly to implement and is infeasible for whole-of-population implementation (Snow & Van Hemel, 2008). Although adult report measures are generally more cost-effective as they are quick and simple to administer and do not rely on developmental expertise to be delivered (Ertem et al., 2017), often such tools are based on developmental milestones with a pass or fail outcome, lacking the sensitivity required to detect changes in development over time (Mustard, 2007). Although tools that produce globally comparable results (McCoy et al., 2016) allow for monitoring of ECD across countries, such instruments may not be aligned with local culture or early learning and development frameworks, and thus the information they produce may have limited utility locally (Keller & Kärtner, 2013). Local 'ownership' of results are crucial for local action to invest in ECD.

1.2. *Measuring ECD at a population-level*

Existing population-level measures of ECD include the Early Development Instrument (EDI) and UNICEF's Multiple Indicator Cluster Survey-Early Childhood Development Index (MICS-ECDI). The EDI was the first population-level measure of ECD to be implemented in multiple countries including Canada, the United States, Jamaica, and Australia (Janus, Harrison, Goldfeld, Guhn, & Brinkman, 2016). The 100-item checklist is completed by teachers of children in the first year of full-time school. The EDI formed the basis of the Australian Early Development Census (AEDC), a triennial national census of ECD which has captured over 1.2 million children across Australia to date (Brinkman, Gregory, Goldfeld, Lynch, & Hardy, 2014). The most widely utilized population-level measure of ECD is the MICS-ECDI. Consisting of 10 caregiver-reported items for children aged 3–4 years, the ECDI has been embedded in the MICS (i.e. a household survey) to collect globally comparable ECD information in more than 80 countries (UNICEF, 2017). Characteristics of both tools pose challenges to their use in diverse, low resource settings. For instance, the EDI licensing requirements and the specific intention of the ECDI for international comparison limit adaptation to local culture and context.

1.3. *The eHCI*

The eHCI was developed to facilitate program evaluation in Tonga by monitoring population-level child development (Brinkman & Thanh Vu, 2017; MacDonald et al., 2017). The tool has been adapted to support a range of early child education and development projects in several LMICs (Brinkman, Sincovich, & Danchev, 2016; Brinkman, Sincovich, & Than Vu, 2017; Brinkman, Sincovich, & Thanh Vu, 2017, 2017a; Santos & Zanon, 2017; Zhao et al., under review). Utilizing these data, the internal structure of the eHCI was explored, with results demonstrating a similar factor structure of nine theoretically-based developmental domains across countries (Sincovich et al., 2019). The current study examines the tool's convergent, divergent, discriminant, and concurrent validity using data from seven LMICs: Brazil, China, Kiribati, Lao People's Democratic Republic (PDR), Samoa, Tonga, and Tuvalu. We examine associations between eHCI domain scores for evidence of convergent and divergent validity; the ability of eHCI domain scores to discriminate among children of different age, sex, stunting status, preschool attendance, maternal education, and home learning environments (discriminant validity); and associations between scores on eHCI domains and direct assessment of literacy and numeracy for evidence of concurrent validity.

2. Method

2.1. *Participants*

Data from seven country samples were utilized (Table 1). Sample sizes ranged from 549 in Tuvalu to 12,191 in Samoa, with children aged between 2 and 6 years. Data were collected from 2013 to 17 through studies funded for program evaluation and population monitoring purposes by local, national, and international agencies that utilized different sampling techniques and data collection methods (Supplementary Table 1). In Brazil, data were collected from teachers for children attending all 37 public schools in a city in Southern Brazil, to identify areas of intervention and enable evaluation of ECD programs (Santos & Zanon, 2017). In China, data were collected for children in two Northern provinces selected by the China Development Research Foundation on the basis of pre-existing programs in these areas (Zhao et al., under review). Respondents were a combination of children's caregivers and teachers, with data collected to explore development across different population groups, as well as to form a baseline of ECD before further intervention. In Lao PDR, data were collected from children's caregivers to form the baseline for a randomized control trial. Five Northern provinces were selected by the Government of Lao PDR based on high levels of poverty, and 14 districts within these provinces

Convergent validity is evident when measures of theoretically related constructs are highly correlated.

Divergent validity is evident when measures of theoretically less related constructs are not highly correlated.

Discriminant validity is evident when measures produce scores that differ according to participant characteristics, such as sex and age, as theoretically expected.

Concurrent validity is evident when measures of constructs are highly correlated with previously established measures of the same construct.

Fig. 1. Types of validity explored in current study.

Table 1
Sample characteristics for data collected in 7 LMICs from 2013 to 2017.

	Brazil n (%)	China n (%)	Kiribati n (%)	Lao PDR n (%)	Samoa n (%)	Tonga n (%)	Tuvalu n (%)
Sex							
Female	855 (47.2)	5338 (46.7)	3915 (46.9)	3669 (49.0)	5898 (48.4)	2967 (47.7)	277 (50.5)
Missing	102 (5.6)	496 (4.3)	155 (1.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Age							
2 years	57 (3.1)	54 (0.5)	948 (11.4)	1410 (18.8)	1159 (9.5)	13 (0.2)	1 (0.2)
3 years	571 (31.5)	3396 (29.7)	2185 (26.2)	1749 (23.3)	4377 (35.9)	1609 (25.9)	163 (29.7)
4 years	760 (42.0)	3329 (29.1)	2136 (25.6)	1867 (24.9)	4616 (37.9)	2058 (33.1)	180 (32.8)
5 years	420 (23.2)	3360 (29.4)	2013 (24.1)	1599 (21.3)	2039 (16.7)	2038 (32.8)	195 (35.5)
6 years	0 (0.0)	1226 (10.7)	173 (2.1)	16 (0.2)	0 (0.0)	443 (7.1)	7 (1.3)
Missing	2 (0.1)	56 (0.5)	884 (10.6)	852 (11.4)	0 (0.0)	53 (0.9)	3 (0.5)
Stunted							
Yes	–	1602 (14.0)	1935 (23.2)	3652 (48.7)	2332 (19.1)	374 (6.0)	132 (24.0)
Missing	–	1342 (11.8)	1401 (16.8)	852 (11.4)	0 (0.0)	91 (1.5)	98 (17.9)
Preschool							
Attended	1810 (100.0)	9159 (80.2)	7665 (91.9)	1738 (23.2)	4657 (38.2)	2701 (43.5)	498 (90.7)
Missing	0 (0.0)	86 (0.8)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Maternal Education							
No school	–	307 (2.7)	–	2265 (30.2)	–	–	–
Started Primary	–	1242 (10.9)	714 (8.6)	2045 (27.3)	–	–	–
Finished Primary	–	3372 (29.5)	1438 (17.2)	2182 (29.1)	222 (1.8)	81 (1.3)	47 (8.6)
Started Secondary	–	3563 (31.2)	2319 (27.8)	–	603 (4.9)	2399 (38.6)	89 (16.2)
Finished Secondary	–	1358 (11.9)	2710 (32.5)	754 (10.1)	10037 (82.3)	2621 (42.2)	107 (19.5)
Tertiary	–	751 (6.6)	785 (9.4)	242 (3.2)	1329 (10.9)	1107 (17.8)	117 (21.3)
Missing	–	828 (7.2)	373 (4.5)	5 (0.1)	0 (0.0)	0 (0.0)	189 (34.4)
Home learning environment							
Read to	–	–	3930 (47.1)	2149 (28.7)	2424 (19.9)	3566 (57.4)	297 (54.1)
Missing	–	–	6 (0.1)	0 (0.0)	3470 (28.5)	704 (11.3)	164 (29.9)
Tell story	–	–	5886 (70.6)	1903 (25.4)	2856 (23.4)	3836 (61.7)	337 (61.4)
Missing	–	–	6 (0.1)	0 (0.0)	3470 (28.5)	702 (11.3)	165 (30.1)
Sing/dance	–	–	4964 (59.5)	2215 (29.6)	5043 (41.4)	5025 (80.9)	353 (64.3)
Missing	–	–	6 (0.1)	0 (0.0)	3470 (28.5)	700 (11.3)	165 (30.1)
Play	–	–	4774 (57.2)	1341 (17.9)	4800 (39.4)	5180 (83.4)	359 (65.4)
Missing	–	–	6 (0.1)	0 (0.0)	3470 (28.5)	699 (11.2)	165 (30.1)
Go outside	–	–	3455 (41.4)	1775 (23.7)	4722 (38.7)	5174 (83.3)	338 (61.6)
Missing	–	–	6 (0.1)	0 (0.0)	3470 (28.5)	700 (11.3)	165 (30.1)
Name, count, draw	–	–	3380 (40.5)	2280 (30.4)	2720 (22.3)	3152 (50.7)	309 (56.3)
Missing	–	–	6 (0.1)	0 (0.0)	3470 (28.5)	701 (11.3)	166 (30.2)
Books in home	–	–	3948 (47.3)	–	6253 (51.3)	4129 (66.4)	–
Missing	–	–	267 (3.2)	–	0 (0.0)	11 (0.2)	–
Total n	1810	11421	8339	7493	12191	6214	549

Note. Children's height and weight, maternal education, and home learning environment data were not collected in Brazil. Home learning environment data were not collected in China. Information regarding books in the home was not collected in Laos PDR and Tuvalu. In Samoa, Tonga, and Tuvalu, when responding to the maternal education item, respondents could select only one response option pertaining to primary school and so the proportions represented against the 'finished primary school' category for these countries may include a combination of children for whom their mother either started or finished primary school.

were selected on the basis of presence of a district level education office. All villages within selected districts in which at least 20 children resided were sampled, and random sampling methods were used to select 20 households in each village for data collection (Brinkman, Sincovich, & Danchev, 2016). In Kiribati, Samoa, Tonga, and Tuvalu, a census approach was employed with data collection seeking to capture information from a combination of caregivers and teachers for every child aged 3–5 years nationally (Brinkman, Sincovich, & Danchev, 2016; Brinkman, Sincovich, & Than Vu, 2017; Brinkman, Sincovich, & Than Vu, 2017; 2017a).

2.2. *Measures*

2.2.1. *The eHCI*

The eHCI is unlicensed and free to use, requires minimal enumerator training, and can be completed quickly by an adult familiar with the child. Thus, the eHCI can be implemented feasibly across large populations in low resource settings. The tool captures both positive and negative aspects of how a child is developing, rather than developmental delay only. The eHCI places children on a developmental continuum which improves ability to detect changes in development over time and/or through intervention. Further, the eHCI was designed to be adapted to local culture and context for a range of purposes, including population monitoring, evaluation of early years policies and programs, and longitudinal studies seeking to predict children's future capabilities.

The eHCI includes approximately 60 items (ranging from 56 in Lao PDR to 66 in Tuvalu) designed to measure ECD across nine domains: Physical Health, Verbal Communication, Cultural Knowledge, Social and Emotional Skills, Perseverance, Approaches to Learning, Numeracy, Reading, and Writing. The eHCI underwent a local adaptation process to ensure the tool's content and face validity in each country. Thus, although many items are similar across different adaptations of the eHCI, some items and domains differ across countries. For instance, the Perseverance domain is measured by the same four items across all adaptations of the instrument, while the Physical Health domain varies from 2 items in Brazil to 5 items in Kiribati and Tuvalu, while the Laotian version of the eHCI does not capture physical health (each country adaptation is available in Sincovich et al., 2019). In each country, all items are applied to all children (i.e. rather than a sub-group of items for different age groups). Response options for each item are binary ("yes"/"no", "able"/"unable", "can do already"/"cannot do yet"). Most items are positively worded so that the "yes"/"able"/"can do already" responses were scored as 1, and the "no"/"unable"/"cannot do yet" responses were scored as 0. A small number of items (ranging from 4 in Kiribati and Lao PDR to 6 in Tonga) are negatively worded and were reverse-scored. Individual item scores were averaged so that children received a score for each domain ranging from 0 to 1, with higher scores indicative of better development.

2.2.2. Demographic characteristics

Children's demographic characteristics were collected alongside the eHCI in each country, including child age, sex, previous or current preschool attendance, and their mother's highest level of education (except in Brazil).

2.2.3. Stunting

Children's height and weight were also recorded at the time of data collection in each country (except in Brazil). Anthropometric measures were converted into World Health Organization Child Growth Standards height-for-age z-scores, and stunting was defined as a height-for-age z-score < -2 (WHO Multicentre Growth Reference Study Group, 2006).

2.2.4. Home learning environment

Information about caregiver-child engagement in six types of learning activities at home was also collected alongside the eHCI in each country (except in China and Brazil). These binary questions ("yes"/"no") were based on items from the Multiple Indicator Cluster Surveys (MICS) questionnaire for children under five (UNICEF, 2019) and asked if, in the last 3 days (or in the last 7 days in Lao PDR and Tonga), a member of the household aged 15 years or above had: read books or looked at picture books with the child; told stories to the child; sang songs or danced with the child; played with the child; took the child outside of the home; and named, counted or drew with the child. In Kiribati, Samoa, and Tonga, respondents were also asked if there were reading materials in the home (also based on the MICS with a "yes"/"no" response option).

2.2.5. Literacy and numeracy

Concurrent to administration of the eHCI in Lao PDR, children's literacy and numeracy were measured via direct assessment. These assessments were based on items from the Early Grade Reading Assessment (EGRA) and the Early Grade Maths Assessment (EGMA), and adapted to the context in Lao PDR (Brinkman, Sincovich, & Danchev, 2016). The EGRA and EGMA have been adapted for use in a number of countries, with the reliability and validity of the EGRA in particular, well established (Dubeck, Gove, & Alexander, 2016). In Lao PDR, children were assessed on six aspects of early literacy and eight aspects of early numeracy (Supplementary Table 2). Correct item responses were scored as 1, and incorrect responses were scored as 0. For domains measured via multiple items, individual item scores were averaged so that children received a score for each domain, ranging from 0 to 1, with higher scores indicative of better literacy and numeracy. All literacy domain scores were averaged to provide a total literacy score, and all numeracy domain scores were averaged to provide a total numeracy score.

2.3. Statistical analysis

Convergent and divergent validity were tested by examining patterns of correlation (Spearman's rho) amongst eHCI domains in each country. We predicted that the strongest associations would be observed between eHCI domains measuring children's literacy and numeracy, with smaller correlations between remaining domains. The ability of eHCI domains to discriminate amongst children's development by a range of child and family characteristics was tested in each country. Discriminant ability according to child age (2–6 years), child sex (female, male), child stunting status (yes, no), child preschool attendance (yes, no), maternal education (no school, started primary school, finished primary school, started secondary school, finished secondary school, tertiary education), and home learning environment items (yes, no) were examined using linear regressions. Children with missing data were excluded from relevant analyses. Density plots for eHCI domains were generated to further explore the distribution of domain scores by child and family characteristics in each country, and unstandardized regression coefficients were graphed to examine associations amongst eHCI domains and child and family characteristics in each country. We expected that higher domain scores would be observed among females, older children, those not stunted, children who attended preschool, children of more educated mothers, and children with learning opportunities at home. Concurrent validity was tested by exploring correlations (Spearman's rho) among eHCI domains and literacy and numeracy direct assessment in Lao PDR. We hypothesized that the strongest associations would be observed between direct assessment scores and eHCI domains measuring literacy and numeracy, with smaller correlations between direct assessment scores and remaining eHCI domains.

3. Results

Correlations amongst eHCI domains are shown in Table 2. The largest correlations were observed amongst Numeracy, Reading, and Writing domains in all countries (ranging from $r_s = 0.54$ in Lao PDR to $r_s = 0.85$ in Tuvalu between Numeracy and Reading). Smaller correlations were observed between Physical Health and other domains in Brazil and China ($r_s = 0.16$ with Writing in China), and between Perseverance and other domains in Kiribati, Lao PDR, Samoa, and Tonga ($r_s = 0.08$ with Approaches to Learning in Samoa).

Mean eHCI domain scores according to child and family characteristics are presented in Supplementary Tables 3–8. Regression coefficients demonstrating associations among eHCI domain scores and these characteristics are shown in Supplementary Table 9, with Figs. 2–3 providing examples for Tonga and Lao PDR. Data for remaining countries are shown in Supplementary Figs. 1–5. Results show a positive association between child age and eHCI domain scores in all countries. Generally, the largest differences between older and younger children were found on Numeracy, Reading, and Writing domains, with smaller differences on remaining domains (e.g. a year increase in age was associated with a score 0.23 (95% CI: 0.22, 0.24) points higher on the Writing domain and 0.04 (95% CI: 0.03, 0.04) points higher on Social and Emotional Skills in Tonga). Fig. 4 shows a visual representation of these age gradients for Numeracy and Social and Emotional Skills domains in each country. On average, girls scored slightly higher than boys across domains, except in Brazil. This association was strongest in Tuvalu (e.g. female sex was associated with a score 0.10 (95% CI: 0.06, 0.14) points higher on Social and Emotional Skills and 0.11 (95% CI 0.05, 0.17) points higher on Reading). A negative association between child stunting and eHCI domain scores was observed, with stunted children scoring lower across domains in all countries, compared to children not stunted. Differences in scores tended to be larger on Numeracy, Reading, and Writing domains, and smaller on Social and Emotional Skills and Perseverance domains (e.g. being stunted was associated with a score 0.14 (95% CI: -0.22, -0.07) points lower on the Writing domain and 0.02 (95% CI: -0.07, 0.03) points lower on Social and Emotional Skills in Tuvalu), though this association was less clear in Tonga whereby prevalence of stunting was low. Children who attended preschool had better development than those who did not, with the largest differences in scores on Numeracy, Reading, and Writing domains (e.g. preschool attendance was associated with a score 0.44 (95% CI: 0.43, 0.45) points higher on the Numeracy domain and 0.17 (95% CI: 0.16, 0.19) points higher on Approaches to Learning in China). Fig. 5 demonstrates these differences in eHCI scores for children who did and did not attend preschool on the Numeracy and Social and Emotional Skills domains in each country.

Results also demonstrate a positive association between maternal education and eHCI domain scores, with children of more educated mothers scoring slightly higher across eHCI domains relative to children of less educated mothers in all countries but Tuvalu (e.g. a category increase in maternal education was associated with a score 0.09 (95% CI: 0.07, 0.10) points higher on the Reading domain in Samoa). A

Table 2
Convergent and divergent validity: correlations among scores on eHCI domains in 7 LMICs.

	Phys	Comm	Cult	Soc	Persev	Appr	Num	Read
Brazil								
Comm	.13							
Cult	.10	.37						
Soc	.08	.26	.19					
Persev	.14	.38	.32	.43				
Appr	.12	.43	.30	.29	.29			
Num	.16	.47	.50	.41	.34	.34		
Read	.15	.37	.38	.31	.24	.67	.67	
Writ	.15	.32	.43	.15	.29	.22	.63	.66
China								
Comm	.20							
Cult	.32	.48						
Soc	.32	.43	.70					
Persev	.12	.28	.34	.45				
Appr	.23	.54	.57	.54	.34			
Num	.27	.40	.47	.45	.25	.47		
Read	.42	.39	.47	.45	.32	.44	.67	
Writ	.16	.29	.32	.32	.26	.31	.62	.66
Kiribati								
Comm	.40							
Cult	.46	.61						
Soc	.42	.53	.71					
Persev	.21	.21	.27	.30				
Appr	.31	.39	.48	.48	.21			
Num	.43	.53	.61	.52	.24	.46		
Read	.39	.47	.54	.45	.22	.41	.76	
Writ	.37	.43	.51	.42	.20	.36	.72	.70
Lao PDR								
Comm	–							
Cult	–	.31						
Soc	–	.31	.46					
Persev	–	.09	.24	.27				
Appr	–	.43	.36	.45	.14			
Num	–	.30	.42	.50	.25	.37		
Read	–	.16	.38	.41	.20	.27	.52	
Writ	–	.12	.35	.42	.21	.28	.50	.55
Samoa								
Comm	.56							
Cult	.57	.61						
Soc	.57	.64	.81					
Persev	.16	.10	.25	.22				
Appr	.45	.63	.53	.61	.08			
Num	.48	.51	.69	.70	.26	.47		
Read	.46	.54	.68	.71	.20	.51	.78	
Writ	.46	.46	.63	.64	.25	.43	.75	.77
Tonga								
Comm	.34							
Cult	.48	.51						
Soc	.46	.38	.48					
Persev	.22	.21	.21	.24				
Appr	.34	.38	.46	.36	.20			
Num	.46	.37	.49	.40	.25	.43		
Read	.39	.33	.42	.36	.25	.38	.76	
Writ	.38	.26	.40	.32	.21	.34	.76	.73
Tuvalu								
Comm	.44							
Cult	.55	.67						
Soc	.48	.51	.67					
Persev	.36	.39	.45	.56				
Appr	.29	.36	.44	.41	.32			
Num	.52	.67	.69	.54	.44	.40		
Read	.56	.66	.69	.55	.42	.43	.85	
Writ	.47	.54	.57	.44	.42	.44	.74	.76

Note. $p < .001$ for all correlations (Spearman's rho). The Lao PDR version of the eHCI does not include a Physical Health domain. Phys = Physical Health, Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

positive association was also observed between home learning activities and eHCI domain scores in each country. Of all home environment items, the largest differences in scores were observed between children who did and did not have access to books at home on Numeracy, Reading and Writing domains (e.g. having books in the home was associated with a score 0.21 (95% CI: 0.19, 0.22) points higher on the Writing domain and 0.12 (95% CI: 0.11, 0.13) points higher on Physical Health in Kiribati). Differences in scores between children who were and were not read to, told stories, and named, counted and drew with were also larger across domains, while differences in scores were smaller for remaining activities (e.g. naming, counting, or drawing was associated with a score 0.18 (95% CI: 0.16, 0.19) points higher on the Numeracy domain, while taking the child outside was associated with a score 0.16 (95% CI: 0.05, 0.08) points higher on Numeracy in Lao PDR).

Correlations between eHCI domains and literacy and numeracy direct assessment in Lao PDR are shown in Table 3. Both direct assessment scores had the largest positive correlations with eHCI Numeracy, Reading, and Writing ($r_s = 0.50$ between direct assessment and eHCI numeracy domains), and smaller correlations with remaining domains ($r_s = 0.18$ between literacy direct assessment and eHCI Perseverance).

4. Discussion

We examined the convergent, divergent, and discriminant, validity of the eHCI in Brazil, China, Kiribati, Lao PDR, Samoa, Tuvalu, and Tonga, and the concurrent validity of the tool in Lao PDR. Despite different sampling methods and items used across countries, the overall pattern of results suggests the eHCI provides valid measurement of ECD in these contexts.

4.1. Evidence for validity of the eHCI

The strongest associations were observed amongst eHCI Numeracy, Reading, and Writing domains in all countries. Literacy and numeracy are often intertwined; indeed some ECD tools combine items measuring these skills into one domain because of their strong relationship (Janus et al., 2016). Overall results provided evidence for the convergent and divergent validity of the eHCI.

eHCI domain scores discriminated between children's development according to age and sex in all countries. This is consistent with results of other measures of ECD in LMICs, including the East Asia Pacific Early Child Development Scales (EAP-ECDS), a direct assessment of development in children aged 3–5 years, across Cambodia, China, Mongolia, Timor-Leste, Papua New Guinea, and Vanuatu (Rao, Sun, et al., 2018). Although the eHCI was designed to measure ECD in children aged 3–5 years, results suggest it can be validly applied to children aged 2–6 years. Stunting impairs children's development with effects particularly detrimental to cognitive abilities (Alderman & Fernald, 2017). In all countries eHCI scores were lower among stunted children, with the largest differences in scores on cognitive domains capturing literacy and numeracy. This is also aligned with results of other tools, such as the Caregiver Reported Early Development Instrument (CREDI) for children aged 0–2 years across 17 low, middle, and high middle income countries (McCoy, Waldman, Team, & Fink, 2018). A positive association between early education and ECD was also observed, with differences in scores between children who did and did not attend preschool largest on Numeracy, Reading, and Writing. In contrast, results from a direct assessment, Measure of Development and Early Learning (MODEL), of children aged 4–8 years in Tanzania did not find a relationship between pre-primary education and children's development as was expected (Raikes et al., 2019), whereas cognitive, language and socio-emotional development as measured by the EAP-ECDS was higher amongst children aged 3–5 years who had attended early education in Cambodia, China, Mongolia, and Vanuatu (Rao et al., 2018). Higher eHCI scores were generally observed among children born to better educated mothers, except in Tuvalu. This sample represents poor families from a

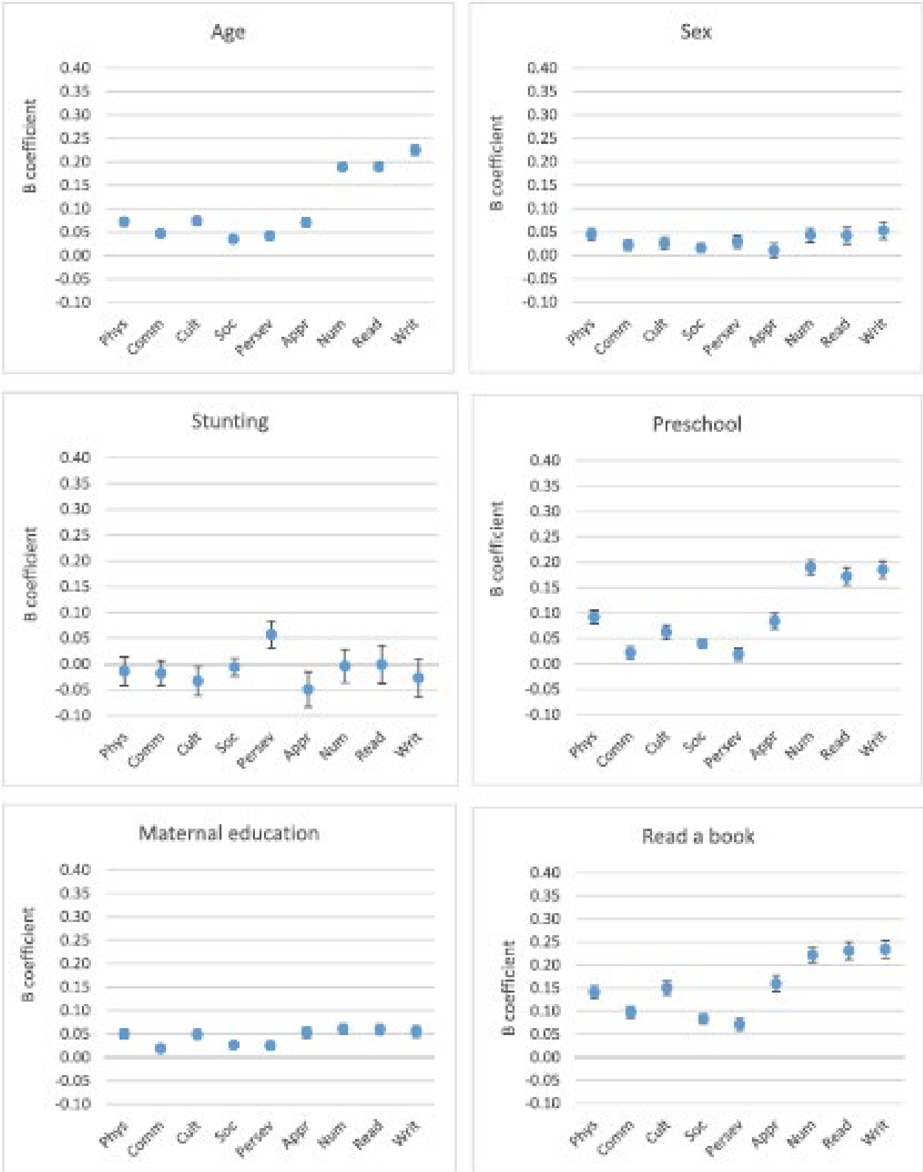
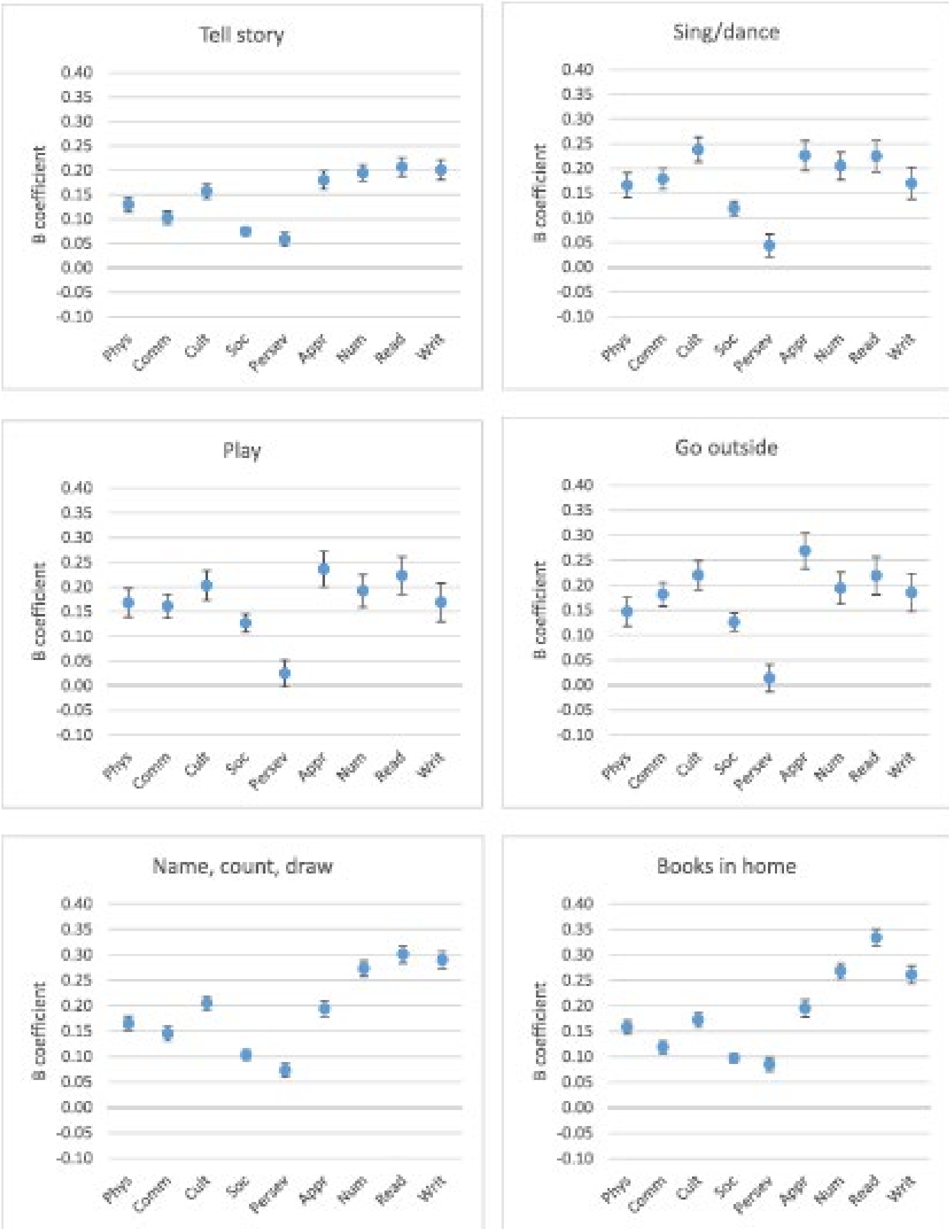


Fig. 2. Discriminant validity: linear regression coefficients and 95% confidence intervals of eHCI domain scores on demographic and contextual variables in Tonga (n = 6214, 2013/14).
Note. Female, stunted, attended preschool, home learning environment yes = 1; male, not stunted, did not attend preschool, home learning environment items no = 0. Phys = Physical Health, Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

small island nation with a community-based economy. It is likely that advantages of maternal education are transferred to the whole community rather than biological children only and thus the association may not be seen. Similar results were observed when using the International Development and Early Learning Assessment (IDELA) in Ethiopia. A direct assessment of children aged 3–6 years, maternal education did not predict IDELA scores in this context (Wolf et al., 2017). Stimulating home environments have positive effects on ECD and this pattern was also evident among eHCI scores in all countries, with the strongest associations observed for literacy and numeracy. This finding was aligned with analysis of the CREDI in Brazil amongst children aged 0–2 years (Altafim et al., 2018).

The strongest associations with direct assessment of literacy and numeracy in Lao PDR were observed for eHCI Numeracy, Reading, and Writing domains as expected. Few similar ECD measures have published concurrent validity evidence with the IDELA an exception. Specifically, the IDELA and the Ages and Stages Questionnaire (ASQ) were used to measure development of children aged 4–5 years in Bangladesh, with



Note. Female, stunted, attended preschool, home learning environment yes = 1; male, not stunted, did not attend preschool, home learning environment items no = 0. Phys=Physical Health, Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing.

Fig. 2. (*continued*).

results indicating medium correlations between respective IDELA and ASQ domains when examining children's literacy (r = 0.36) and larger correlations between domains measuring numeracy (r = 0.48) (Pisani et al., 2018). Concurrent validity of eHCI scores were stronger than those reported for the IDELA. Together, results provide evidence for the concurrent validity of the eHCI in this setting.

4.2. *Implications of findings*

Inclusion of ECD in the Sustainable Development Agenda raises issues regarding how this goal should be monitored. Relative to other measures of ECD, the eHCI can be implemented feasibly in low resource settings, captures development across domains and abilities, and produces information relevant to local policy and practice. Results demonstrate that the eHCI discriminated between the development of children and captured the intended aspects of ECD within a range of

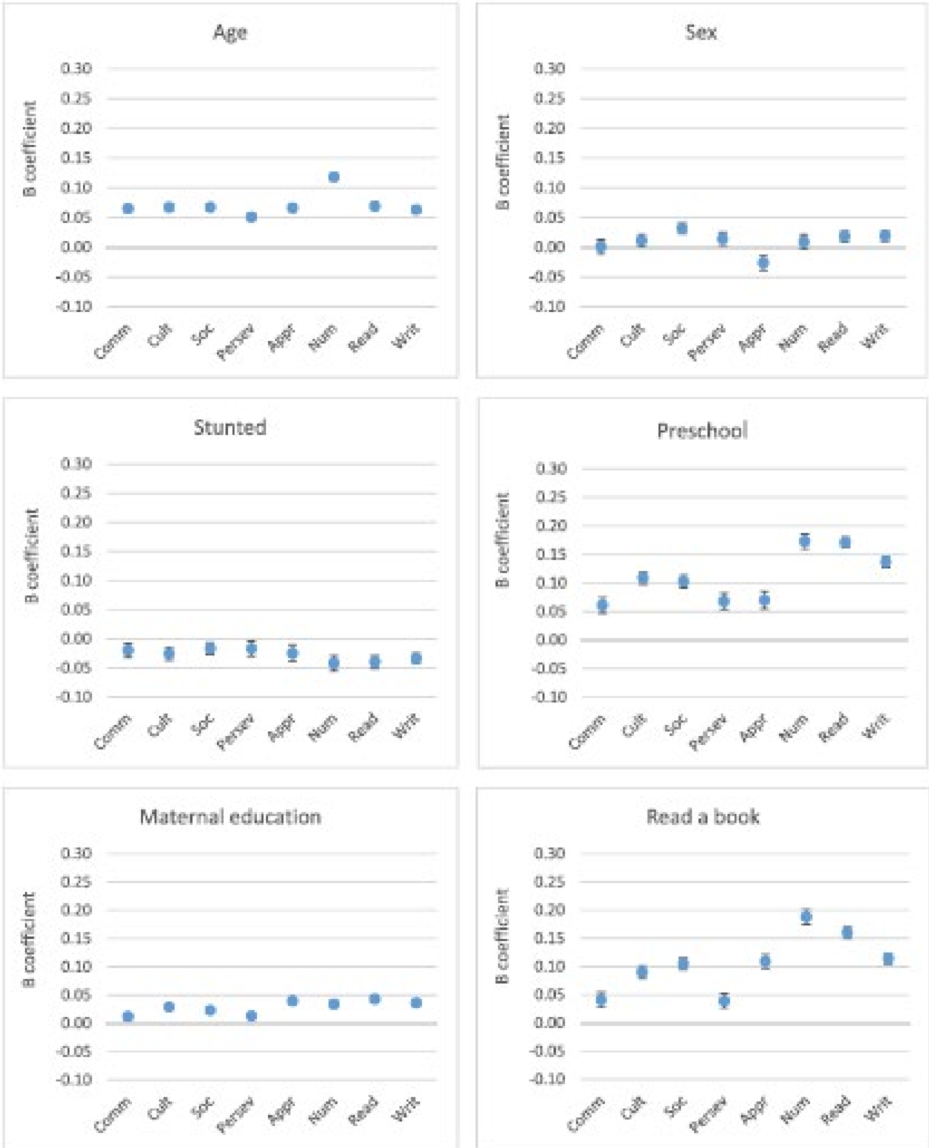
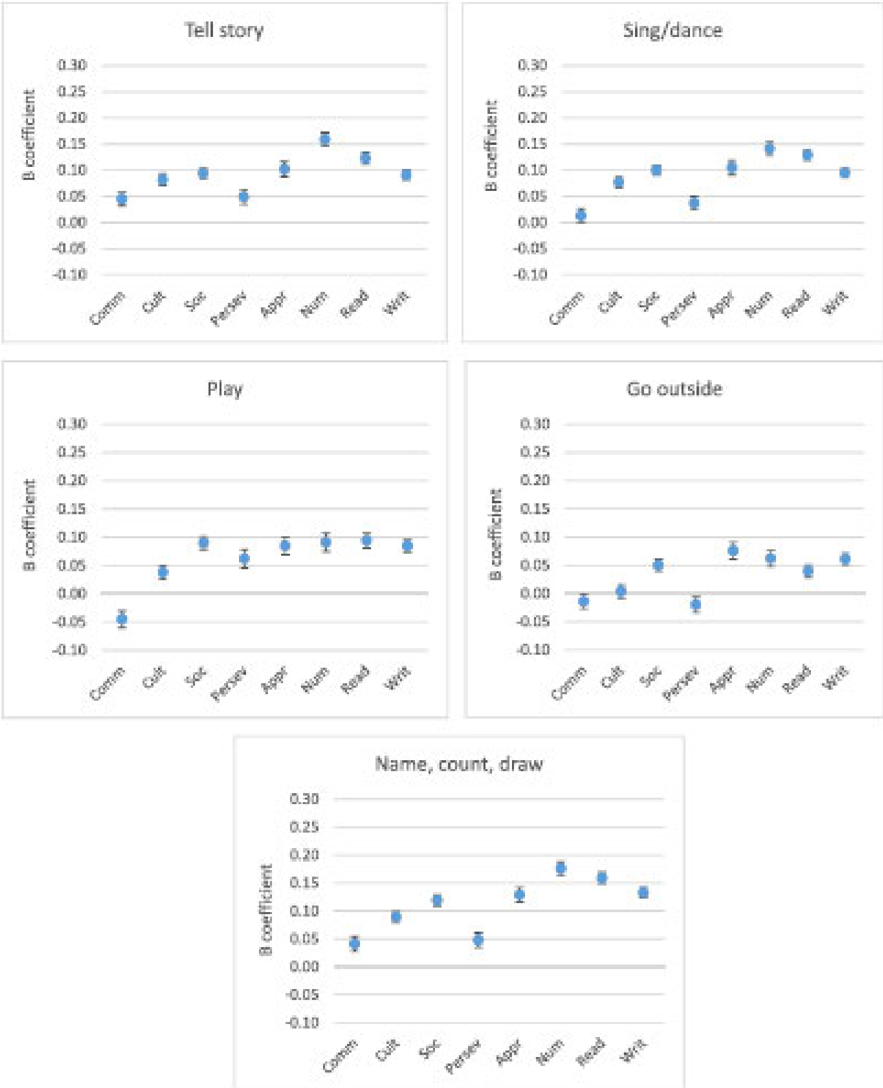


Fig. 3. Discriminant validity: linear regression coefficients and 95% confidence intervals of eHCI domain scores on demographic and contextual variables in Lao PDR (n = 7493, 2015/16).
Note: The Lao PDR version of the eHCI does not include a Physical Health domain. Information regarding books in the home were not collected in Lao PDR. Female, stunted, attended preschool, home learning environment items yes = 1; male, not stunted, did not attend preschool, home learning environment items no = 0. Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

LMICs. Together with previous research (Sincovich et al., 2019), findings indicate the tool can provide valid measurement of ECD in diverse contexts.

Although associations observed among eHCI domains and with child and family characteristics were relatively consistent overall, some variation in results across countries highlights the context-specific nature of ECD measurement. For instance, varied strength of the association between caregiver-child interactions and eHCI scores could be a reflection of cultural and/or contextual factors, including those related to caregiving practices and early years service provision. This lends support to the need for ECD measures to produce information that reflects local settings. Global comparability and cultural neutrality are the focus for tracking progress toward SDG 4.2 (Richter et al., 2019), however this approach will not have the sensitivity to capture change in ECD as a result of local interventions or policy shifts, and will not reflect aspects of ECD important to the local context. Experts in ECD measurement continue to be challenged by striking a balance between producing globally comparable data and producing information relevant to local policy and practice (Rao, Mirpuri, Sincovich, & Brinkman, 2020).

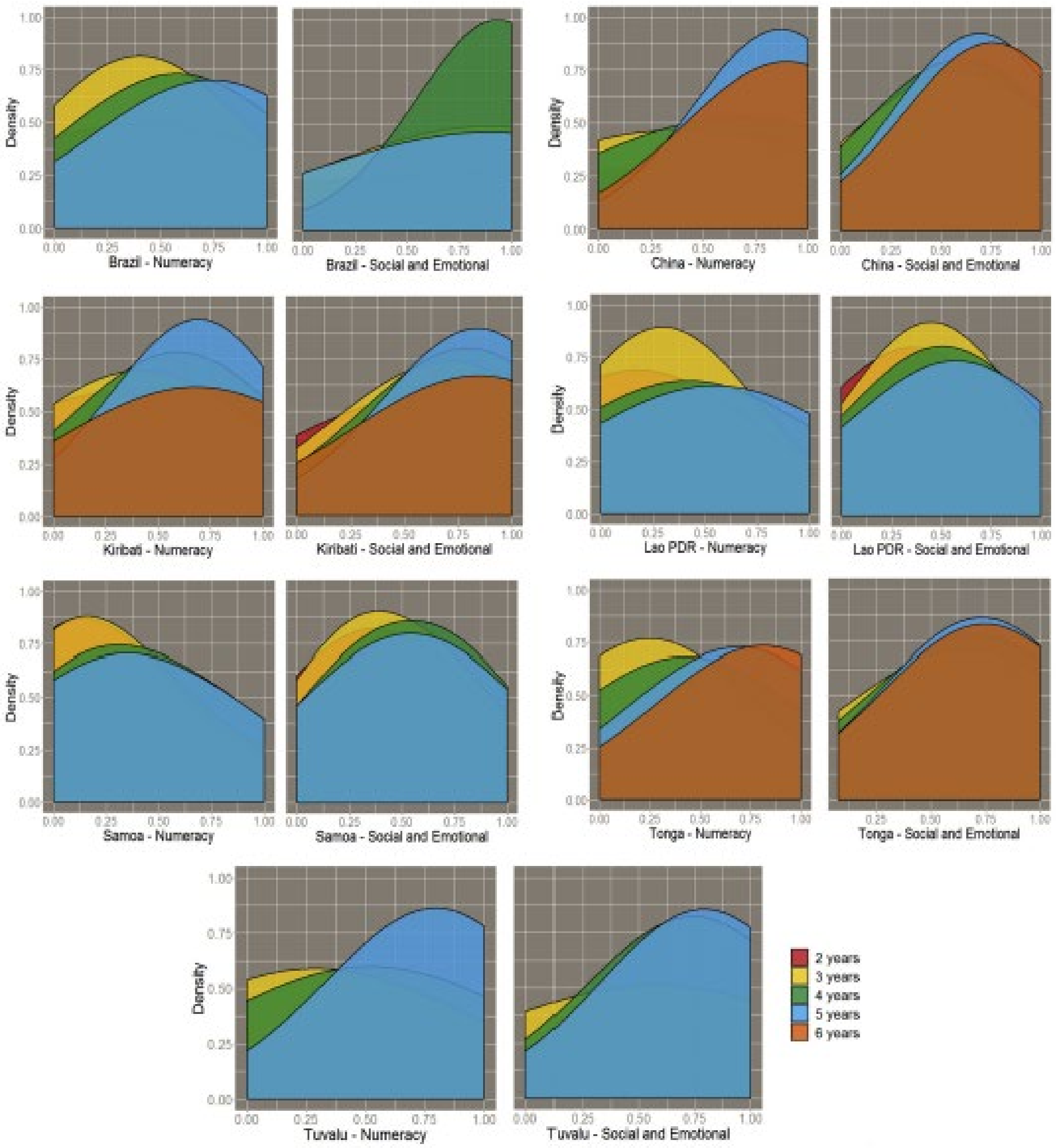


Note: The Lao PDR version of the eHCI does not include a Physical Health domain. Information regarding books in the home were not collected in Lao PDR. Female, stunted, attended preschool, home learning environment items yes = 1; male, not stunted, did not attend preschool, home learning environment items no = 0. Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing.

Fig. 3. (*continued*).

5. Conclusion

Results demonstrate the eHCI is psychometrically robust in diverse country contexts and could enable the evaluation of early years policies and programs, as well as monitoring of children's development to track progress towards the Sustainable Development Agenda in LMICs. Existing ECD measurement tools range from short adult-report tools designed for population monitoring (e.g. MICS-ECDI), to detailed, multi-domain direct assessment batteries designed to aid program evaluation. Findings indicate the eHCI is suitable for both applications. Indeed, the eHCI is a feasible and valid population monitoring measure when applied through either a census or sample approach. The eHCI is free, requires minimal implementation resources, captures development across domains and abilities, and allows local culture and context to be reflected. Results presented in this paper together with previously published evidence demonstrate that the tool is able to provide valid measurement of ECD. The next step in assessing the validity of the eHCI is exploring how well it predicts later outcomes of interest such as academic achievement and social and emotional skills.



Note. X-axis = eHCI domain score, y-axis = proportion of children e.g. in Brazil the largest proportion of 3-year-olds scored approximately 0.40 on the Numeracy domain.

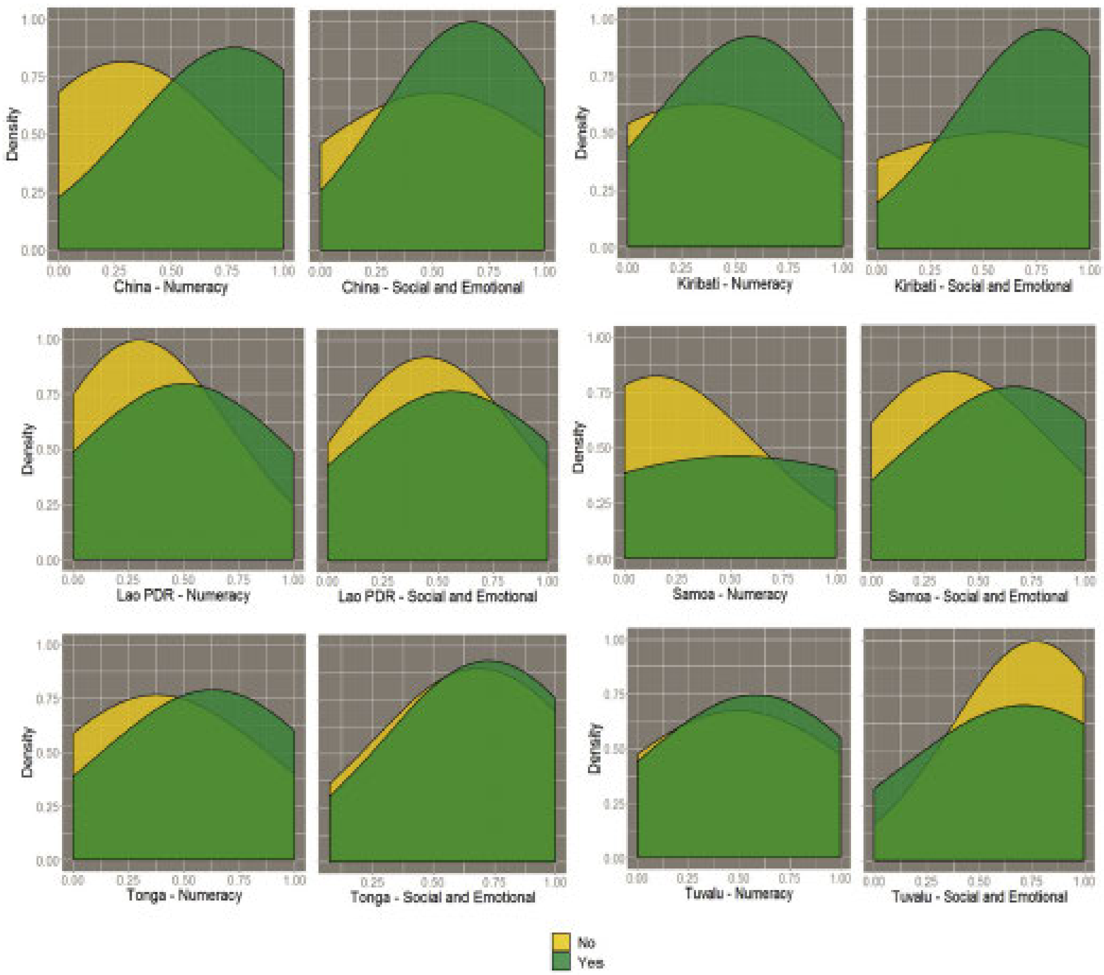
Fig. 4. Discriminant validity: distribution of eHCI Numeracy and Social and Emotional Skills domains by age (years) in 7 LMICs.
Note. X-axis = eHCI domain score, y-axis = proportion of children e.g. in Brazil the largest proportion of 3-year-olds scored approximately 0.40 on the Numeracy domain.

Declaration of competing interest

We declare no competing interests.

Role of the funding source

Funding sources supported the collection of data but had no role in data analysis or development of this article. The corresponding author



Note. X-axis = eHCI domain score, y-axis = proportion of children e.g. in China the largest proportion of children who attended preschool scored approximately 0.75 on the Numeracy domain. Results are not presented for Brazil as 100% of the sample were attending preschool.

Fig. 5. Discriminant validity: distribution of eHCI Numeracy and Social and Emotional Skills domains by preschool attendance in 7 LMICs.
Note. X-axis = eHCI domain score, y-axis = proportion of children e.g. in China the largest proportion of children who attended preschool scored approximately 0.75 on the Numeracy domain. Results are not presented for Brazil as 100% of the sample were attending preschool.

Table 3
Concurrent validity: correlations among eHCI domain scores and literacy and numeracy direct assessment scores in Lao PDR (n = 7493, 2015/16).

	Comm	Cult	Soc	Persev	Appr	Num	Read	Writ
Total numeracy score	.18	.29	.33	.22	.24	.50	.40	.37
Total literacy score	.06	.24	.26	.18	.16	.42	.39	.39

Notes: p < .05 for all correlations (Spearman's rho). Phys = Physical Health, Comm = Verbal Communication, Cult = Cultural Knowledge, Soc = Social and Emotional, Persev = Perseverance, Appr = Approaches to Learning, Num = Numeracy, Read = Reading, and Writ = Writing.

had full access to all data in the study and had final responsibility for the decision to submit for publication.

Research ethics statement

The current study is a secondary analysis of pre-existing, de-identified data and thus was deemed exempt from requiring ethical review by the University of Adelaide Human Research Ethics Committee.

CRediT authorship contribution statement

Alanna Sincovich: Conceptualization, Methodology, Formal analysis, Writing - original draft. **Tess Gregory:** Conceptualization, Methodology, Writing - review & editing. **Cristian Zanon:** Data curation, Writing - review & editing. **Daniel D. Santos:** Data curation, Writing - review & editing. **John Lynch:** Conceptualization, Methodology, Writing - review & editing. **Sally A. Brinkman:** Conceptualization, Methodology, Writing - review & editing, Supervision.

Acknowledgments

We wish to acknowledge the agencies who supported the collection of early Human Capability Index (eHCI) data which were utilized in the current study. Data collection was funded by the Fundação Maria Cecilia Souto Vidigal in Brazil, the China Development Research Foundation in China, the World Bank Group through the Early Childhood Education Project in Lao PDR, and the Global Partnership for Education through the Pacific Early Age Readiness and Learning Program in Kiribati, Samoa, Tonga, and Tuvalu.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ssmph.2020.100613.

References

Alderman, H., & Fernald, L. (2017). The nexus between nutrition and early childhood development. *Annual Review of Nutrition, 37*(1), 447–476. https://doi.org/10.1146/annurev-nutr-071816-064627.

Altafim, E. R., McCoy, D. C., Brentani, A., Escobar, A. M., Grisi, S. J., & Fink, G. (2018). Measuring early childhood development in Brazil: validation of the caregiver reported Early Development instruments (CREDI). *Jornal de Pediatria, 96*(1), 66–75. https://doi.org/10.1016/j.jped.2018.07.008.

Bennetts, S. K., Mensah, F. K., Westrupp, E. M., Hackworth, N. J., & Reilly, S. (2016). The agreement between parent-reported and directly measured child language and parenting behaviors. *Frontiers in Psychology, 7*(1710). https://doi.org/10.3389/fpsyg.2016.01710.

Brinkman, S., Gregory, T., Goldfeld, S., Lynch, J., & Hardy, M. (2014). Data resource profile: The Australian early development Index (AEDI). *International Journal of Epidemiology, 43*(4), 1089–1096. https://doi.org/10.1093/ije/dyu085.

Brinkman, S., Sincovich, A., & Danchev, P. (2016). *The status of early childhood health and development in northern Lao PDR: Baseline results from the early childhood education project.* Washington, D.C: World Bank Group. Available from: http://documents.worldbank.org/curated/en/228221488446175113/The-status-of-early-childhood-health-and-development-in-northern-Lao-PDR.

Brinkman, S., Sincovich, A., & Than Vu, B. (2017). *Early childhood development in Samoa: Baseline results from the Samoan early Human Capability Index.* Washington, D.C: World Bank Group. Available from: http://documents.worldbank.org/curated/en/363761563248716398/pdf/Main-Report.pdf.

Brinkman, S., Sincovich, A., & Thanh Vu, B. (2017a). *How are Tuvalu's children developing? Evidence-based policy recommendations for better early childhood development in Tuvalu.* Washington, D.C: World Bank Group. Available from: http://documents.worldbank.org/curated/en/373411563250298102/How-are-Tuvalu-s-children-developing-Evidence-based-policy-recommendations-for-better-early-childhood-development-of-Tuvaluan-children.

Brinkman, S., Sincovich, A., & Thanh Vu, B. (2017b). *The status of early childhood health and development in Kiribati: Results from a population wide census.* Washington, D.C: World Bank Group. Available from: http://documents.worldbank.org/curated/en/853791563251577440/pdf/The-Status-of-Early-Childhood-Health-and-Development-in-Kiribati-Results-from-a-Population-Wide-Census.pdf.

Brinkman, S., & Thanh Vu, B. (2017). *Early child development in Tonga.* Washington: D.C: World Bank Group. Retrieved from https://files.eric.ed.gov/fulltext/ED589882.pdf.

Dubeck, M. M., Gove, A., & Alexander, K. (2016). *School-based assessments: What and how to assess reading. Understanding what works in oral reading assessments.*

Ertem, I. O., Krishnamurthy, V., Mulaudzi, M. C., Sguassero, Y., Balta, H., Gulumser, O., … Forsyth, B. W. C. (2017). Similarities and differences in child development from birth to age 3 years by sex and across four countries: A cross-sectional, observational study. *Lancet Glob Health, 6*, e279–e291.

Fernald, L., Prado, E., Kariger, P., & Raikes, A. (2017). *A toolkit for measuring early childhood development in low- and middle-income countries.* Washington: D.C: World Bank Group. Retrieved from http://documents.worldbank.org/curated/en/384681513101293811/pdf/WB-SIEFECD-MEASUREMENT-TOOLKIT.pdf.

Gove, A., & Black, M. M. (2016). Measurement of early childhood development and learning under the sustainable development goals. *Journal of Human Development and Capabilities, 17*(4), 599–605. https://doi.org/10.1080/19452829.2016.1243520.

Janus, M., Harrison, L. J., Goldfeld, S., Guhn, M., & Brinkman, S. (2016). International research utilizing the early development instrument (EDI) as a measure of early child development: Introduction to the special issue. *Early Childhood Research Quarterly, 35*, 1–5.

Keller, H., & Kärtner, J. (2013). *The cultural solution of universal developmental tasks* (Vol. 3). New York, NY: Oxford University Press.

MacDonald, K., Brinkman, S., Jarvie, W., Machuca-Sierra, M., McDonall, K., Messaoud-Galusi, S., et al. (2017). *Pedagogy versus school readiness: The impact of a randomized reading instruction intervention and community-based playgroup intervention on early grade reading outcomes in Tonga.* Retrieved from http://documents.worldbank.org/curated/en/651771484764766333/pdf/WPS7944.pdf.

McCoy, D. C., Peet, E. D., Ezzati, M., Danaei, G., Black, M. M., Sudfeld, C. R., … Fink, G. (2016). Early childhood developmental status in low- and middle-income countries: National, regional, and global prevalence estimates using predictive modeling. *PLoS Medicine, 13*(6). https://doi.org/10.1371/journal.pmed.1002034.

McCoy, D. C., Waldman, M., Team, C. F., & Fink, G. (2018). Measuring early childhood development at a global scale: Evidencefrom the caregiver-reported early development instruments. *Early Childhood Research Quarterly, 45*, 58–68. https://doi.org/10.1016/j.ecresq.2018.05.002.

Mustard, J. F. (2007). Experience-based brain development: Scientific underpinnings of the importance of early child development in a global World. In M. E. Young, & L. M. Richardson (Eds.), *Early child development: From measurement to action* (pp. 43–84). Washington DC, USA: The World Bank.

Pisani, L., Borisova, I., & Dowd, A. J. (2018). Developing and validating the international development and early learning assessment (IDELA). *International Journal of Educational Research, 91*, 1–15. https://doi.org/10.1016/j.ijer.2018.06.007.

Raikes, A., Koziol, N., Janus, M., Platas, L., Weatherholt, T., Smeby, A., et al. (2019). Examination of school readiness constructs in Tanzania: Psychometric evaluation of the MELQO scales. *Journal of Applied Developmental Psychology, 62*, 122–134. https://doi.org/10.1016/j.appdev.2019.02.003.

Rao, N., Mirpuri, S., Sincovich, A., & Brinkman, S. A. (2020). Overcoming conceptual and methodological challenges in measuring early childhood development across cultures. *Lancet Child and Adolescent Health.* https://doi.org/10.1016/S2352-4642(20)30026-2. Published online February 11.

Rao, N., Richards, B., Sun, J., Weber, A., & Sincovich, A. (2018a). Early childhood education and child development in four countries inEast Asia and the Pacific. *Early Childhood Research Quarterly, 47*(2019), 169–181. https://doi.org/10.1016/j.ecresq.2018.08.011.

Rao, N., Sun, J., Richards, B., Weber, A. M., Sincovich, A., Darmstadt, G. L., et al. (2018). Assessing diversity in early childhood development in the East Asia-Pacific. *Child Indicators Research.* https://doi.org/10.1007/s12187-018-9528-5.

Richter, L., Black, M., Britto, P., Daelmans, B., Desmond, C., Devercelli, A., … Vargas-Barón, E. (2019). Early childhood development: An imperative for action and measurement at scale. *BMJ Global Health, 5*, i154–i160. https://doi.org/10.1136/bmjgh-2018-001302.

Santos, D. D., & Zanon, C. (2017). *The Brazilian experience with the eHCI: The adaptation process, reduction proceedings, and validation study.*

Sincovich, A., Gregory, T., Zanon, C., Santos, D. D., Lynch, J., & Brinkman, S. (2019). Measuring early childhood development in multiple contexts: The factor structure and reliability of the early Human Capability Index in seven low and middle income countries. *BMC Pediatrics, 19*(471). https://doi.org/10.1186/s12887-019-1852-5.

Snow, C. E., & Van Hemel, S. B. (2008). *Early childhood assessment: Why, what, and how.* Washington, DC: National Academies Press.

UNICEF. (2017). *Development of the early childhood development index in MICS surveys, MICS methodological papers.* Paper No. 6.

UNICEF. (2019). *MICS6 TOOLS. Retrieved from multiple indicator cluster surveys (MICS).* http://mics.unicef.org/tools. http://mics.unicef.org/tools.

WHO Multicentre Growth Reference Study Group. (2006). *WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development.* Geneva.

Wolf, S., Halpin, P., Yoshikawa, H., Dowd, A. J., Pisani, L., & Borisova, I. (2017). Measuring school readiness globally: Assessing the construct validity and measurement invariance of the International Development and Early Learning Assessment (IDELA) in Ethiopia. *Early Childhood Research Quarterly, 41*, 21–36.

Young, M. E. (2007). *Early child development: From measurement to action.* Washington DC., USA: The World Bank.

Zhao, J., Brinkman, S. A., Song, Y., Lu, C., Young, M. E., Zhang, Y., … Zhang, Y. (under review). Measuring early childhood development with the early human capability index (eHCI): A reliability and validity study in China *J pediatr.*

Appendix E: Supplementary materials from published journal article 2

Supplementary Table 4.1. Data collection contexts and procedures in 7 LMICs from 2013-17

	Country context	Year/s	Respondent/s	Method	Data collection purpose	Sample
Brazil	Occupies a large area on the eastern coast of South America. In 2017, population was approx. 209000 and GNI was USD8600 per capita.	2015	Preschool teachers	Pen and paper	Adaptation of the eHCI for the Brazilian context [107]	Children from one city in Southwest Brazil
China	In East Asia, the world's most populous country. In 2017, population was approx. 1.4 billion and GNI was USD8690 per capita.	2015/16	Preschool teachers and caregivers	Pen and paper	Evaluate inequality in children's outcomes across population groups[105]	Children from two provinces across Northern China
Kiribati	Comprised of 33 coral atolls and isles in the Central Pacific. In 2017, population was approx. 116000 and GNI was USD3010 per capita.	2017	Preschool teachers and caregivers	Tablet	National baseline of child development to guide policy and programs [106]	Population; aimed to collect data for all children aged 3-5 years nationally
Lao PDR	A landlocked country in the Southeast of Asia. In 2017, population was approx. 6.9 million and GNI was USD2270 per capita.	2015/16	Caregivers	Tablet	Baseline data for an RCT designed to support children's early development [102]	Children from five provinces across Northern Lao PDR
Samoa	Comprised of 2 main islands in the South Pacific. In 2017, population was approx. 196000 and GNI was USD4090 per capita.	2016	Preschool teachers and caregivers	Tablet	National baseline of child development to guide policy and programs [103]	Population; aimed to collect data for all children aged 3-5 years nationally
Tonga	An archipelago of more than 170 South Pacific islands (36 inhabited). In 2017, population was approx. 108000 and GNI was USD4010 per capita.	2013/14	Preschool teachers and caregivers	Pen and paper	Baseline data for an RCT designed to support children's school readiness [99]	Population; aimed to collect data for all children aged 3-5 years nationally
Tuvalu	An island country in the South Pacific, comprising 9 small islands. In 2017, population was approx. 11000 and GNI was USD4970 per capita.	2015	Preschool teachers and caregivers	Pen and paper	National baseline of child development to guide policy and programs [104]	Population; aimed to collect data for all children aged 3-5 years nationally

Note. GNI=Gross National Income. RCT=Randomised Control Trial. Population and Gross National Income figures sourced from World Bank (28).

Supplementary Table 4.2. Direct assessment items concurrent to the eHCI in Lao PDR (n = 7493, 2015/16)

Domain	Aspects of development	Number of items
Early literacy	Familiarity with print	3
	Sound discrimination	3
	Sound identification	5
	Letter name knowledge	20
	Listening comprehension	5
	Name writing	1
Early numeracy	Measurement vocabulary	4
	Shape naming	4
	Spatial vocabulary	4
	Verbal counting	1
	Number identification	20
	Producing a set	3
	Addition with two sets	1
	Spatial visualisation	1

Supplementary Table 4.3. Means and standard deviations of eHCI domain scores by age (years) in 7 LMICs

		n	Phys M (SD)	Comm M (SD)	Cult M (SD)	Soc M (SD)	Persev M (SD)	Appr M (SD)	Num M (SD)	Read M (SD)	Writ M (SD)
Brazil											
	2	57	.92 (.18)	.83 (.20)	.83 (.18)	.86 (.17)	.74 (.31)	.88 (.19)	.40 (.27)	.15 (.15)	.23 (.09)
	3	571	.94 (.14)	.89 (.18)	.85 (.17)	.88 (.19)	.74 (.32)	.92 (.17)	.41 (.24)	.20 (.17)	.28 (.17)
	4	760	.94 (.14)	.92 (.16)	.90 (.17)	.90 (.18)	.77 (.32)	.93 (.16)	.56 (.27)	.32 (.20)	.61 (.23)
	5	420	.95 (.13)	.93 (.14)	.93 (.14)	.88 (.20)	.77 (.33)	.92 (.17)	.71 (.26)	.45 (.21)	.78 (.17)
China											
	2	54	.73 (.26)	.76 (.25)	.70 (.30)	.62 (.27)	.45 (.33)	.73 (.27)	.37 (.23)	.25 (.23)	.17 (.24)
	3	3,396	.71 (.27)	.77 (.26)	.65 (.33)	.58 (.26)	.41 (.31)	.74 (.29)	.40 (.25)	.26 (.25)	.12 (.21)
	4	3,329	.76 (.26)	.77 (.23)	.67 (.30)	.59 (.24)	.37 (.31)	.78 (.24)	.66 (.23)	.33 (.23)	.22 (.29)
	5	3,360	.82 (.24)	.86 (.19)	.79 (.25)	.68 (.21)	.41 (.30)	.86 (.20)	.85 (.16)	.54 (.23)	.50 (.31)
	6	1,226	.84 (.24)	.90 (.18)	.86 (.21)	.74 (.19)	.49 (.26)	.90 (.18)	.89 (.15)	.66 (.24)	.60 (.31)
Kiribati											
	2	948	.54 (.30)	.60 (.30)	.53 (.33)	.62 (.27)	.40 (.25)	.66 (.29)	.31 (.27)	.32 (.29)	.24 (.31)
	3	2,185	.61 (.28)	.70 (.27)	.64 (.31)	.69 (.25)	.42 (.25)	.71 (.27)	.42 (.26)	.40 (.31)	.34 (.32)
	4	2,136	.69 (.25)	.78 (.24)	.75 (.28)	.76 (.23)	.45 (.26)	.76 (.24)	.58 (.23)	.55 (.31)	.57 (.34)
	5	2,013	.74 (.22)	.84 (.22)	.83 (.24)	.81 (.20)	.49 (.26)	.81 (.21)	.66 (.22)	.64 (.30)	.72 (.30)
	6	173	.73 (.23)	.83 (.23)	.84 (.25)	.82 (.22)	.51 (.25)	.79 (.24)	.66 (.24)	.63 (.30)	.71 (.31)
Lao PDR											
	2	1,410	-	.70 (.32)	.58 (.25)	.37 (.18)	.42 (.25)	.55 (.30)	.19 (.20)	.04 (.11)	.09 (.11)
	3	1,749	-	.85 (.24)	.69 (.20)	.45 (.19)	.49 (.27)	.65 (.28)	.31 (.21)	.09 (.17)	.12 (.13)
	4	1,867	-	.89 (.22)	.75 (.18)	.51 (.19)	.52 (.26)	.73 (.24)	.42 (.23)	.15 (.21)	.18 (.17)
	5	1,599	-	.91 (.20)	.79 (.17)	.57 (.20)	.58 (.24)	.76 (.24)	.54 (.25)	.25 (.27)	.28 (.25)
	6	16	-	-	-	-	-	-	-	-	-
Samoa											
	2	1,159	.62 (.25)	.44 (.37)	.31 (.33)	.41 (.26)	.32 (.26)	.51 (.38)	.22 (.29)	.27 (.33)	.24 (.35)
	3	4,377	.67 (.26)	.50 (.38)	.37 (.35)	.45 (.27)	.32 (.27)	.57 (.38)	.26 (.30)	.31 (.35)	.29 (.37)
	4	4,616	.72 (.27)	.62 (.36)	.48 (.37)	.54 (.28)	.38 (.30)	.65 (.37)	.39 (.34)	.45 (.38)	.43 (.40)
	5	2,039	.73 (.27)	.63 (.36)	.48 (.38)	.54 (.28)	.40 (.29)	.64 (.37)	.40 (.33)	.46 (.37)	.45 (.40)
Tonga											
	2	13	-	-	-	-	-	-	-	-	-
	3	1,609	.64 (.28)	.78 (.26)	.58 (.29)	.65 (.17)	.39 (.25)	.62 (.35)	.27 (.23)	.25 (.26)	.12 (.20)

4	2,058	.72 (.27)	.84 (.22)	.67 (.27)	.69 (.16)	.45 (.25)	.71 (.33)	.43 (.26)	.39 (.31)	.28 (.28)
5	2,038	.79 (.24)	.89 (.19)	.75 (.24)	.72 (.15)	.49 (.24)	.77 (.30)	.67 (.25)	.64 (.32)	.58 (.32)
6	443	.84 (.21)	.91 (.18)	.78 (.24)	.74 (.14)	.50 (.23)	.82 (.27)	.78 (.24)	.79 (.29)	.77 (.30)
Tuvalu										
2	1	-	-	-	-	-	-	-	-	-
3	163	.59 (.32)	.62 (.26)	.49 (.32)	.57 (.28)	.40 (.31)	.78 (.29)	.31 (.27)	.26 (.28)	.28 (.28)
4	180	.71 (.25)	.75 (.24)	.68 (.26)	.72 (.21)	.46 (.32)	.83 (.21)	.52 (.27)	.50 (.32)	.51 (.33)
5	195	.83 (.21)	.86 (.17)	.80 (.20)	.78 (.18)	.56 (.28)	.88 (.25)	.76 (.22)	.70 (.29)	.81 (.26)
6	7	-	-	-	-	-	-	-	-	-

Note: Cells where n < 20 have been omitted. The Lao PDR version of the eHCI does not include a Physical Health domain. Phys=Physical Health, Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing.

Supplementary Table 4.4 Means and standard deviations of eHCI domain scores by sex in 7 LMICs

	n	Phys M (SD)	Comm M (SD)	Cult M (SD)	Soc M (SD)	Persev M (SD)	Appr M (SD)	Num M (SD)	Read M (SD)	Writ M (SD)
Brazil										
Female	855	.95 (.13)	.94 (.14)	.90 (.14)	.92 (.15)	.83 (.27)	.93 (.16)	.56 (.29)	.32 (.22)	.55 (.27)
Male	853	.94 (.14)	.89 (.18)	.87 (.19)	.86 (.21)	.69 (.35)	.92 (.16)	.52 (.28)	.28 (.21)	.51 (.28)
China										
Female	5,338	.78 (.26)	.82 (.22)	.74 (.29)	.64 (.24)	.42 (.30)	.81 (.24)	.68 (.28)	.42 (.28)	.32 (.33)
Male	5,587	.77 (.26)	.81 (.23)	.71 (.30)	.62 (.24)	.40 (.30)	.80 (.25)	.66 (.29)	.40 (.28)	.31 (.33)
Kiribati										
Female	3,915	.67 (.27)	.76 (.27)	.72 (.31)	.74 (.24)	.47 (.26)	.76 (.26)	.53 (.27)	.52 (.32)	.52 (.37)
Male	4,269	.64 (.27)	.73 (.27)	.70 (.31)	.72 (.25)	.43 (.26)	.74 (.26)	.50 (.28)	.48 (.33)	.49 (.36)
Lao PDR										
Female	3,669	-	.84 (.25)	.71 (.22)	.50 (.21)	.52 (.26)	.66 (.28)	.37 (.26)	.14 (.22)	.18 (.20)
Male	3,824	-	.84 (.26)	.70 (.21)	.47 (.20)	.50 (.26)	.69 (.27)	.37 (.26)	.12 (.22)	.16 (.18)
Samoa										
Female	5,898	.70 (.27)	.57 (.38)	.43 (.37)	.50 (.28)	.36 (.28)	.61 (.38)	.34 (.33)	.39 (.37)	.37 (.40)
Male	6,293	.69 (.26)	.56 (.38)	.41 (.37)	.49 (.28)	.35 (.28)	.60 (.38)	.33 (.33)	.37 (.37)	.36 (.39)
Tonga										
Female	2,967	.75 (.26)	.86 (.22)	.69 (.27)	.70 (.16)	.47 (.25)	.72 (.33)	.51 (.30)	.49 (.35)	.40 (.35)
Male	3,247	.71 (.27)	.84 (.23)	.67 (.28)	.68 (.16)	.44 (.25)	.71 (.33)	.47 (.30)	.44 (.35)	.35 (.34)
Tuvalu										
Female	277	.76 (.25)	.78 (.23)	.71 (.29)	.75 (.21)	.53 (.30)	.85 (.25)	.58 (.31)	.56 (.34)	.60 (.36)
Male	272	.67 (.29)	.72 (.26)	.62 (.29)	.65 (.25)	.43 (.31)	.81 (.26)	.51 (.32)	.45 (.35)	.50 (.36)

Note: The Lao PDR version of the eHCI does not include a Physical Health domain. Phys=Physical Health, Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing.

Supplementary Table 4.5. Means and standard deviations of eHCI domain scores by stunting status (yes/no) in 7 LMICs

		n	Phys M (SD)	Comm M (SD)	Cult M (SD)	Soc M (SD)	Persev M (SD)	Appr M (SD)	Num M (SD)	Read M (SD)	Writ M (SD)
China											
	Yes	1,602	.78 (.25)	.76 (.23)	.69 (.31)	.61 (.24)	.40 (.28)	.77 (.26)	.62 (.29)	.38 (.28)	.30 (.34)
	No	8,477	.76 (.26)	.82 (.22)	.72 (.29)	.63 (.24)	.40 (.30)	.81 (.24)	.67 (.28)	.41 (.28)	.31 (.33)
Kiribati											
	Yes	1,935	.63 (.28)	.70 (.29)	.67 (.31)	.71 (.25)	.44 (.25)	.72 (.27)	.46 (.28)	.45 (.32)	.43 (.36)
	No	5,003	.67 (.26)	.77 (.26)	.73 (.30)	.75 (.24)	.45 (.26)	.76 (.25)	.55 (.27)	.53 (.32)	.54 (.36)
Lao PDR											
	Yes	3,652	-	.83 (.26)	.70 (.22)	.47 (.20)	.50 (.26)	.67 (.28)	.35 (.26)	.12 (.20)	.16 (.17)
	No	2,989	-	.85 (.25)	.72 (.21)	.49 (.21)	.52 (.26)	.69 (.27)	.40 (.26)	.16 (.23)	.19 (.20)
Samoa											
	Yes	2,332	.63 (.29)	.49 (.38)	.39 (.38)	.45 (.27)	.39 (.29)	.51 (.37)	.28 (.30)	.32 (.35)	.32 (.37)
	No	9,859	.71 (.26)	.58 (.37)	.43 (.37)	.50 (.28)	.35 (.26)	.63 (.39)	.34 (.33)	.40 (.37)	.38 (.40)
Tonga											
	Yes	374	.72 (.28)	.83 (.25)	.65 (.30)	.69 (.18)	.50 (.24)	.67 (.36)	.49 (.35)	.46 (.38)	.35 (.37)
	No	5,749	.73 (.27)	.85 (.22)	.68 (.27)	.69 (.16)	.45 (.25)	.72 (.32)	.49 (.30)	.47 (.35)	.37 (.35)
Tuvalu											
	Yes	132	.67 (.30)	.70 (.27)	.62 (.31)	.69 (.26)	.47 (.33)	.76 (.29)	.50 (.35)	.43 (.36)	.44 (.38)
	No	319	.73 (.27)	.76 (.23)	.67 (.27)	.70 (.23)	.51 (.31)	.85 (.24)	.57 (.30)	.53 (.34)	.58 (.36)

Note: Children's height and weight were not collected in Brazil. The Lao PDR version of the eHCI does not include a Physical Health domain. Phys=Physical Health, Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing.

Supplementary Table 4.6. Means and standard deviations of eHCI domain scores by preschool attendance (yes/no) in 7 LMICs

	n	Phys M (SD)	Comm M (SD)	Cult M (SD)	Soc M (SD)	Persev M (SD)	Appr M (SD)	Num M (SD)	Read M (SD)	Writ M (SD)
China										
Yes	9,159	.79 (.25)	.83 (.21)	.76 (.27)	.65 (.23)	.42 (.30)	.84 (.21)	.75 (.23)	.47 (.27)	.37 (.34)
No	2,176	.68 (.29)	.70 (.28)	.56 (.33)	.52 (.26)	.35 (.30)	.66 (.31)	.31 (.22)	.16 (.18)	.07 (.16)
Kiribati										
Yes	7,665	.66 (.27)	.75 (.26)	.72 (.30)	.75 (.24)	.45 (.26)	.77 (.23)	.53 (.27)	.51 (.32)	.52 (.36)
No	674	.57 (.33)	.64 (.34)	.52 (.36)	.57 (.30)	.38 (.24)	.51 (.37)	.36 (.29)	.34 (.33)	.35 (.35)
Lao PDR										
Yes	1,738	-	.89 (.21)	.79 (.18)	.56 (.20)	.56 (.27)	.73 (.23)	.50 (.26)	.26 (.26)	.27 (.22)
No	5,755	-	.83 (.27)	.68 (.22)	.46 (.20)	.49 (.26)	.66 (.29)	.33 (.25)	.09 (.18)	.14 (.16)
Samoa										
Yes	4,657	.80 (.25)	.70 (.34)	.60 (.35)	.63 (.26)	.36 (.32)	.77 (.31)	.52 (.32)	.58 (.35)	.60 (.38)
No	7,534	.63 (.25)	.48 (.37)	.31 (.33)	.41 (.26)	.35 (.26)	.50 (.38)	.21 (.27)	.26 (.33)	.22 (.33)
Tonga										
Yes	2,701	.78 (.25)	.85 (.22)	.72 (.26)	.72 (.15)	.46 (.26)	.76 (.30)	.60 (.28)	.56 (.34)	.48 (.33)
No	3,513	.69 (.28)	.84 (.23)	.65 (.28)	.68 (.17)	.44 (.24)	.68 (.34)	.41 (.30)	.39 (.34)	.29 (.34)
Tuvalu										
Yes	498	.71 (.28)	.76 (.25)	.66 (.30)	.69 (.25)	.49 (.31)	.83 (.25)	.55 (.32)	.51 (.35)	.56 (.36)
No	51	.81 (.20)	.71 (.23)	.75 (.21)	.76 (.14)	.44 (.28)	.86 (.21)	.50 (.30)	.48 (.29)	.44 (.37)

Note: Results are not presented for Brazil as 100% of the sample were attending preschool. The Lao PDR version of the eHCI does not include a Physical Health domain. Phys=Physical Health, Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing.

Supplementary Table 4.7. Means and standard deviations of eHCI domain scores by maternal education in 7 LMICs

	n	Phys M (SD)	Comm M (SD)	Cult M (SD)	Soc M (SD)	Persev M (SD)	Appr M (SD)	Num M (SD)	Read M (SD)	Writ M (SD)
China										
No school	307	.80 (.23)	.83 (.23)	.79 (.26)	.69 (.24)	.46 (.31)	.83 (.22)	.51 (.29)	.35 (.25)	.23 (.28)
Started Primary	1,242	.72 (.26)	.80 (.22)	.70 (.29)	.61 (.24)	.41 (.31)	.78 (.24)	.59 (.29)	.40 (.28)	.31 (.35)
Finished Primary	3,372	.78 (.25)	.81 (.23)	.77 (.28)	.66 (.23)	.45 (.31)	.82 (.23)	.66 (.29)	.42 (.27)	.30 (.32)
Started Secondary	3,563	.76 (.27)	.80 (.24)	.67 (.31)	.60 (.24)	.38 (.31)	.77 (.27)	.66 (.29)	.38 (.27)	.30 (.32)
Finished Secondary	1,358	.78 (.27)	.82 (.22)	.71 (.30)	.61 (.24)	.36 (.28)	.82 (.23)	.73 (.36)	.43 (.29)	.33 (.35)
Tertiary	751	.80 (.26)	.85 (.19)	.75 (.29)	.62 (.24)	.38 (.28)	.83 (.20)	.73 (.24)	.43 (.28)	.30 (.32)
Kiribati										
Started Primary	714	.62 (.27)	.70 (.29)	.68 (.32)	.70 (.27)	.44 (.24)	.72 (.28)	.48 (.29)	.47 (.34)	.47 (.37)
Finished Primary	1,438	.64 (.27)	.72 (.28)	.68 (.33)	.72 (.26)	.45 (.24)	.72 (.27)	.49 (.28)	.46 (.33)	.48 (.37)
Started Secondary	2,319	.66 (.28)	.76 (.27)	.71 (.30)	.73 (.24)	.45 (.26)	.74 (.26)	.52 (.27)	.51 (.32)	.51 (.36)
Finished Secondary	2,710	.66 (.27)	.75 (.27)	.72 (.30)	.74 (.24)	.44 (.26)	.76 (.25)	.53 (.28)	.51 (.32)	.51 (.37)
Tertiary	785	.69 (.25)	.77 (.25)	.74 (.28)	.76 (.23)	.47 (.26)	.79 (.22)	.54 (.26)	.52 (.33)	.51 (.36)
Lao PDR										
No school	2,265	-	.83 (.27)	.67 (.22)	.45 (.20)	.49 (.26)	.63 (.30)	.32 (.24)	.08 (.17)	.13 (.15)
Started Primary	2,045	-	.83 (.26)	.69 (.22)	.48 (.20)	.51 (.26)	.65 (.28)	.36 (.25)	.12 (.20)	.15 (.18)
Finished Primary	2,182	-	.85 (.25)	.73 (.21)	.50 (.21)	.52 (.26)	.70 (.26)	.40 (.27)	.16 (.23)	.19 (.20)
Finished Secondary	754	-	.86 (.24)	.75 (.20)	.52 (.20)	.52 (.26)	.75 (.21)	.42 (.28)	.21 (.25)	.23 (.21)
Tertiary	242	-	.89 (.21)	.81 (.18)	.52 (.21)	.55 (.27)	.79 (.28)	.46 (.30)	.26 (.28)	.31 (.27)
Samoa										
Finished Primary	222	.57 (.33)	.55 (.36)	.38 (.35)	.46 (.26)	.33 (.25)	.57 (.38)	.37 (.34)	.36 (.40)	.41 (.39)
Started Secondary	603	.63 (.26)	.56 (.36)	.33 (.30)	.43 (.26)	.25 (.29)	.54 (.34)	.22 (.25)	.27 (.31)	.26 (.34)
Finished Secondary	10,037	.69 (.26)	.56 (.38)	.41 (.37)	.49 (.28)	.36 (.28)	.60 (.38)	.32 (.32)	.37 (.37)	.35 (.39)
Tertiary	1,329	.75 (.25)	.59 (.38)	.54 (.36)	.58 (.27)	.37 (.30)	.67 (.37)	.47 (.33)	.52 (.37)	.49 (.40)
Tonga										
Finished Primary	81	.60 (.29)	.81 (.28)	.62 (.30)	.61 (.20)	.42 (.25)	.62 (.38)	.41 (.33)	.37 (.34)	.30 (.36)
Started Secondary	2,399	.69 (.28)	.83 (.23)	.64 (.28)	.67 (.17)	.43 (.25)	.67 (.34)	.44 (.30)	.41 (.34)	.33 (.34)
Finished Secondary	2,621	.74 (.26)	.85 (.22)	.70 (.27)	.70 (.16)	.46 (.24)	.73 (.32)	.52 (.30)	.50 (.35)	.39 (.35)
Tertiary	1,107	.79 (.24)	.87 (.22)	.74 (.27)	.72 (.15)	.48 (.26)	.78 (.30)	.56 (.31)	.52 (.35)	.43 (.35)
Tuvalu										
Finished Primary	47	.76 (.26)	.78 (.26)	.71 (.30)	.72 (.22)	.47 (.27)	.89 (.23)	.63 (.30)	.55 (.34)	.63 (.34)

Started Secondary	89	.70 (.26)	.73 (.25)	.64 (.31)	.68 (.25)	.41 (.28)	.79 (.29)	.54 (.32)	.50 (.35)	.53 (.36)
Finished Secondary	107	.73 (.28)	.77 (.23)	.67 (.29)	.69 (.23)	.50 (.32)	.82 (.26)	.55 (.29)	.51 (.34)	.59 (.37)
Tertiary	117	.73 (.27)	.77 (.23)	.65 (.31)	.67 (.26)	.50 (.33)	.85 (.24)	.58 (.33)	.54 (.35)	.59 (.37)

Note: Children's maternal education data were not collected in Brazil. The Lao PDR version of the eHCI does not include a Physical Health domain. Phys=Physical Health, Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing.

Supplementary Table 4.8. Means and standard deviations of eHCI domain scores by engagement in home learning activities (yes/no) in 7 LMICs

		n	Phys M (SD)	Comm M (SD)	Cult M (SD)	Soc M (SD)	Persev M (SD)	Appr M (SD)	Num M (SD)	Read M (SD)	Writ M (SD)
Kiribati											
Read to	Yes	3,930	.69 (.26)	.78 (.26)	.76 (.29)	.76 (.23)	.48 (.26)	.78 (.23)	.56 (.27)	.55 (.32)	.55 (.36)
	No	4,403	.63 (.28)	.71 (.28)	.66 (.32)	.70 (.26)	.42 (.25)	.71 (.27)	.47 (.28)	.45 (.32)	.46 (.37)
Tell story	Yes	5,886	.66 (.27)	.77 (.26)	.72 (.30)	.75 (.24)	.45 (.26)	.77 (.24)	.54 (.27)	.51 (.32)	.52 (.36)
	No	2,447	.63 (.28)	.69 (.29)	.67 (.32)	.70 (.26)	.45 (.25)	.70 (.30)	.47 (.29)	.46 (.33)	.46 (.37)
Sing/dance	Yes	4,964	.67 (.27)	.76 (.26)	.73 (.29)	.75 (.24)	.45 (.26)	.77 (.24)	.53 (.27)	.50 (.32)	.50 (.37)
	No	3,369	.64 (.28)	.71 (.29)	.67 (.32)	.71 (.26)	.44 (.25)	.71 (.28)	.50 (.28)	.49 (.33)	.50 (.37)
Play	Yes	4,774	.66 (.27)	.77 (.26)	.72 (.30)	.74 (.24)	.45 (.26)	.77 (.24)	.53 (.27)	.50 (.32)	.50 (.37)
	No	3,559	.64 (.28)	.71 (.28)	.68 (.32)	.71 (.26)	.44 (.25)	.72 (.28)	.50 (.28)	.49 (.33)	.50 (.37)
Go outside	Yes	3,455	.68 (.27)	.78 (.26)	.75 (.29)	.76 (.24)	.47 (.26)	.78 (.23)	.55 (.27)	.53 (.32)	.53 (.36)
	No	4,878	.64 (.28)	.72 (.28)	.68 (.32)	.71 (.25)	.43 (.26)	.72 (.27)	.49 (.28)	.48 (.33)	.48 (.37)
Name, count, draw	Yes	3,380	.69 (.25)	.79 (.24)	.76 (.28)	.77 (.23)	.47 (.26)	.79 (.22)	.57 (.26)	.54 (.31)	.54 (.36)
	No	4,953	.63 (.28)	.71 (.28)	.67 (.32)	.70 (.26)	.43 (.25)	.71 (.28)	.48 (.28)	.47 (.33)	.47 (.37)
Books in home	Yes	3,948	.71 (.24)	.80 (.24)	.80 (.26)	.80 (.21)	.48 (.25)	.82 (.20)	.60 (.25)	.61 (.31)	.60 (.35)
	No	4,124	.60 (.29)	.69 (.29)	.62 (.33)	.67 (.27)	.41 (.26)	.68 (.29)	.43 (.27)	.39 (.30)	.40 (.35)
Lao PDR											
Read to	Yes	2,149	-	.87 (.23)	.77 (.17)	.56 (.19)	.54 (.25)	.75 (.22)	.50 (.26)	.25 (.26)	.25 (.22)
	No	5,344	-	.83 (.27)	.68 (.22)	.45 (.20)	.50 (.27)	.64 (.29)	.32 (.24)	.09 (.17)	.14 (.16)
Tell story	Yes	1,903	-	.88 (.22)	.77 (.18)	.55 (.19)	.55 (.25)	.75 (.22)	.49 (.27)	.22 (.26)	.24 (.23)
	No	5,590	-	.83 (.27)	.69 (.22)	.46 (.20)	.50 (.26)	.65 (.29)	.33 (.25)	.10 (.19)	.14 (.17)
Sing/dance	Yes	2,215	-	.85 (.24)	.76 (.19)	.55 (.19)	.54 (.26)	.75 (.22)	.47 (.27)	.22 (.26)	.23 (.22)
	No	5,278	-	.84 (.26)	.69 (.22)	.45 (.20)	.50 (.26)	.64 (.29)	.33 (.25)	.09 (.18)	.14 (.16)
Play	Yes	1,341	-	.81 (.26)	.74 (.21)	.56 (.20)	.56 (.25)	.74 (.22)	.45 (.26)	.21 (.26)	.24 (.21)
	No	6,152	-	.85 (.26)	.70 (.21)	.47 (.20)	.50 (.26)	.66 (.28)	.35 (.26)	.12 (.20)	.15 (.18)
Go outside	Yes	1,775	-	.83 (.25)	.71 (.20)	.52 (.20)	.50 (.28)	.73 (.24)	.42 (.25)	.16 (.23)	.21 (.19)
	No	5,718	-	.85 (.26)	.71 (.22)	.47 (.20)	.52 (.26)	.66 (.28)	.36 (.26)	.12 (.21)	.15 (.18)
Name, count, draw	Yes	2,280	-	.87 (.22)	.77 (.18)	.57 (.19)	.54 (.26)	.76 (.21)	.49 (.26)	.24 (.26)	.26 (.23)
	No	5,213	-	.83 (.27)	.68 (.22)	.45 (.20)	.50 (.26)	.63 (.29)	.32 (.24)	.08 (.17)	.13 (.15)
Samoa											
Read to	Yes	2,424	.76 (.25)	.68 (.34)	.53 (.37)	.59 (.27)	.39 (.29)	.77 (.31)	.39 (.33)	.47 (.37)	.48 (.43)
	No	6,297	.60 (.25)	.44 (.37)	.28 (.32)	.38 (.24)	.34 (.25)	.45 (.37)	.20 (.25)	.23 (.31)	.20 (.30)

Tell story	Yes	2,856	.75 (.25)	.67 (.34)	.51 (.36)	.57 (.27)	.38 (.29)	.74 (.32)	.37 (.32)	.43 (.37)	.43 (.43)
	No	5,865	.59 (.25)	.43 (.37)	.27 (.32)	.37 (.24)	.34 (.25)	.43 (.37)	.19 (.25)	.23 (.31)	.20 (.30)
Sing/dance	Yes	5,043	.68 (.25)	.57 (.36)	.38 (.35)	.48 (.27)	.32 (.28)	.63 (.36)	.28 (.29)	.34 (.35)	.31 (.39)
	No	3,678	.59 (.26)	.42 (.38)	.31 (.35)	.38 (.25)	.40 (.24)	.40 (.38)	.21 (.28)	.24 (.33)	.22 (.33)
Play	Yes	4,800	.69 (.25)	.59 (.36)	.40 (.35)	.49 (.27)	.32 (.29)	.65 (.35)	.29 (.30)	.36 (.35)	.33 (.39)
	No	3,921	.59 (.26)	.40 (.38)	.29 (.34)	.37 (.25)	.40 (.23)	.39 (.37)	.20 (.27)	.22 (.32)	.21 (.32)
Go outside	Yes	4,722	.69 (.25)	.56 (.37)	.37 (.36)	.47 (.27)	.32 (.27)	.63 (.36)	.27 (.30)	.34 (.35)	.31 (.39)
	No	3,999	.60 (.27)	.44 (.38)	.33 (.34)	.40 (.26)	.39 (.25)	.43 (.38)	.22 (.28)	.25 (.33)	.23 (.33)
Name, count, draw	Yes	2,720	.75 (.25)	.68 (.34)	.50 (.38)	.57 (.28)	.39 (.28)	.75 (.32)	.38 (.33)	.46 (.38)	.48 (.42)
	No	6,001	.60 (.25)	.43 (.37)	.28 (.32)	.38 (.24)	.33 (.26)	.44 (.37)	.19 (.25)	.22 (.30)	.19 (.30)
Books in home	Yes	6,253	.77 (.25)	.70 (.33)	.59 (.35)	.63 (.25)	.38 (.30)	.74 (.33)	.49 (.34)	.60 (.34)	.57 (.39)
	No	5,938	.61 (.26)	.42 (.37)	.25 (.31)	.35 (.24)	.34 (.26)	.46 (.37)	.17 (.21)	.15 (.23)	.15 (.26)
Tonga											
Read to	Yes	3,566	.77 (.25)	.89 (.19)	.73 (.25)	.72 (.15)	.47 (.25)	.76 (.30)	.55 (.29)	.53 (.34)	.44 (.35)
	No	1,944	.63 (.28)	.79 (.26)	.58 (.29)	.64 (.18)	.40 (.24)	.60 (.36)	.33 (.26)	.30 (.30)	.20 (.28)
Tell story	Yes	3,836	.76 (.26)	.88 (.19)	.73 (.25)	.71 (.15)	.47 (.24)	.76 (.30)	.53 (.30)	.51 (.35)	.42 (.35)
	No	1,676	.63 (.28)	.78 (.27)	.57 (.29)	.64 (.18)	.41 (.25)	.58 (.36)	.33 (.27)	.30 (.30)	.22 (.29)
Sing/dance	Yes	5,025	.74 (.26)	.87 (.20)	.70 (.26)	.70 (.16)	.45 (.25)	.73 (.32)	.49 (.30)	.47 (.34)	.37 (.35)
	No	489	.57 (.28)	.69 (.32)	.46 (.30)	.57 (.21)	.41 (.25)	.50 (.37)	.28 (.29)	.24 (.31)	.20 (.30)
Play	Yes	5,180	.73 (.27)	.86 (.21)	.69 (.26)	.70 (.16)	.45 (.25)	.72 (.32)	.48 (.30)	.46 (.35)	.37 (.35)
	No	335	.56 (.27)	.70 (.33)	.49 (.32)	.57 (.21)	.42 (.23)	.49 (.38)	.29 (.29)	.24 (.30)	.20 (.30)
Go outside	Yes	5,174	.73 (.27)	.86 (.21)	.69 (.26)	.70 (.16)	.45 (.25)	.72 (.32)	.48 (.30)	.46 (.35)	.37 (.35)
	No	340	.58 (.28)	.68 (.34)	.47 (.32)	.57 (.21)	.43 (.24)	.45 (.38)	.29 (.28)	.24 (.30)	.18 (.29)
Name, count, draw	Yes	3,152	.79 (.24)	.91 (.16)	.77 (.23)	.73 (.14)	.48 (.24)	.79 (.29)	.59 (.28)	.58 (.34)	.48 (.35)
	No	2,361	.63 (.28)	.77 (.26)	.56 (.28)	.63 (.18)	.41 (.25)	.60 (.35)	.31 (.25)	.28 (.28)	.19 (.27)
Books in home	Yes	4,129	.78 (.24)	.89 (.19)	.74 (.24)	.72 (.14)	.48 (.25)	.78 (.28)	.58 (.29)	.58 (.33)	.46 (.35)
	No	2,074	.62 (.28)	.77 (.27)	.57 (.29)	.63 (.18)	.39 (.25)	.58 (.37)	.31 (.26)	.24 (.26)	.20 (.28)
Tuvalu											
Read to	Yes	297	.74 (.26)	.78 (.24)	.68 (.30)	.70 (.24)	.49 (.31)	.85 (.25)	.59 (.32)	.56).36)	.61 (.36)
	No	88	.67 (.31)	.74 (.26)	.60 (.32)	.64 (.28)	.49 (.31)	.79 (.28)	.50 (.32)	.45 (.35)	.53 (.36)
Tell story	Yes	337	.73 (.27)	.78 (.24)	.68 (.30)	.70 (.24)	.49 (.31)	.86 (.25)	.59 (.32)	.55 (.35)	.61 (.36)
	No	47	.63 (.27)	.67 (.30)	.51 (.32)	.60 (.28)	.52 (.32)	.72 (.30)	.45 (.32)	.39 (.35)	.46 (.38)
Sing/dance	Yes	353	.72 (.27)	.77 (.24)	.67 (.31)	.69 (.25)	.48 (.31)	.85 (.25)	.58 (.32)	.54 (.36)	.60 (.36)
	No	31	.75 (.24)	.71 (.32)	.57 (.31)	.64 (.28)	.59 (.25)	.72 (.31)	.51 (.33)	.43 (.35)	.52 (.39)

Play	Yes	359	.72 (.27)	.77 (.25)	.67 (.31)	.69 (.25)	.48 (.31)	.84 (.26)	.58 (.32)	.54 (.36)	.60 (.36)
	No	25	.71 (.28)	.70 (.29)	.56 (.32)	.64 (.29)	.65 (.25)	.84 (.20)	.50 (.32)	.46 (.35)	.50 (.40)
Go outside	Yes	338	.72 (.27)	.77 (.25)	.68 (.31)	.69 (.25)	.49 (.32)	.85 (.26)	.58 (.33)	.54 (.36)	.61 (.36)
	No	46	.72 (.25)	.74 (.24)	.55 (.31)	.65 (.25)	.49 (.28)	.77 (.23)	.50 (.30)	.45 (.34)	.47 (.37)
Name, count, draw	Yes	309	.72 (.27)	.78 (.25)	.68 (.31)	.71 (.24)	.49 (.31)	.85 (.26)	.59 (.32)	.55 (.36)	.62 (.36)
	No	74	.72 (.27)	.73 (.26)	.59 (.32)	.60 (.26)	.48 (.33)	.81 (.25)	.51 (.33)	.46 (.36)	.47 (.36)

Notes: Home learning environment data were not collected in Brazil and China. The Lao PDR version of the eHCI does not include a Physical Health domain. Phys=Physical Health, Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing.

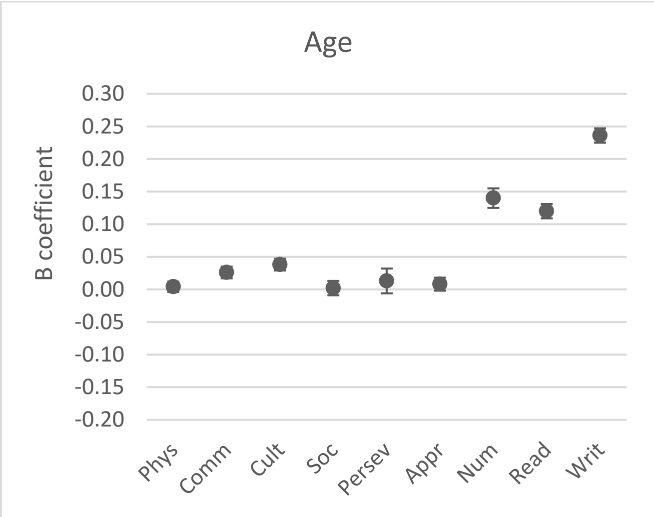
Supplementary Table 4.9. Discriminant validity: linear regression of eHCI domain scores on demographic and contextual variables in 7 LMICs

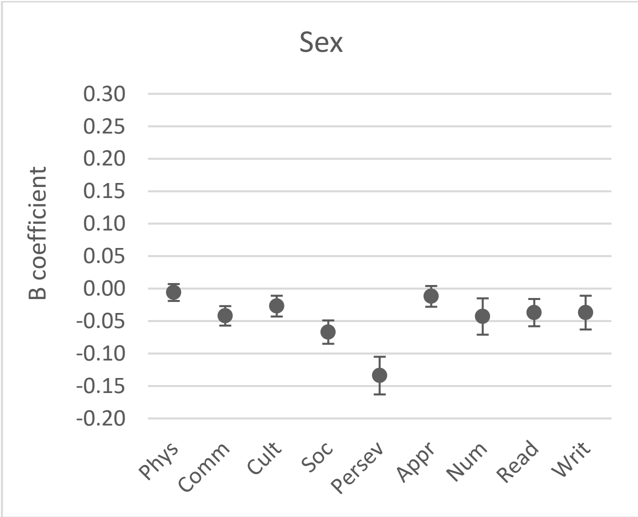
	Phys Coefficient (95% CI)	Comm Coefficient (95% CI)	Cult Coefficient (95% CI)	Soc Coefficient (95% CI)	Persev Coefficient (95% CI)	Appr Coefficient (95% CI)	Num Coefficient (95% CI)	Read Coefficient (95% CI)	Writ Coefficient (95% CI)
Brazil									
Sex	-.01 (-.02 to .01)	-.04 (-.06 to -.03)	-.03 (-.04 to -.01)	-.07 (-.08 to -.05)	-.13 (-.16 to -.11)	-.01 (-.03 to .00)	-.04 (-.07 to -.02)	-.04 (.06 to -.02)	-.04 (.06 to -.01)
Age (years)	.00 (.00 to .01)	.03 (.02 to .04)	.04 (.03 to .05)	.00 (-.01 to .01)	.01 (-.01 to .03)	.01 (.01 to .02)	.14 (.12 to .16)	.12 (.11 to .13)	.24 (.22 to .25)
China									
Sex	.01 (.00 to .02)	.01 (.00 to .02)	.02 (.01 to .03)	.02 (.01 to .03)	.02 (.01 to .03)	.01 (.00 to .02)	02 (.01 to .03)	02 (.01 to .03)	.01 (.00 to .02)
Age (years)	.05 (.04 to .05)	.05 (.04 to .05)	.07 (.07 to .08)	.05 (.05 to .06)	.02 (.01 to .03)	.06 (.05 to .06)	.18 (.18 to .19)	.14 (.13 to .14)	.18 (.17 to .18)
Stunted	.02 (.00 to .03)	-.05 (-.06 to -.04)	-.04 (-.05 to -.02)	-.01 (-.03 to .00)	.00 (-.02 to .02)	-.04 (-.05 to -.03)	-.06 (-.07 to -.04)	-.03 (-.05 to -.02)	-.02 (-.03 to .00)
Preschool	.11 (.10 to .12)	.13 (.12 to .14)	.19 (.18 to .21)	.14 (.13 to .15)	.07 (.05 to .08)	.17 (.16 to .19)	.44 (.43 to .45)	.31 (.30 to .32)	.30 (.28 to .31)
Maternal education	.01 (.01 to .01)	.01 (.00 to .01)	-.01 (-.02 to -.01)	-.01 (-.02 to -.01)	-.02 (-.03 to -.02)	.00 (.00 to .01)	.04 (.03 to .04)	.01 (.00 to .01)	.01 (.00 to .01)
Kiribati									
Sex	.03 (.02 to .04)	.04 (.02 to .05)	.02 (.01 to .04)	.02 (.01 to .04)	.04 (.03 to .05)	.02 (.01 to .03)	.03 (.02 to .04)	.04 (.02 to .05)	.03 (.02 to .05)
Age (years)	.06 (.06 to .07)	.07 (.07 to .08)	.10 (.09 to .10)	.06 (.06 to .07)	.03 (.02 to .04)	.05 (.04 to .05)	.12 (.11 to .12)	.11 (.10 to .11)	.16 (.16 to .17)
Stunted	-.04 (-.06 to -.03)	-.07 (-.08 to -.05)	-.06 (-.07 to -.04)	-.04 (-.05 to -.02)	-.01 (-.02 to .01)	-.03 (-.05 to -.02)	-.09 (-.10 to -.07)	-.08 (-.09 to -.06)	-.11 (-.13 to -.09)
Preschool	.09 (.07 to .11)	.11 (.09 to .13)	.20 (.18 to .22)	.18 (.16 to .20)	.08 (.06 to .10)	.26 (.24 to .28)	.17 (.14 to .19)	.17 (.15 to .20)	.17 (.14 to .20)
Maternal education	.02 (.01 to .02)	.02 (.01 to .02)	.02 (.01 to .02)	.01 (.01 to .02)	.00 (.00 to .01)	.02 (.01 to .02)	.02 (.01 to .02)	.02 (.01 to .02)	.01 (.01 to .02)
Read to	.06 (.05 to .07)	.06 (.05 to .08)	.09 (.08 to .11)	.06 (.05 to .07)	.05 (.04 to .06)	.07 (.06 to .08)	.09 (.08 to .10)	.10 (.09 to .12)	.09 (.07 to .11)
Tell story	.03 (.02 to .05)	.08 (.07 to .09)	.06 (.04 to .07)	.05 (.03 to .06)	.00 (-.01 to .01)	.07 (.06 to .08)	.07 (.06 to .09)	.06 (.04 to .07)	.06 (.04 to .08)
Sing/dance	.03 (.02 to .04)	.05 (.04 to .06)	.06 (.05 to .07)	.04 (.03 to .06)	.01 (.00 to .02)	.06 (.04 to .07)	.03 (.02 to .04)	.01 (.00 to .03)	.00 (-.02 to .01)
Play	.02 (.01 to .03)	.05 (.04 to .06)	.04 (.03 to .05)	.03 (.02 to .04)	.01 (.00 to .02)	.05 (.04 to .06)	.02 (.01 to .04)	.01 (-.01 to .02)	.00 (-.02 to .02)
Go outside	.05 (.04 to .06)	.06 (.05 to .08)	.07 (.06 to .08)	.05 (.04 to .06)	.03 (.02 to .04)	.07 (.05 to .08)	.06 (.05 to .07)	.05 (.03 to .06)	.05 (.03 to .06)
Name, count, draw	.06 (.05 to .07)	.08 (.07 to .09)	.10 (.08 to .11)	.07 (.06 to .08)	.04 (.03 to .05)	.08 (.07 to .09)	.09 (.08 to .11)	.08 (.06 to .09)	.07 (.06 to .09)
Books in home	.12 (.11 to .13)	.12 (.11 to .13)	.18 (.17 to .19)	.13 (.12 to .14)	.07 (.05 to .08)	.14 (.13 to .15)	.18 (.16 to .19)	.23 (.21 to .24)	.21 (.19 to .22)
Lao PDR									
Sex	-	.00 (-.01 to .01)	.01 (.00 to .02)	.03 (.02 to .04)	.01 (.00 to .03)	-.03 (-.04 to -.01)	.01 (.00 to .02)	.02 (.01 to .03)	.02 (.01 to .03)
Age (years)	-	.07 (.06 to .07)	.07 (.06 to .07)	.07 (.06 to .07)	.05 (.05 to .06)	.07 (.06 to .07)	.12 (.11 to .12)	.07 (.07 to .07)	.06 (.06 to .07)
Stunted	-	-.02 (-.03 to -.01)	-.03 (-.04 to -.02)	-.02 (-.03 to -.01)	-.02 (-.03 to .00)	-.03 (-.04 to -.01)	-.04 (-.05 to -.03)	-.04 (-.05 to -.03)	-.03 (-.04 to -.03)
Preschool	-	.06 (.05 to .08)	.11 (.10 to .12)	.10 (.09 to .11)	.07 (.05 to .08)	.07 (.06 to .09)	.17 (.16 to .19)	.17 (.16 to .18)	.14 (.13 to .15)
Maternal education	-	.01 (.01 to .02)	.03 (.03 to .03)	.02 (.02 to .03)	.01 (.01 to .02)	.04 (.03 to .05)	.03 (.03 to .04)	.04 (.04 to .05)	.04 (.03 to .04)
Read to	-	.04 (.03 to .05)	.09 (.08 to .10)	.11 (.10 to .12)	.04 (.03 to .05)	.11 (.10 to .12)	.19 (.18 to .20)	.16 (.15 to .17)	.11 (.11 to .12)
Tell story	-	.05 (.03 to .06)	.08 (.07 to .09)	.09 (.08 to .10)	.05 (.04 to .06)	.10 (.09 to .12)	.16 (.15 to .17)	.12 (.11 to .13)	.09 (.08 to .10)
Sing/dance	-	.01 (.00 to .03)	.08 (.07 to .09)	.10 (.09 to .011)	.04 (.02 to .05)	.11 (.09 to .12)	.14 (.13 to .15)	.13 (.12 to .14)	.10 (.09 to .10)
Play	-	-.05 (-.06 to -.03)	.04 (.03 to .05)	.09 (.08 to .10)	.06 (.05 to .08)	.09 (.07 to .10)	.09 (.08 to .11)	.09 (.08 to .11)	.09 (.07 to .10)
Go outside	-	-.01 (-.03 to .00)	.00 (-.01 to .02)	.05 (.04 to .06)	-.02 (-.03 to -.01)	.08 (.06 to .09)	.06 (.05 to .08)	.04 (.03 to .05)	.06 (.05 to .07)
Name, count, draw	-	.04 (.03 to .05)	.09 (.08 to .10)	.12 (.11 to .13)	.05 (.04 to .06)	.13 (.12 to .014	.18 (.16 to .19)	.16 (.15 to .17)	.13 (.12 to .14)
Samoa									
Sex	.01 (.00 to .02)	.01 (.00 to .03)	.01 (.00 to .03)	.02 (.01 to .03)	.01 (.00 to .02)	.02 (.00 to .03)	.01 (.00 to .02)	.02 (.00 to .03)	.01 (.00 to .03)
Age (years)	.04 (.04 to .05)	.07 (.07 to .08)	.07 (.06 to .07)	.05 (.05 to .06)	.04 (.03 to .04)	.05 (.04 to .06)	.07 (.07 to .08)	.08 (.07 to .09)	.09 (.08 to .09)
Stunted	-.08 (-.09 to -.07)	-.09 (-.11 to -.08)	-.03 (-.05 to -.02)	-.05 (-.07 to -.04)	.04 (.03 to .05)	-.11 (- 13 to -.09)	-.06 (-.07 to -.05)	-.08 (-.10 to -.06)	-.06 (-.08 to -.04)
Preschool	.17 (.16 to .18)	.22 (.21 to .23)	.29 (.27 to .30)	.23 (.22 to .23)	.01 (.00 to .02)	.27 (.26 to .28)	.31 (.30 to .32)	.33 (.31 to .34)	.38 (.37 to .39)
Maternal education	.06 (.05 to .07)	.02 (.00 to .03)	.08 (.07 to .09)	.06 (.05 to .07)	.03 (.02 to .04)	.05 (.03 to .06)	.09 (.07 to .10)	.09 (.08 to .11)	.08 (.06 to .09)
Read to	.16 (.15 to .17)	.25 (.23 to .26)	.25 (.24 to .27)	.22 (.20 to .23)	.05 (.03 to .06)	.32 (.31 to .34)	.20 (.18 to .21)	.24 (.22 to .25)	.28 (.27 to .30)
Tell story	.16 (.15 to .17)	.25 (.23 to .26)	.23 (.22 to .25)	.21 (.19 to .22)	.04 (.03 to .05)	.31 (.30 to .33)	.18 (.17 to .19)	.20 (.19 to .22)	.23 (.21 to .24)

Sing/dance	.10 (.09 to .11)	.15 (.14 to .17)	.07 (.06 to .08)	.10 (.08 to .11)	-.08 (-.09 to -.07)	.23 (.22 to .25)	.06 (.05 to .08)	.10 (.09 to .12)	.09 (.08 to .11)
Play	.11 (.10 to .12)	.19 (.18 to .21)	.11 (.10 to .13)	.12 (.11 to .13)	-.08 (-.10 to -.07)	.26 (.24 to .28)	.09 (.08 to .10)	.14 (.12 to .15)	.12 (.11 to .14)
Go outside	.09 (.08 to .10)	.13 (.11 to .14)	.04 (.03 to .06)	.07 (.05 to .08)	-.06 (-.07 to -.05)	.20 (.18 to .21)	.05 (.04 to .06)	.09 (.08 to .10)	.08 (.06 to .10)
Name, count, draw	.16 (.15 to .17)	.25 (.23 to .26)	.22 (.21 to .24)	.19 (.18 to .20)	.06 (.05 to .07)	.31 (.29 to .32)	.19 (.18 to .21)	.24 (.22 to .25)	.29 (.27 to .30)
Books in home	.17 (.16 to .18)	.27 (.26 to .28)	.34 (.33 to .35)	.28 (.27 to .29)	.04 (.03 to .05)	.28 (.27 to .30)	.32 (.31 to .33)	.45 (.44 to .46)	.42 (.41 to .43)
Tonga									
Sex	.05 (.03 to .06)	.02 (.01 to .03)	.03 (.01 to .04)	.02 (.01 to .03)	.03 (.02 to .04)	.01 (-.01 to .03)	.04 (.03 to .06)	.04 (.03 to .06)	.05 (.04 to .07)
Age (years)	.07 (.07 to .08)	.05 (.04 to .05)	.07 (.07 to .08)	.04 (.03 to .04)	.04 (.04 to .05)	.07 (.06 to .08)	.19 (.18 to .20)	.19 (.18 to .20)	.23 (.22 to .24)
Stunted	-.01 (-.04 to .01)	-.02 (-.04 to .01)	-.03 (-.06 to -.01)	-.01 (-.02 to -.01)	.06 (.03 to .08)	-.05 (-.08 to -.02)	.00 (-.04 to .03)	.00 (-.04 to .04)	-.03 (-.06 to .01)
Preschool	.09 (.08 to .11)	.02 (.01 to .03)	.06 (.05 to .08)	.04 (.03 to .05)	.02 (.01 to .03)	.08 (.07 to .10)	.19 (.18 to .20)	.17 (.16 to .19)	.19 (.17 to .20)
Maternal education	.05 (.04 to .06)	.02 (.01 to .03)	.05 (.04 to .06)	.03 (.02 to .03)	.03 (.02 to .03)	.05 (.04 to .06)	.06 (.05 to .07)	.06 (.05 to .07)	.06 (.04 to .07)
Read to	.14 (.13 to .16)	.10 (.09 to .11)	.15 (.14 to .17)	.08 (.07 to .09)	.07 (.06 to .08)	.16 (.14 to .18)	.22 (.21 to .24)	.23 (.21 to .25)	.23 (.22 to .25)
Tell story	.13 (.12 to .15)	.10 (.09 to .12)	.16 (.14 to .17)	.08 (.07 to .08)	.06 (.05 to .07)	.18 (.16 to .20)	.20 (.18 to .21)	.21 (.19 to .23)	.20 (.18 to .22)
Sing/dance	.17 (.14 to .19)	.18 (.16 to .20)	.24 (.21 to .26)	.12 (.10 to .13)	.04 (.02 to .07)	.23 (.20 to .26)	21 (.18 to .23)	.23 (.19 to .26)	.17 (.14 to .20)
Play	.17 (.14 to .20)	.16 (.14 to .19)	.20 (.17 to .23)	.13 (.11 to .15)	.03 (.00 to .05)	.24 (.20 to .27)	.19 (.16 to .23)	.22 (.19 to .26)	.17 (.13 to .21)
Go outside	.15 (.12 to .18)	.18 (.16 to .21)	.22 (.19 to .25)	.13 (.11 to .14)	.01 (-.01 to .04)	.27 (.23 to .31)	.19 (.16 to .23)	.22 (.18 to .26)	.19 (.15 to .22)
Name, count, draw	.17 (.15 to .18)	.15 (.13 to .16)	.21 (.19 to .22)	.10 (.10 to .11)	.07 (.06 to .09)	.19 (.18 to .21)	.27 (.26 to .29)	.30 (.28 to .32)	.29 (.27 to .31)
Books in home	.16 (.14 to .17)	.12 (.11 to .13)	.17 (.16 to .19)	.10 (.09 to .11)	.09 (.07 to .10)	.20 (.18 to .21)	.27 (.25 to .28)	.33 (.32 to .35)	.26 (.24 to .28)
Tuvalu									
Sex	.08 (.04 to .13)	.06 (.02 to .11)	.08 (.04 to .13)	.10 (.06 to .14)	.09 (.04 to .15)	.04 (.00 to .08)	.08 (.03 to .13)	.11 (.05 to .17)	.11 (.05 to .17)
Age (years)	.12 (.09 to .14)	.12 (.10 to .14)	.15 (.13 to .18)	.10 (.08 to .12)	.08 (.05 to .11)	.05 (.02 to .07)	.22 (.20 to .25)	.22 (.19 to .25)	.26 (.23 to .29)
Stunted	-.06 (-.11 to .00)	-.06 (-.11 to -.01)	-.06 (-.11 to .00)	-.02 (-.07 to .03)	-.04 (-.10 to .03)	-.09 (- 14 to -.04)	-.06 (-.13 to .00)	-.10 (-.17 to -.03)	-.14 (-.22 to -.07)
Preschool	-.10 (-.18 to -.02)	.04 (-.03 to .12)	-.09 (-.18 to -.01)	-.07 (-.14 to .00)	.04 (-.05 to .13)	-.03 (-.11 to .04)	.04 (-.05 to .14)	.03 (-.07 to 13)	.12 (.02 to .23)
Maternal education	.00 (-.03 to .03)	.01 (-.02 to .03)	-.01 (-.04 to .02)	-.01 (-.04 to .01)	.02 (-.04 to .01)	.00 (-.03 to .03)	.00 (-.04 to .03)	.00 (-.03 to .04)	.00 (-.04 to .04)
Read to	.07 (.00 to .13)	.04 (-.02 to .10)	.08 (.00 to .15)	.06 (.01 to .12)	.00 (-.08 to .07)	.06 (.00 to .13)	.09 (.01 to .17)	.11 (.03 to .20)	.08 (-.01 to .16)
Tell story	.10 (.02 to .19)	.11 (.03 to .18)	.17 (.08 to .27)	.10 (.02 to .17)	-.03 (-.13 to .06)	.14 (.06 to .22)	.14 (.04 to .24)	.16 (.05 to .27)	.15 (.04 to .26)
Sing/dance	-.04 (-.14 to .06)	.06 (-.03 to .15)	.11 (-.01 to .22)	.05 (-.04 to .14)	-.11 (-.22 to .01)	.13 (.04 to .23)	.07 (-.05 to .19)	.11 (-.02 to 24)	.07 (-.06 to .21)
Play	.01 (-.10 to .12)	.07 (-.03 to .17)	.12 (-.01 to .24)	.05 (-.05 to .15)	-.17 (-.30 to -.05)	.00 (- 11 to .10)	.08 (-.05 to .21)	.08 (-.07 to 22)	.09 (-.06 to .24)
Go outside	.01 (-.08 to .09)	.04 (-.04 to .11)	.13 (.03 to .22)	.04 (-.03 to .22)	-.01 (-.10 to .09)	.08 (.00 to .15)	.08 (-.02 to .18)	.09 (-.02 to 20)	.13 (.02 to .24)
Name, count, draw	.00 (-.07 to .07)	.05 (-.01 to .11)	.09 (.01 to .16)	.11 (.04 to .17)	.02 (-.06 to .09)	.04 (-.03 to .10)	.08 (.00 to .16)	.09 (.00 to .18)	.15 (.06 to .24)

Note. Sex: male = 0, female = 1. Stunting: not stunted = 0, stunted = 1. Preschool: did not attend = 0, attended = 1. Maternal education: no school = 0, started primary = 1, finished primary = 2, started secondary = 3, finished secondary = 4, tertiary = 5. Home learning activities: did not do = 0, did do = 1. The Lao PDR version of the eHCI does not include a Physical Health domain. Phys=Physical Health, Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing

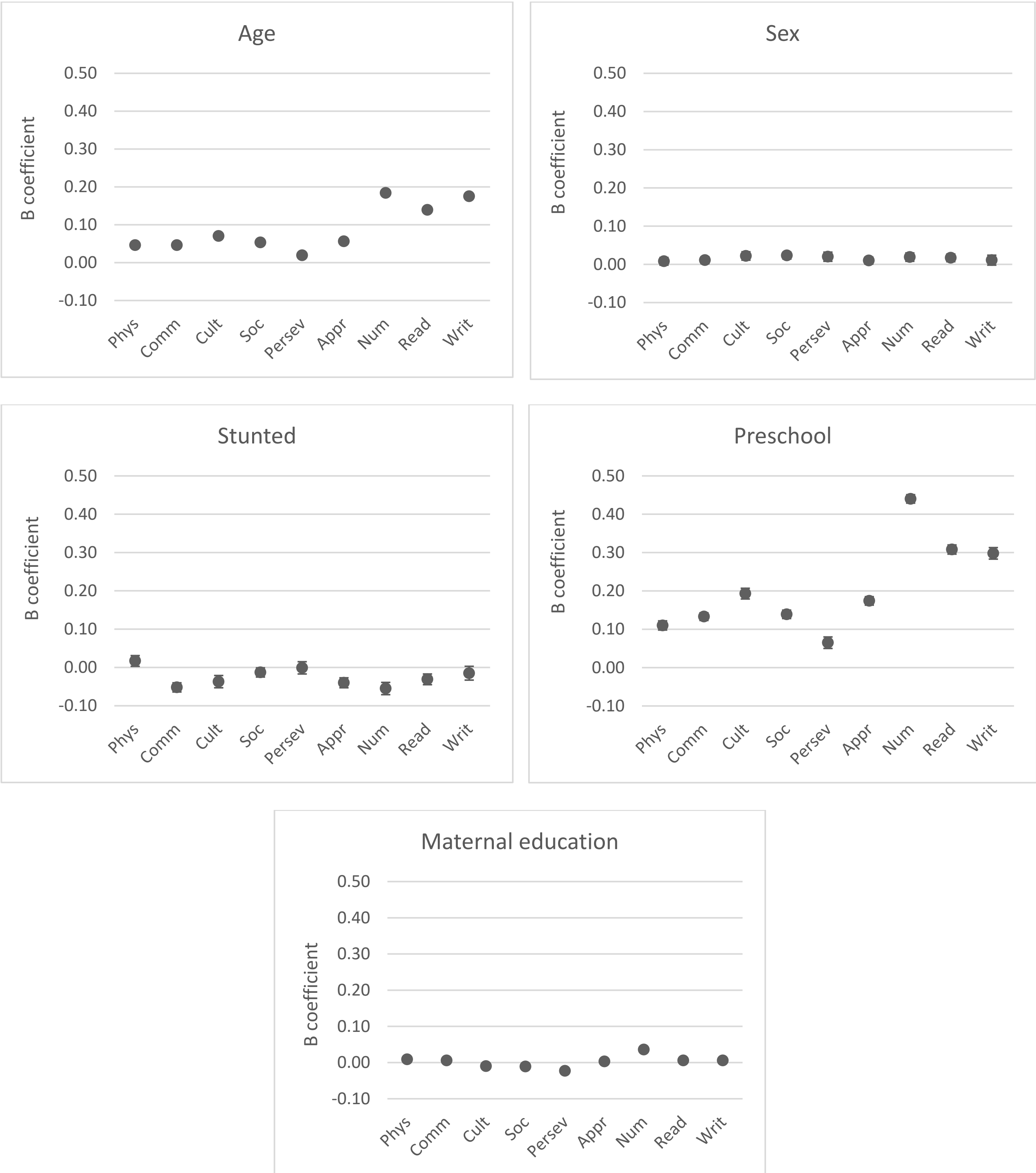
Supplementary Figure 4.1. Discriminant validity: linear regression coefficients and 95% confidence intervals of eHCI domain scores on demographic and contextual variables in Brazil (n = 1810, 2015)





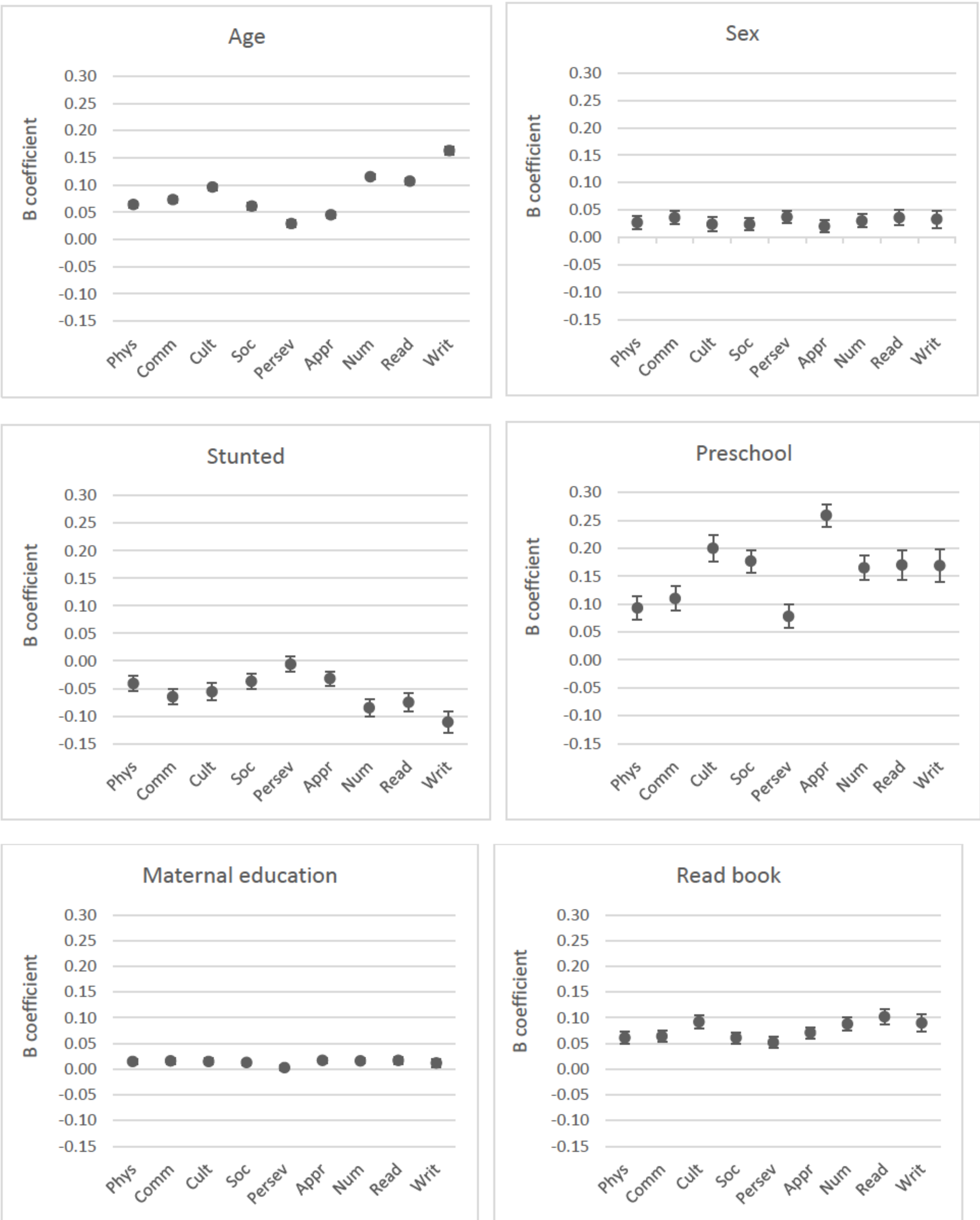
Note: Stunting, maternal education, and home learning environment data were not collected in Brazil. Preschool results are not presented as 100% of the sample were attending preschool. Female = 1, male = 0. Phys=Physical Health, Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing.

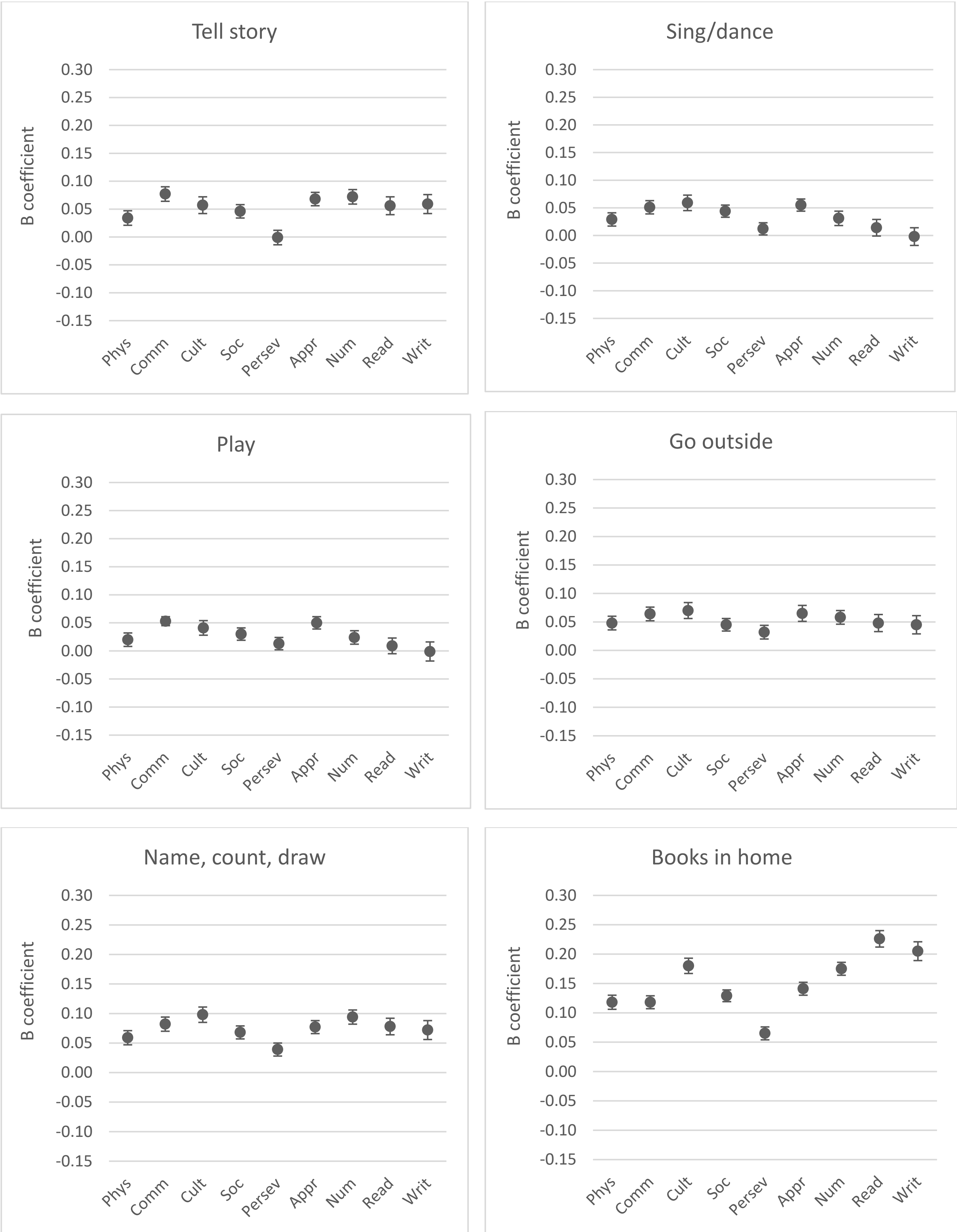
Supplementary Figure 4.2. Discriminant validity: linear regression coefficients and 95% confidence intervals of eHCI domain scores on demographic and contextual variables in China (n = 11421, 2015/16)



Note: Home learning environment data were not collected in China. Female, stunted, attended preschool = 1; male, not stunted, did not attend preschool = 0. Phys=Physical Health, Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing.

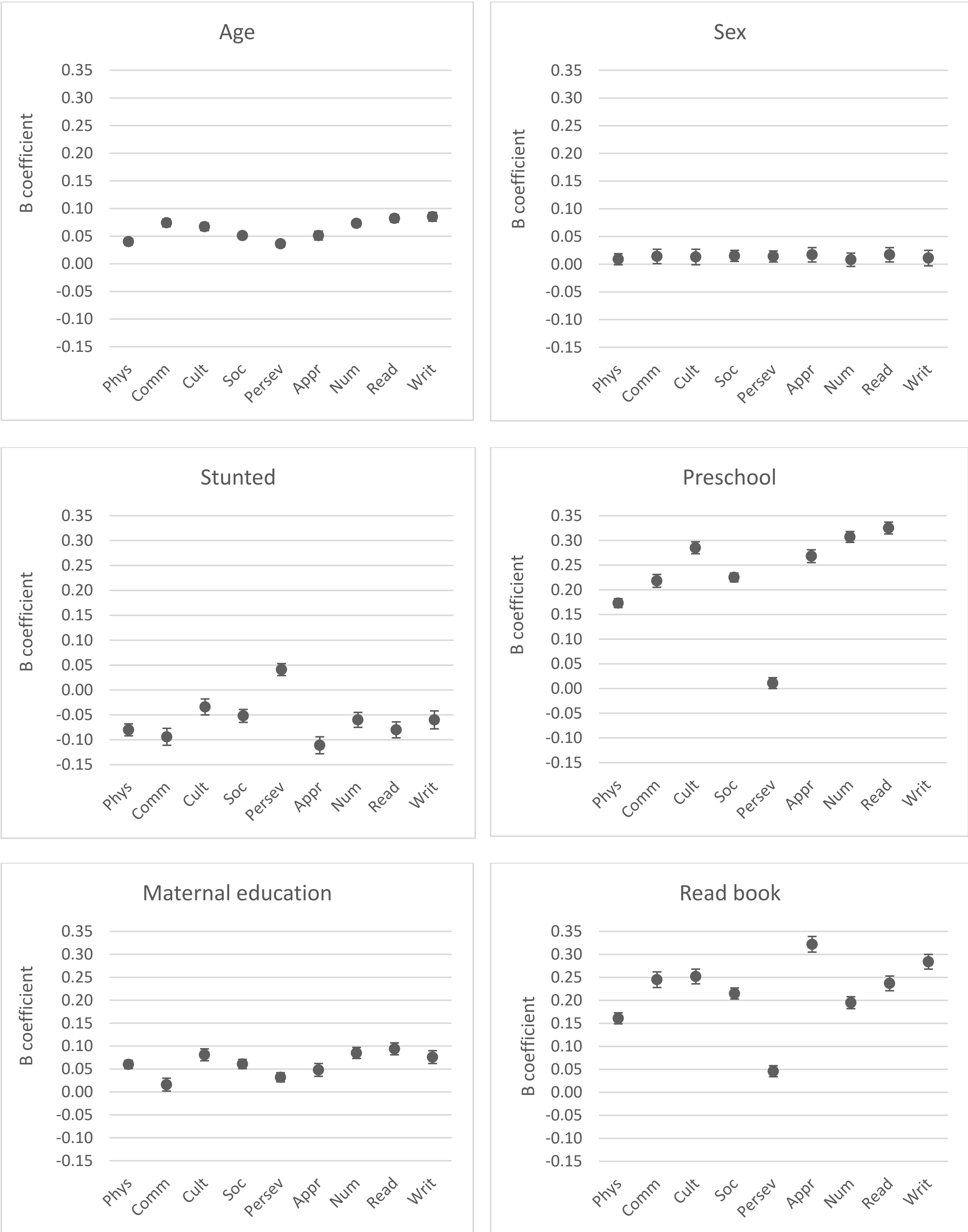
Supplementary Figure 4.3. Discriminant validity: linear regression coefficients and 95% confidence intervals of eHCI domain scores on demographic and contextual variables in Kiribati (n = 8339, 2017)

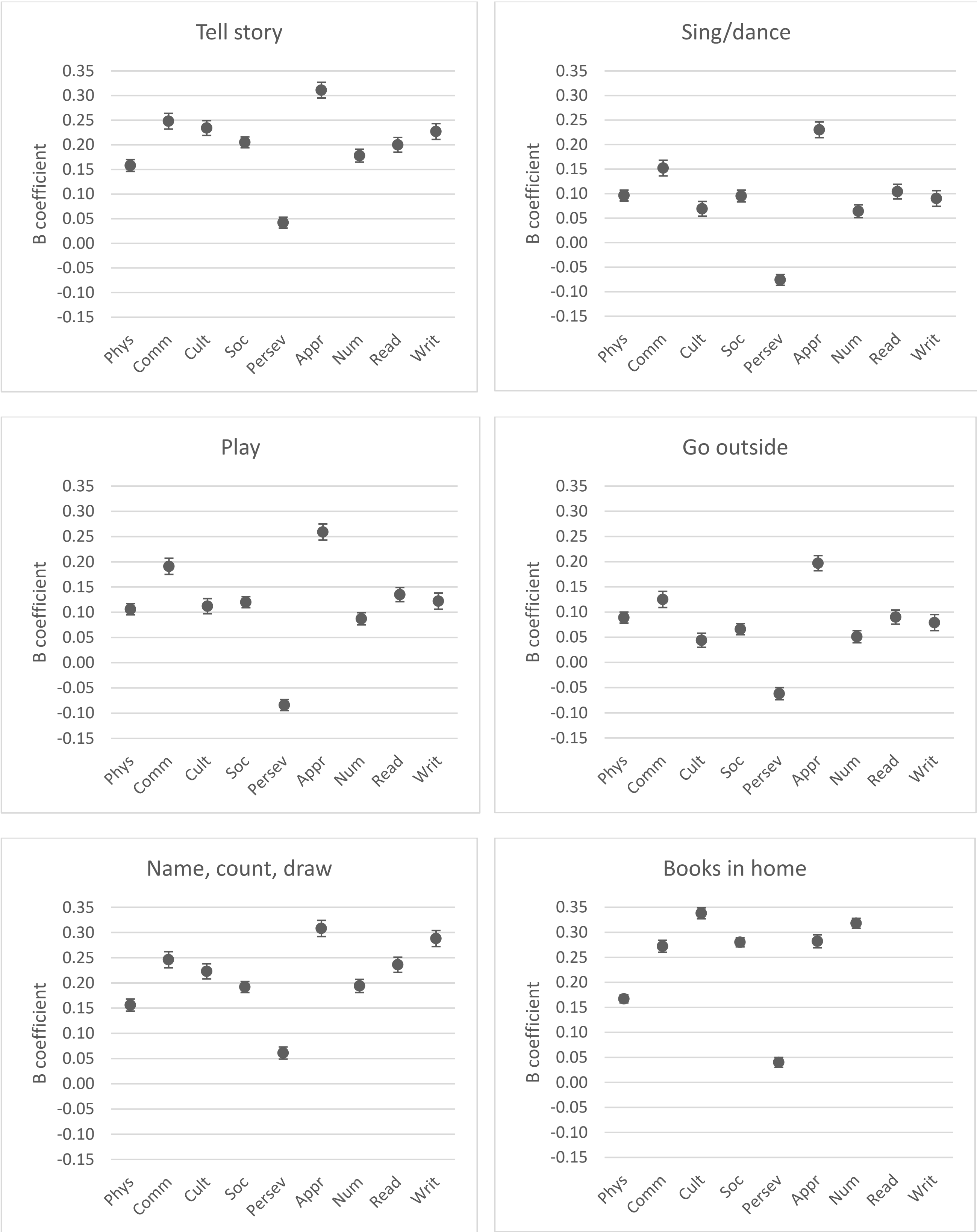




Note. Female, stunted, attended preschool, home learning environment items yes = 1; male, not stunted, did not attend preschool, home learning environment items no = 0. Phys=Physical Health, Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing.

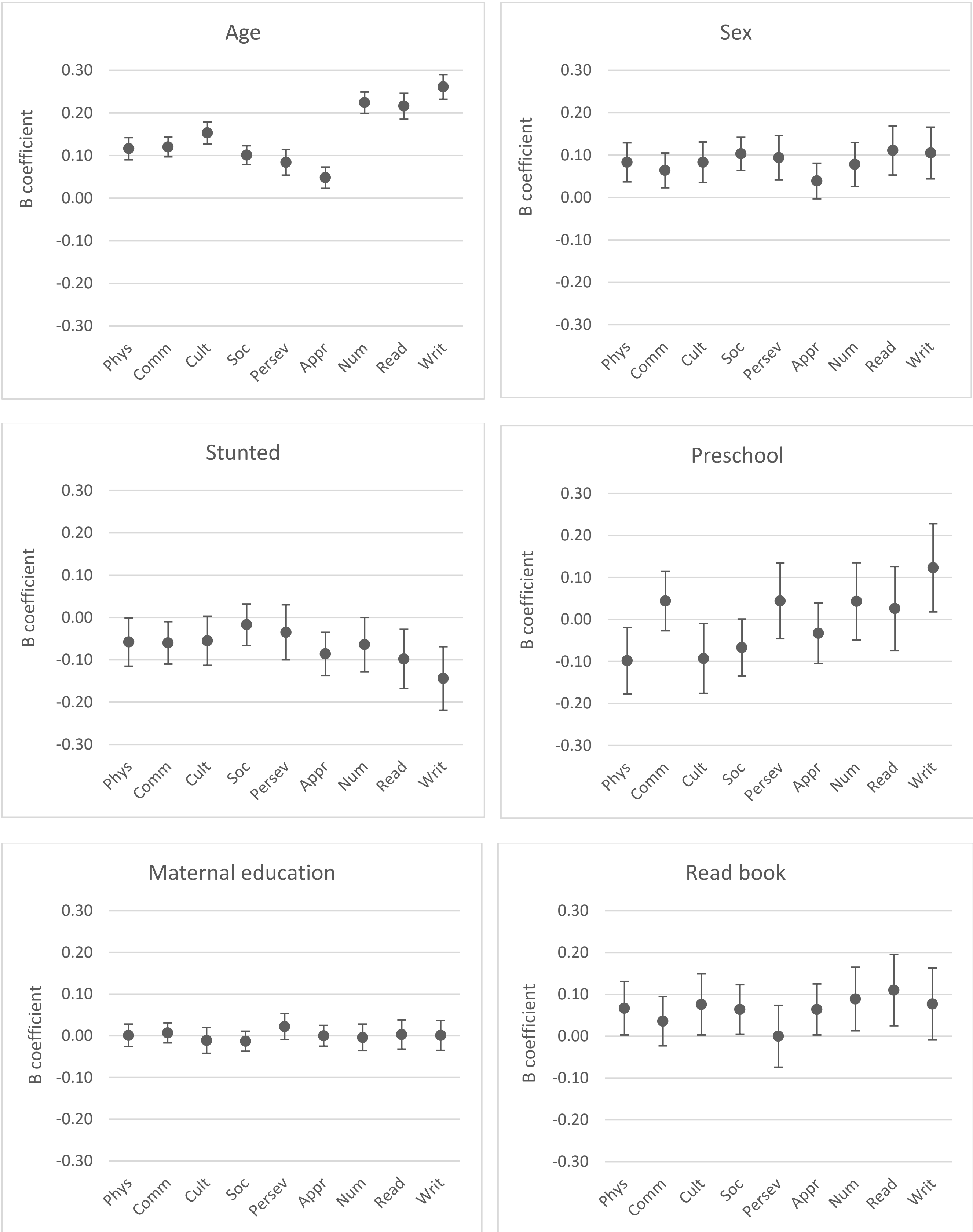
Supplementary Figure 4.4. Discriminant validity: linear regression coefficients and 95% confidence intervals of eHCI domain scores on demographic and contextual variables in Samoa (n = 12191, 2016)

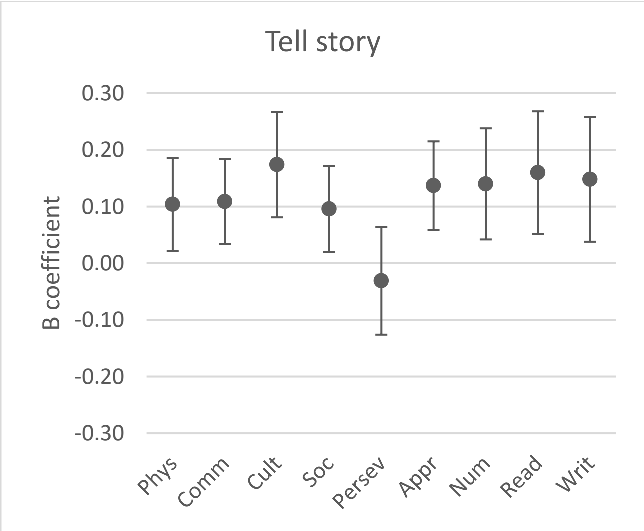


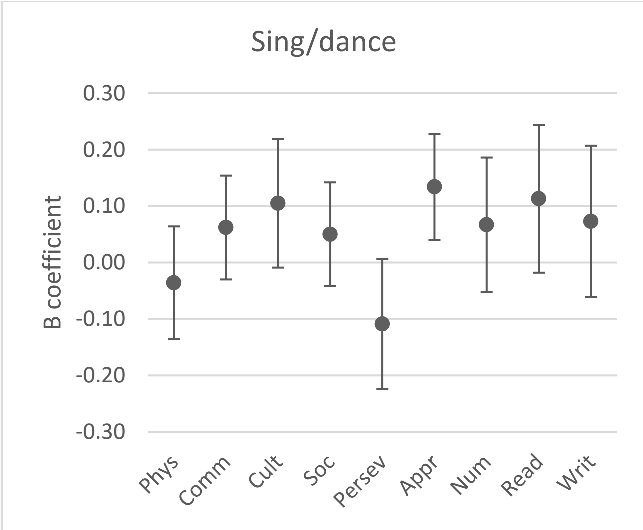


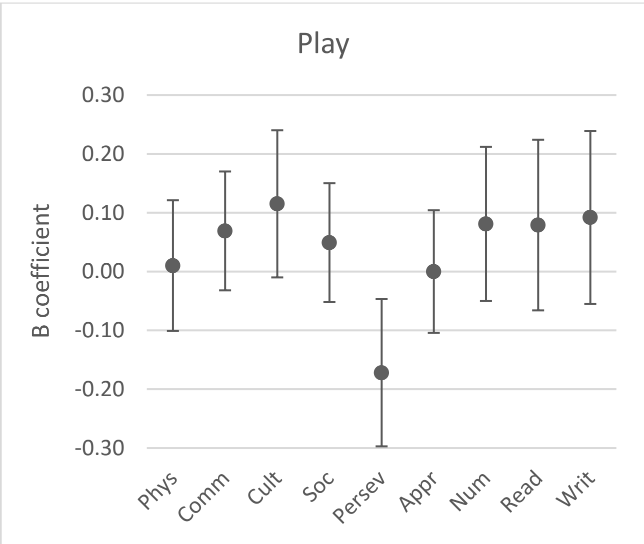
Note. Female, stunted, attended preschool, home learning environment items yes = 1; male, not stunted, did not attend preschool, home learning environment items no = 0. Phys=Physical Health, Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing.

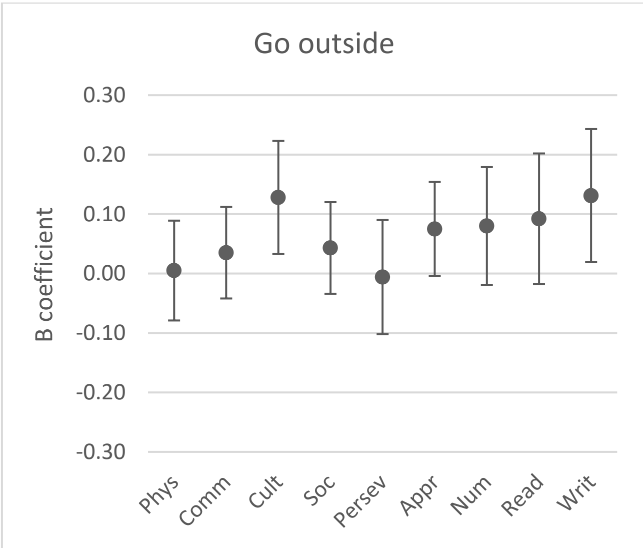
Supplementary Figure 4.5. Discriminant validity: linear regression coefficients and 95% confidence intervals of eHCI domain scores on demographic and contextual variables in Tuvalu (n = 549, 2015)

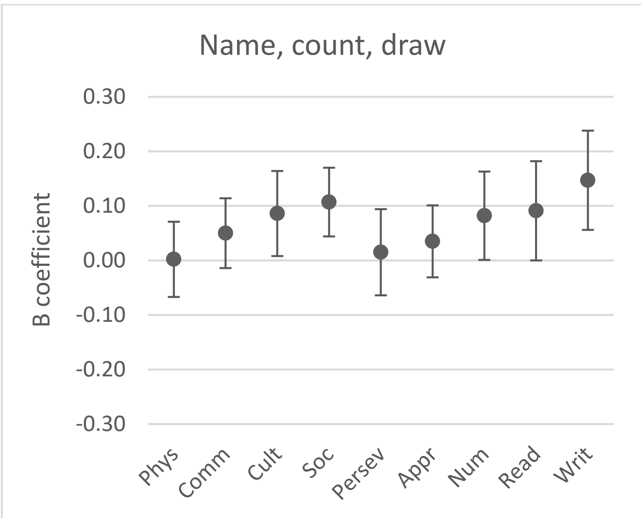












Note: Information regarding books in the home were not collected in Tuvalu. Female, stunted, attended preschool, home learning environment items yes = 1; male, not stunted, did not attend preschool, home learning environment items no = 0. Phys=Physical Health, Comm=Verbal Communication, Cult=Cultural Knowledge, Soc=Social and Emotional, Persev=Perseverance, Appr=Approaches to Learning, Num=Numeracy, Read=Reading, and Writ=Writing.

Appendix F: Supplementary materials from submitted journal article 3

Supplementary Table 5.1. Early Human Capability Index (eHCI) domains and items

Domain	Item
Verbal Communication	Can child tell you what he/she wants? Can child speak a few simple words or sentences to explain what happened to him/her? Can child speak many words or sentences to explain what happened to him/her? Can child communicate well with you on any topics?
Cultural Knowledge	Can child tell a type of animal, at least two animals in the area? Can child tell a food name, at least two dishes that are available in the area? Can child tell a name of a plant/vegetable/fruit, at least two types in the area? Can child sing? Can child participate in traditional events such as giving alms and seeing monks?
Social and Emotional	Is child happy to share his/her toys with others? Does child know how to take care of his/her belongings? Has child shown respect to elders? Does child respect other kids? Is child responsible for his/her own behaviour? Does child consider other people's feelings? Can child help other people? Is child friendly to other kids? Is child hot tempered?* Can child understand the difference between right and wrong? Can child respond to a simple instruction? Is your child very clingy (i.e. does not want to leave their parent's side)?* Does your child understand his/her feelings and is able to describe his/her feelings, for example "I'm happy" or "I'm sad"? Does your child stop an activity when told to do so straight away?
Perseverance	Can child do something on his/her own? When child is doing something, does he/she finish it? Does child have to be told several times so that then he/she can finish what he/she is doing?* When child is doing something, does he/she lose focus easily?*
Approaches to Learning	Does child show any sign of interest to learn new things? Does child try to learn how to play with new toys? Has child ever used any object to role play (using imagination) e.g. banana stem horse, sword fighting etc? Does child show interest in playing games such as collecting pebbles, jumping with rubber bands, hide and seek, ball throwing etc? When the kids go somewhere with you or another household member, are they brave to survey or ask some questions with another person? When child is doing a particular activity, they will pay intense attention to the activity.

	Does child think things out before starting a task?
Numeracy	Can child distinguish between a triangle, circle and rectangle?
	Can child tell at least three different colours?
	Can child distinguish objects based on shape, colour, and size?
	Can child count from 1 to 10?
	Can child count from 1 to 20?
	Does child know that a tiger is taller than a cat?'
	Does child know morning, afternoon and evening?
	Can child tell you if it is yesterday, today or tomorrow?
	Does child know that an elephant is heavier than a pig?
	Does child know that 8 is more than 2?
Reading	Can child read in the correct direction from left to right and from top to bottom (even if they can't read)?
	Can child tell at least 3 letters of the alphabet?
	Can child tell at least 10 letters of the alphabet?
	Can child read at least 4 simple words?
	Can child read difficult words such as axe, buffalo?
	Can child read simple sentences?
Writing	Can child write or draw with pencil, coloured pencils or pen?
	Can child draw a picture that you can tell what it is?
	Can child write at least three letters such as A, B, C?
	Can child write his/her own name?
	Can child write a simple word?
	Can child write a simple sentence?

Note. * = reverse scored items.

Supplementary Table 5.2. Direct assessment domains and sub-domains

Domain	Sub-domain	Number of items
Literacy	Familiarity with print	3
	Sound discrimination	3
	Sound identification	5
	Letter name knowledge	20
	Listening comprehension	5
	Name writing	1
Numeracy	Quantitative comparison	1
	Measurement vocabulary	4
	Shape naming	4
	Spatial vocabulary	4
	Verbal counting	1
	Number identification	20
	Producing a set	3
	Addition with two sets	1
	Spatial visualisation	1
Executive function	Forward digit span	4
	Backward digit span	7
	Head-toes-knees-shoulders	5

Supplementary Table 5.3. Descriptive statistics for eHCI and DA scores at ages 2-5 and ages 6-9

		Baseline											
		2 years (n = 1,325)			3 years (n = 1,685)			4 years (n = 1,739)			5 years (n = 473)		
	Domains	M	SD	IQR	M	SD	IQR	M	SD	IQR	M	SD	IQR
eHCI	Verbal Communication	0.71	0.32	0.50	0.85	0.24	0.25	0.89	0.21	0.00	0.92	0.18	0.00
	Cultural Knowledge	0.59	0.25	0.40	0.70	0.20	0.20	0.75	0.18	0.20	0.79	0.18	0.40
	Social and Emotional	0.37	0.18	0.29	0.46	0.19	0.21	0.51	0.19	0.29	0.54	0.20	0.36
	Perseverance	0.43	0.25	0.25	0.49	0.27	0.50	0.52	0.26	0.25	0.58	0.25	0.25
	Approaches to Learning	0.56	0.30	0.57	0.66	0.28	0.43	0.72	0.24	0.29	0.74	0.25	0.29
	Numeracy	0.19	0.20	0.30	0.32	0.22	0.30	0.42	0.24	0.30	0.51	0.26	0.40
	Reading	0.04	0.11	0.00	0.08	0.17	0.17	0.13	0.20	0.17	0.21	0.26	0.33
	Writing	0.09	0.11	0.17	0.13	0.14	0.17	0.18	0.18	0.17	0.27	0.24	0.17
	Overall development	0.34	0.13	0.17	0.43	0.13	0.16	0.48	0.13	0.16	0.54	0.15	0.20
DA	Literacy	0.03	0.05	0.03	0.05	0.08	0.07	0.08	0.11	0.12	0.12	0.16	0.18
	Numeracy	0.12	0.14	0.22	0.20	0.17	0.31	0.31	0.20	0.23	0.39	0.23	0.31
	Executive function	0.06	0.12	0.06	0.10	0.16	0.17	0.18	0.20	0.33	0.24	0.24	0.39
		Follow up											
		6 years			7 years			8 years			9 years		
	Domains	M	SD	IQR	M	SD	IQR	M	SD	IQR	M	SD	IQR
eHCI	Verbal Communication	0.90	0.22	0.00	0.90	0.22	0.00	0.91	0.22	0.00	0.93	0.20	0.00
	Cultural Knowledge	0.82	0.21	0.20	0.84	0.19	0.20	0.86	0.19	0.20	0.89	0.16	0.20
	Social and Emotional	0.70	0.20	0.29	0.73	0.19	0.25	0.76	0.19	0.29	0.79	0.16	0.21
	Perseverance	0.56	0.24	0.25	0.58	0.23	0.25	0.59	0.22	0.25	0.61	0.22	0.25
	Approaches to Learning	0.79	0.23	0.29	0.79	0.24	0.29	0.79	0.23	0.29	0.82	0.21	0.29
	Numeracy	0.83	0.21	0.30	0.87	0.18	0.20	0.89	0.18	0.20	0.93	0.14	0.10
	Reading	0.59	0.34	0.50	0.66	0.33	0.67	0.72	0.31	0.50	0.81	0.27	0.33
	Writing	0.71	0.30	0.50	0.79	0.27	0.33	0.84	0.23	0.33	0.91	0.19	0.17
	Overall development	0.72	0.17	0.25	0.76	0.16	0.22	0.79	0.16	0.19	0.83	0.13	0.12
DA	Literacy	0.37	0.27	0.45	0.47	0.27	0.41	0.54	0.27	0.43	0.65	0.27	0.44
	Numeracy	0.64	0.23	0.37	0.73	0.21	0.29	0.79	0.19	0.24	0.82	0.16	0.19
	Executive function	0.36	0.18	0.25	0.43	0.19	0.28	0.48	0.19	0.29	0.54	0.19	0.25

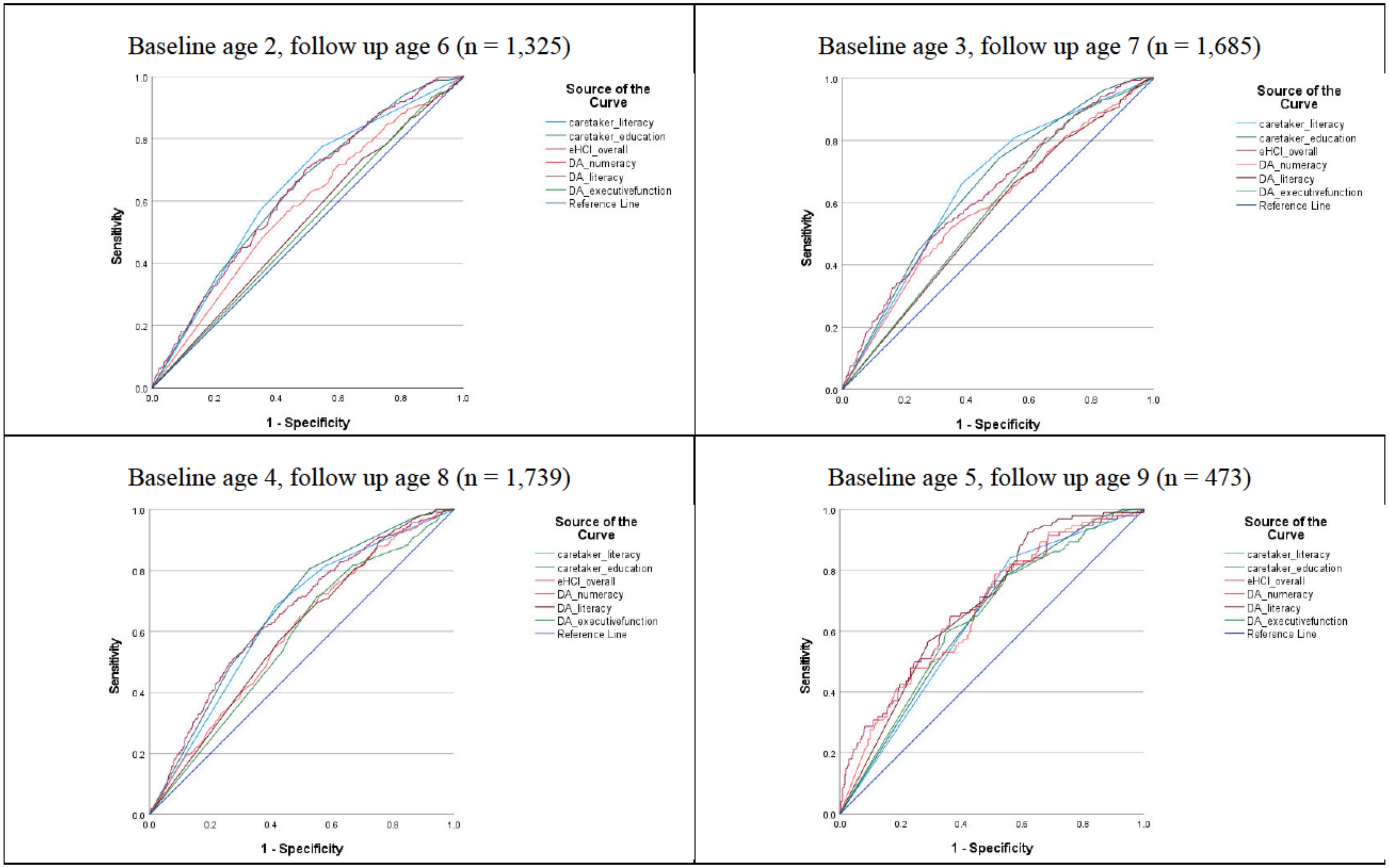
Note. eHCI = early Human Capability index, DA = direct assessment, M = mean, SD = standard deviation, IQR = interquartile range.

Supplementary Table 5.4. AUC (95% CI) for eHCI, DA, and socioeconomic measures at ages 2-5 predicting DA scores in the bottom 20% at ages 6-9

	DA literacy	DA numeracy	DA executive function
Age 2 (n = 1,325)		Age 6	
eHCI overall development	0.63 (0.59-0.66)	0.67 (0.63-0.70)	0.63 (0.60-0.67)
DA literacy	0.53 (0.49-0.57)	0.54 (0.51-0.58)	0.54 (0.50-0.58)
DA numeracy	0.58 (0.54-0.61)	0.60 (0.56-0.64)	0.59 (0.56-0.63)
DA executive function	0.52 (0.48-0.56)	0.55 (0.51-0.59)	0.53 (0.49-0.57)
Caregiver education	0.63 (0.60-0.67)	0.63 (0.59-0.66)	0.60 (0.56-0.64)
Caregiver literacy	0.64 (0.60-0.67)	0.64 (0.61-0.68)	0.60 (0.57-0.64)
Age 3 (n = 1,685)		Age 7	
eHCI overall development	0.63 (0.60-0.66)	0.60 (0.57-0.63)	0.60 (0.56-0.63)
DA literacy	0.56 (0.53-0.59)	0.56 (0.53-0.60)	0.54 (0.51-0.58)
DA numeracy	0.59 (0.56-0.63)	0.59 (0.56-0.63)	0.58 (0.55-0.62)
DA executive function	0.58 (0.55-0.61)	0.56 (0.53-0.60)	0.55 (0.51-0.58)
Caregiver education	0.65 (0.62-0.69)	0.63 (0.60-0.67)	0.59 (0.56-0.63)
Caregiver literacy	0.66 (0.62-0.69)	0.64 (0.60-0.67)	0.61 (0.58-0.64)
Age 4 (n = 1,739)		Age 8	
eHCI overall development	0.65 (0.62-0.69)	0.65 (0.62-0.68)	0.62 (0.58-0.65)
DA literacy	0.60 (0.67-0.63)	0.60 (0.57-0.63)	0.60 (0.57-0.63)
DA numeracy	0.59 (0.56-0.63)	0.63 (0.60-0.66)	0.62 (0.59-0.65)
DA executive function	0.58 (0.54-0.61)	0.59 (0.56-0.62)	0.58 (0.54-0.61)
Caregiver education	0.67 (0.64-0.70)	0.64 (0.61-0.68)	0.62 (0.59-0.65)
Caregiver literacy	0.65 (0.62-0.68)	0.64 (0.61-0.67)	0.61 (0.58-0.65)
Age 5 (n = 473)		Age 9	
eHCI overall development	0.68 (0.62-0.74)	0.66 (0.59-0.72)	0.61 (0.54-0.67)
DA literacy	0.68 (0.63-0.74)	0.65 (0.59-0.71)	0.58 (0.52-0.64)
DA numeracy	0.67 (0.61-0.73)	0.68 (0.62-0.74)	0.60 (0.54-0.66)
DA executive function	0.64 (0.58-0.70)	0.61 (0.55-0.68)	0.52 (0.46-0.59)
Caregiver education	0.64 (0.58-0.70)	0.60 (0.54-0.66)	0.58 (0.52-0.64)
Caregiver literacy	0.64 (0.58-0.70)	0.59 (0.52-0.65)	0.59 (0.52-0.65)

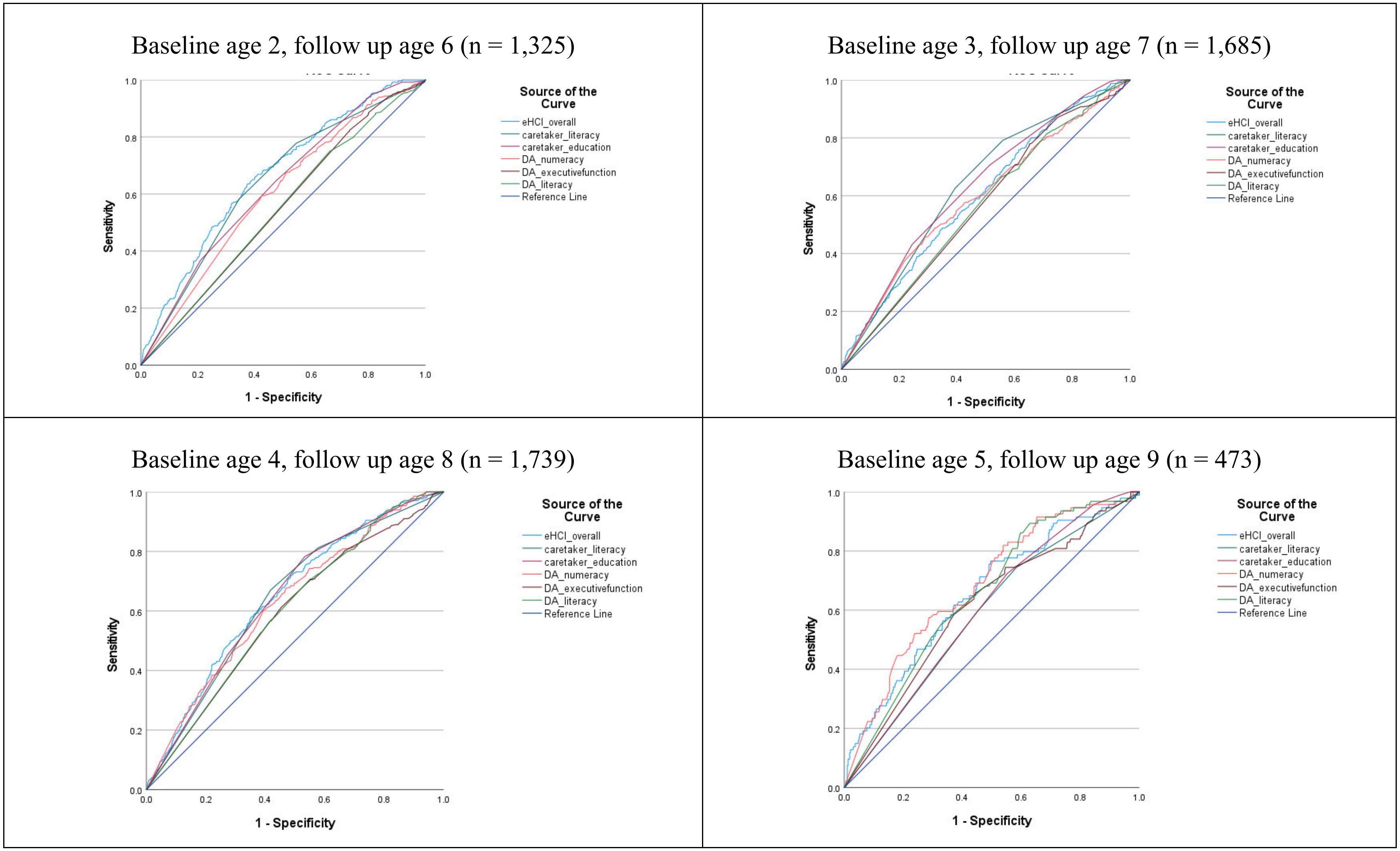
Note. AUC = Area Under the Curve, CI = confidence interval, eHCI = early Human Capability index, DA = Direct Assessment.

Supplementary Figure 5.1. ROC curves for eHCI, DA, and socioeconomic measures at ages 2-5 predicting literacy scores in the bottom 20% at ages 6-9



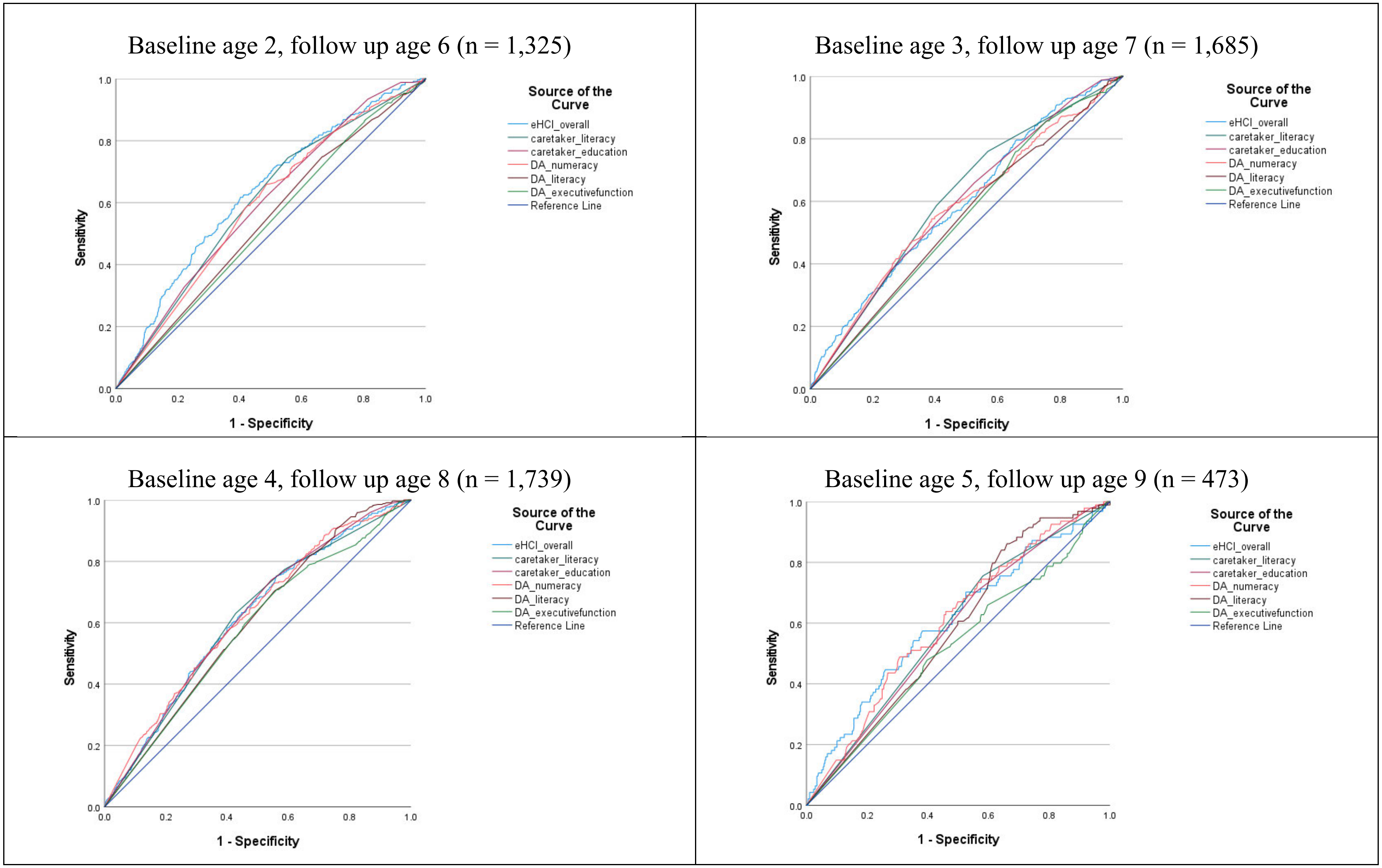
Note. eHCI = early Human Capability Index, DA = direct assessment.

Supplementary Figure 5.2. ROC curves for eHCI, DA, and socioeconomic measures at ages 2-5 predicting numeracy scores in the bottom 20% at ages 6-9



Note. eHCI = early Human Capability Index, DA = direct assessment.

Supplementary Figure 5.3. ROC curves for eHCI, DA, and socioeconomic measures at ages 2-5 predicting executive function scores in the bottom 20% at ages 6-9



Note. eHCI = early Human Capability Index, DA = direct assessment.

Supplementary Table 5.5. AUC (95% CI) for individual eHCI and DA domains at ages 2-5 predicting DA scores in the bottom 10% at ages 6-9

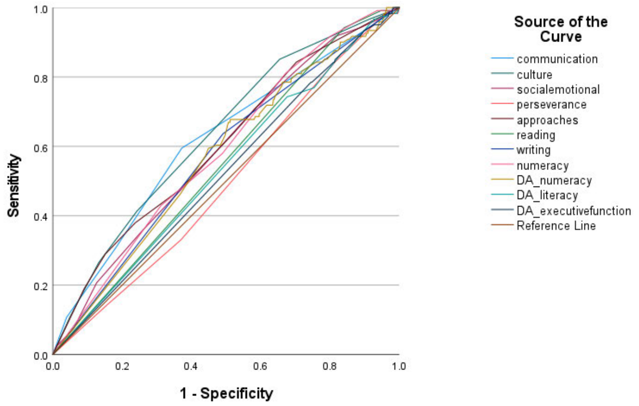
		DA literacy	DA numeracy	DA executive function
Age 2 (n = 1,325)			Age 6	
eHCI	Verbal communication	0.62 (0.56-0.67)	0.63 (0.59-0.68)	0.61 (0.56-0.66)
	Cultural knowledge	0.64 (0.59-0.69)	0.66 (0.61-0.71)	0.62 (0.57-0.67)
	Social and emotional	0.59 (0.54-0.64)	0.60 (0.55-0.65)	0.58 (0.53-0.63)
	Perseverance	0.50 (0.44-0.55)	0.52 (0.47-0.56)	0.50 (0.46-0.55)
	Approaches to learning	0.60 (0.55-0.66)	0.62 (0.57-0.67)	0.61 (0.56-0.66)
	Numeracy	0.55 (0.50-0.60)	0.57 (0.53-0.62)	0.55 (0.51-0.60)
	Reading	0.57 (0.52-0.63)	0.58 (0.53-0.63)	0.57 (0.52-0.61)
	Writing	0.59 (0.54-0.64)	0.60 (0.55-0.64)	0.62 (0.58-0.67)
DA	Literacy	0.53 (0.48-0.59)	0.55 (0.05-0.60)	0.52 (0.47-0.57)
	Numeracy	0.57 (0.52-0.62)	0.62 (0.57-0.67)	0.58 (0.53-0.62)
	Executive function	0.52 (0.47-0.57)	0.55 (0.51-0.60)	0.53 (0.48-0.57)
Age 3 (n = 1,685)			Age 7	
eHCI	Verbal communication	0.55 (0.48-0.63)	0.55 (0.51-0.60)	0.56 (0.51-0.60)
	Cultural knowledge	0.59 (0.52-0.66)	0.60 (0.55-0.64)	0.59 (0.55-0.64)
	Social and emotional	0.60 (0.53-0.66)	0.55 (0.50-0.59)	0.59 (0.54-0.63)
	Perseverance	0.55 (0.48-0.62)	0.51 (0.47-0.56)	0.53 (0.49-0.58)
	Approaches to learning	0.60 (0.52-0.67)	0.55 (0.51-0.60)	0.58 (0.54-0.63)
	Numeracy	0.58 (0.52-0.64)	0.57 (0.53-0.61)	0.53 (0.49-0.57)
	Reading	0.66 (0.60-0.72)	0.59 (0.55-0.64)	0.58 (0.53-0.62)
	Writing	0.65 (0.58-0.71)	0.59 (0.54-0.63)	0.59 (0.55-0.64)
DA	Literacy	0.58 (0.51-0.65)	0.57 (0.52-0.61)	0.56 (0.52-0.60)
	Numeracy	0.60 (0.52-0.67)	0.60 (0.56-0.65)	0.61 (0.56-0.65)
	Executive function	0.56 (0.50-0.63)	0.58 (0.53-0.62)	0.57 (0.53-0.61)
Age 4 (n = 1,739)			Age 8	
eHCI	Verbal communication	0.55 (0.51-0.60)	0.57 (0.52-28)	0.56 (0.52-0.61)
	Cultural knowledge	0.59 (0.54-0.63)	0.59 (0.54-28)	0.59 (0.55-0.63)
	Social and emotional	0.60 (0.56-0.65)	0.61 (0.56-28)	0.57 (0.53-0.61)
	Perseverance	0.54 (0.50-0.58)	0.51 (0.47-28)	0.51 (0.47-0.56)
	Approaches to learning	0.56 (0.51-0.60)	0.59 (0.54-28)	0.57 (0.53-0.62)
	Numeracy	0.61 (0.57-0.65)	0.62 (0.58-28)	0.61 (0.57-0.65)
	Reading	0.62 (0.58-0.66)	0.64 (0.59-28)	0.61 (0.57-0.64)
	Writing	0.63 (0.59-0.67)	0.63 (0.59-28)	0.60 (0.56-0.64)
DA	Literacy	0.59 (0.55-0.63)	0.61 (0.57-28)	0.60 (0.57-0.64)
	Numeracy	0.63 (0.59-0.67)	0.63 (0.58-28)	0.63 (0.59-0.66)
	Executive function	0.57 (0.53-0.62)	0.59 (0.54-28)	0.59 (0.55-0.63)

Age 5 (n = 473)		Age 9		
eHCI	Verbal communication	0.53 (0.45-0.62)	0.58 (0.49-0.67)	0.62 (0.53-0.72)
	Cultural knowledge	0.60 (0.52-0.68)	0.58 (0.50-0.67)	0.65 (0.58-0.73)
	Social and emotional	0.65 (0.56-0.73)	0.60 (0.51-0.68)	0.57 (0.48-0.66)
	Perseverance	0.56 (0.48-0.65)	0.57 (0.48-0.65)	0.56 (0.46-0.65)
	Approaches to learning	0.56 (0.48-0.65)	0.57 (0.49-0.66)	0.60 (0.52-0.69)
	Numeracy	0.67 (0.59-0.74)	0.61 (0.53-0.69)	0.58 (0.49-0.67)
	Reading	0.64 (0.57-0.72)	0.60 (0.52-0.67)	0.58 (0.49-0.67)
	Writing	0.65 (0.57-0.73)	0.60 (0.51-0.68)	0.61 (0.52-0.69)
DA	Literacy	0.68 (0.61-0.74)	0.68 (0.60-0.75)	0.62 (0.54-0.69)
	Numeracy	0.66 (0.59-0.73)	0.69 (0.62-0.77)	0.64 (0.56-0.72)
	Executive function	0.60 (0.52-0.68)	0.62 (0.54-0.70)	0.58 (0.49-0.67)

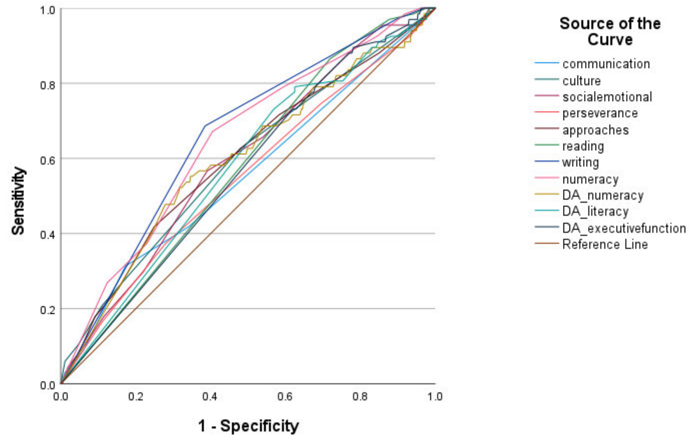
Note. AUC = Area Under the Curve, CI = confidence interval, eHCI = early Human Capability index, DA = Direct Assessment.

Supplementary Figure 5.4. ROC curves for individual eHCI and DA domains at ages 2-5 predicting literacy scores in the bottom 10% at ages 6-9

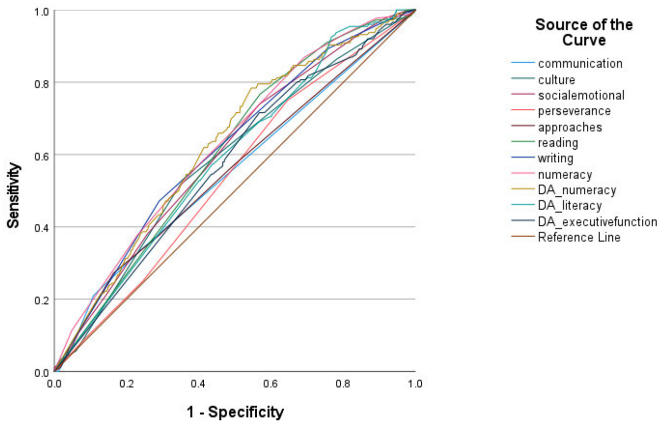
Baseline age 2, follow up age 6 (n = 1,325)



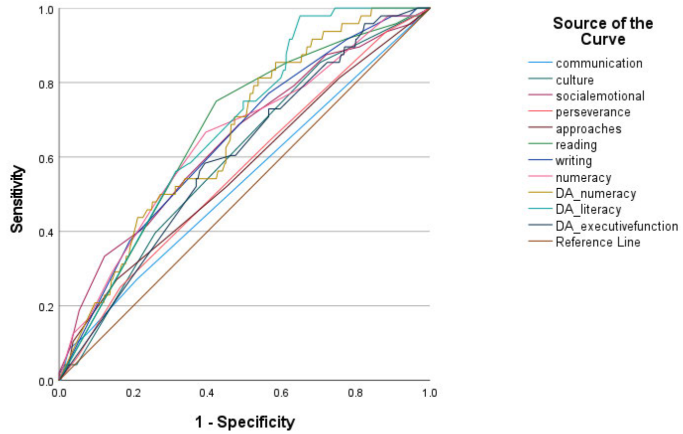
Baseline age 3, follow up age 7 (n = 1,685)



Baseline age 4, follow up age 8 (n = 1,739)



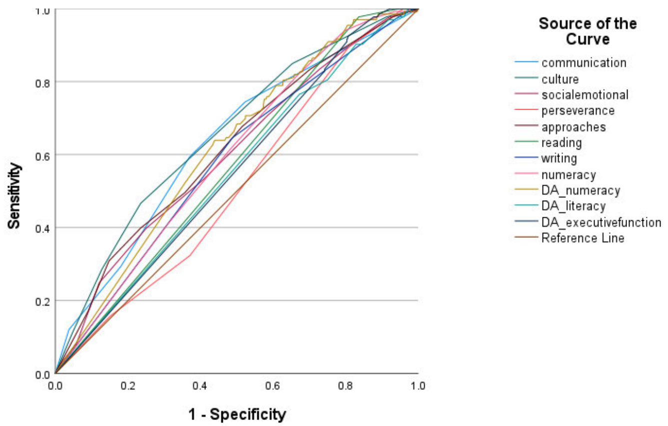
Baseline age 5, follow up age 9 (n = 473)



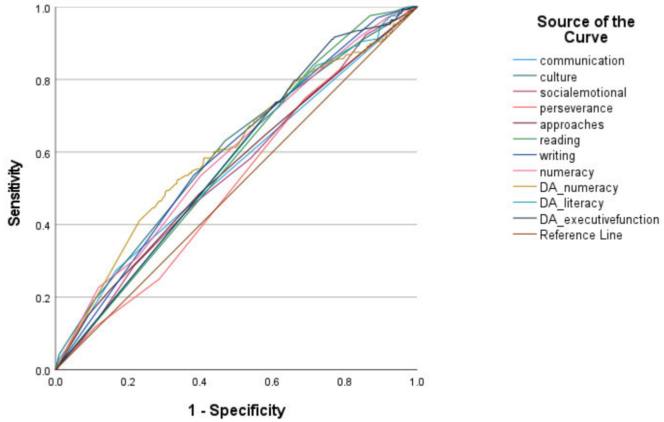
Note. eHCI = early Human Capability Index, DA = direct assessment.

Supplementary Figure 5.5. ROC curves for individual eHCI and DA domains at ages 2-5 predicting numeracy scores in the bottom 10% at ages 6-9

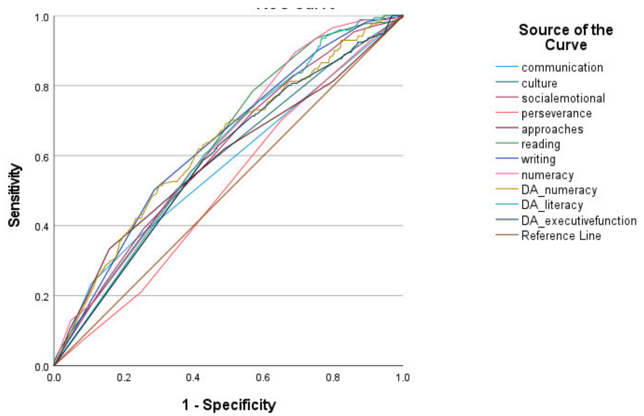
Baseline age 2, follow up age 6 (n = 1,325)



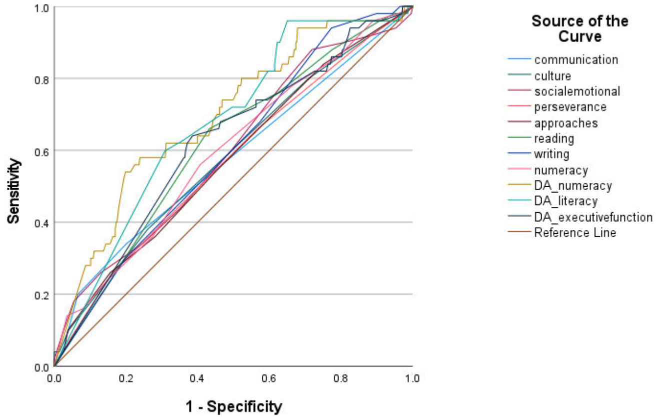
Baseline age 3, follow up age 7 (n = 1,685)



Baseline age 4, follow up age 8 (n = 1,739)

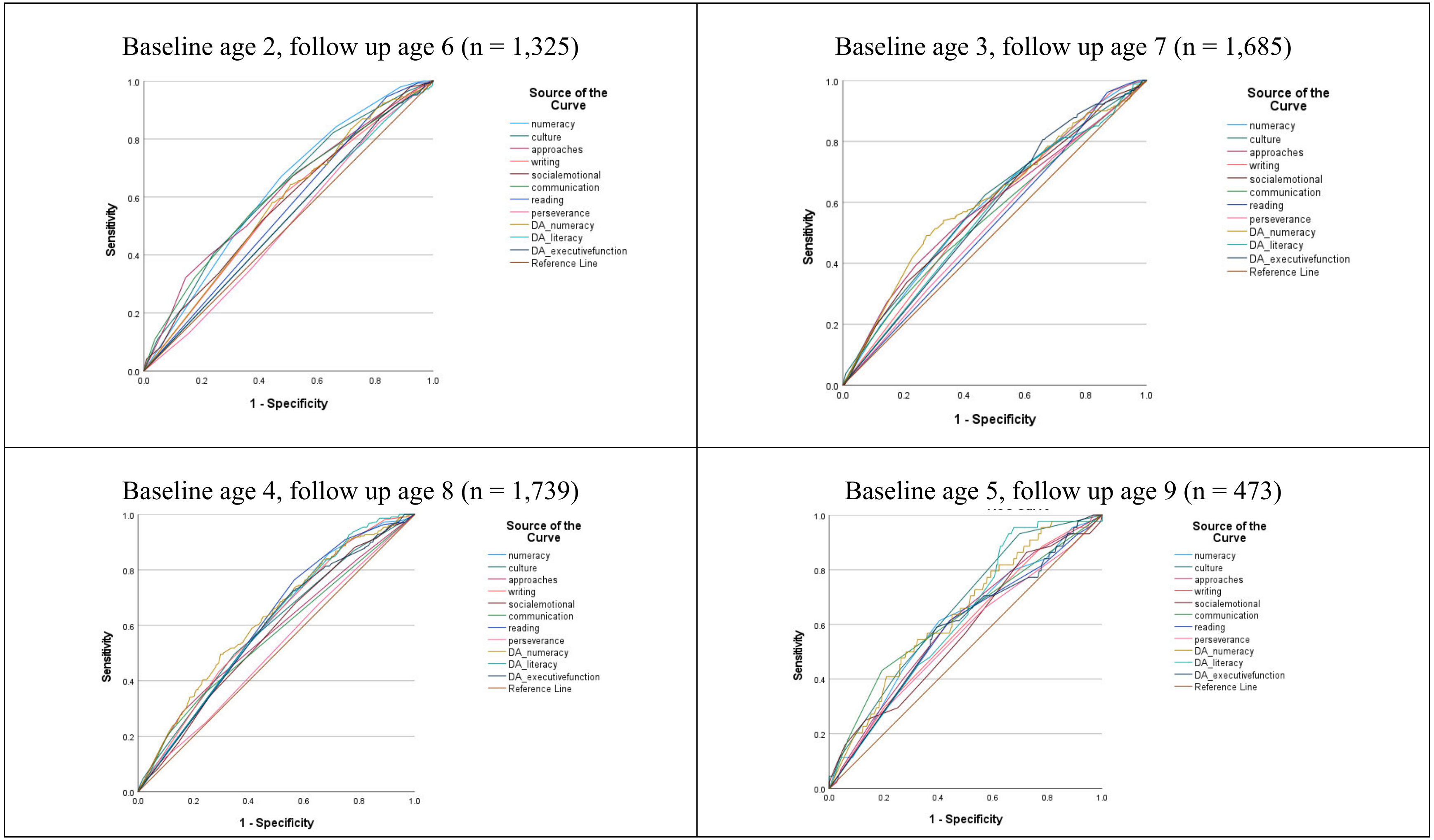


Baseline age 5, follow up age 9 (n = 473)



Note. eHCI = early Human Capability Index, DA = direct assessment.

Supplementary Figure 5.6. ROC curves for individual eHCI and DA domains at ages 2-5 predicting executive function scores in the bottom 10% at ages 6-9



Note. eHCI = early Human Capability Index, DA = direct assessment.

Appendix G: Supplementary materials from submitted journal article 4

Supplementary Table 6.1. Characteristics of full sample (n = 3,322), eligible sample (n = 1,309), and analysis sample (n = 1,168)

Variable		Full sample		Eligible sample		Analysis sample	
		n	(%)	*n*	(%)	*n*	(%)
Age	2 years	550	16.6	65	5.0	60	5.1
	3 years	714	21.5	309	23.6	285	24.4
	4 years	815	24.5	474	36.2	455	39.0
	5 years	687	20.7	407	31.1	319	27.3
	6 years	556	16.7	54	4.1	49	4.2
	Missing	0	0.0	0	0.0	0	0
Sex	Female	1,597	48.1	665	50.8	595	50.9
	Male	1,705	51.3	644	49.2	573	49.1
	Missing	20	0.6	0	0.0	0	0
Ethnicity	Lao Tai	600	18.1	289	22.1	259	22.2
	Khmu	1,224	36.8	517	39.5	475	40.7
	Hmong	799	24.1	245	18.7	212	18.2
	Other	666	20.0	254	19.4	222	19.0
	Missing	33	1.0	4	0.3	0	0.0
Socioeconomic status	1	683	20.6	207	15.8	172	14.7
	2	646	19.4	239	18.3	212	18.2
	3	718	21.6	289	22.1	260	22.3
	4	710	21.4	308	23.5	279	23.9
	5	532	16.0	262	20.0	245	21.0
	Missing	33	1.0	4	0.3	0	0.0
Stunted	Yes	1,444	43.5	551	42.1	500	42.8
	No	1,741	52.4	713	54.5	668	57.2
	Missing	137	4.1	45	3.4	0	0.0
Village accessibility	Yes	1,783	53.7	726	55.5	652	55.8
	No	1,539	46.3	583	44.5	516	44.2
	Missing	0	0.0	0	0.0	0	0.0
Child health status	Very Healthy	1,908	57.4	751	57.4	667	57.1
	Normal	1,209	36.4	478	36.5	425	36.4
	Unhealthy/Often Sick	175	5.3	77	5.9	73	6.3
	Don't know	5	0.2	3	0.2	3	0.3
	Missing	25	0.8	0	0.0	0	0.0

Note. Socioeconomic status 1 = least disadvantaged, 5 = most disadvantaged. Village accessibility based on item "Can cars access the village during rainy seasons?". Child health status based on item “What do you think is the current health status of the child?”.

Supplementary Table 6.2. Summary of adaptations to the MELE in Lao PDR

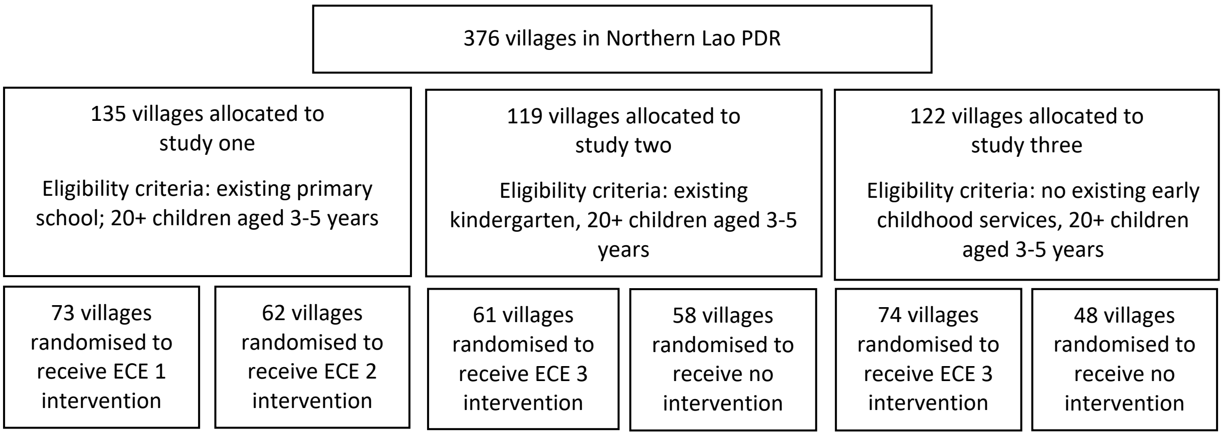
Domain	Adaptation
N/A – basic classroom information	Item around lesson plan now captures both the types of learning teachers use in classrooms (e.g., physical activity, circle activities etc) as well as the topic or subject of learning (e.g., language, maths etc) to align with the early years learning framework in Lao PDR. Items around children and adults present in the classroom now capture details on all children rather than students only (e.g., if there are babies and younger children present) as well as all adults present (e.g., including children's caregivers, CCDG caregivers who lead playgroups but are not teachers, etc).
Learning activities	Music and movement activities are now scored from 1-4 in line with other topics, to reflect its importance in Lao PDR.
Interactions and approaches to learning	Items on child engagement were extended with two additional questions regarding lesson facilitation through explicit instruction and checking on children's understanding – both, through consultations, deemed important aspects of effectively engaging children in learning. An item around creation of supportive learning environments (e.g., using positive language, responding to student needs) has been added. The grouping item has been updated to include two new types of groupings that were observed throughout field testing of the instrument (children working as an individual as part of the whole class, and children working individually as part of a smaller group). The instrument now also requires enumerators record which groupings they observe.
Arrangement, space, and materials	The term portfolio has been replaced with student workbook as this was a term that both enumerators and teachers understood. The item around materials in the classroom now includes additional instructions for enumerators to check if materials are packed away in cupboards. The instrument now also requires enumerators to record the type of materials they observe (e.g., if they select writing utensils, the type of utensils such as pencils needs to be specified). The item on books in the classroom was changed to specify books in Laos language, and books in any other language.
Facilities and safety	Additional safety hazards that are relevant to Lao PDR were incorporated into the safety conditions question, and those that do not apply (e.g., locked doors, items kept in storage) have been removed. The instrument now also requires enumerators to select which (if any) dangerous safety conditions they observe.

Supplementary Table 6.3. Correlations (spearman's rho) among exposure (MELE) and outcome (eHCI, MODEL) variables

Domain	Domain																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
2	.10***																
3	-.12***	.28***															
4	.00	.38***	.28***														
5	-.02	.63***	.70***	.81***													
6	-.09**	.14***	.11***	.16***	.16***												
7	-.06*	.12***	.09***	.18***	.15***	.33***											
8	-.09**	.13***	.14***	.18***	.18***	.35***	.53***										
9	-.02	-.03	-.01	-.01	-.01	.08**	.20***	.27***									
10	-.09**	.12***	.06*	.17***	.13***	.40***	.39***	.43***	.15***								
11	.02	.15***	.15***	.17***	.18***	.29***	.47***	.55***	.24***	.37***							
12	.03	.08**	.06	.09**	.09**	.19***	.39***	.44***	.26***	.34***	.60***						
13	-.06*	.10***	.15***	.13***	.16***	.17***	.41***	.46***	.26***	.31***	.59***	.59***					
14	-.07*	.15***	.14***	.19***	.19***	.51***	.67***	.75***	.45***	.62***	.79***	.73***	.71***				
15	-.11***	.07*	.02	.10***	.08**	.17***	.17***	.23***	.13***	.21***	.32***	.28***	.31***	.35***			
16	-.09**	.11***	.07*	.19***	.16***	.18***	.24***	.27***	.10***	.24***	.35***	.31***	.36***	.39***	.60***		
17	-.01	.10***	.05	.14***	.12***	.15***	.19***	.20***	.08**	.15***	.31***	.30***	.32***	.32***	.56***	.56***	
18	-.08**	.10***	.05	.16***	.13***	.19***	.23***	.26***	.12***	.22***	.38***	.35***	.38***	.40***	.86***	.77***	.87***

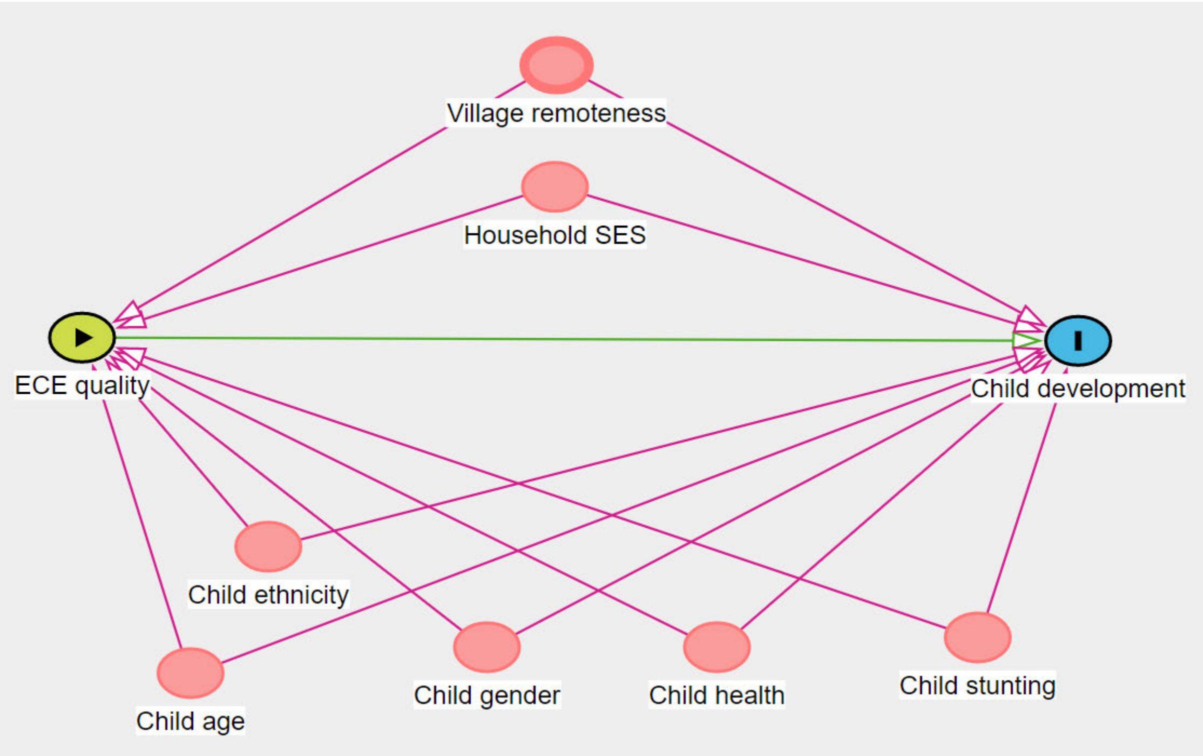
Note. MELE = Measure of Early Learning Environment, eHCI = early Human Capability Index, MODEL = Measurement of Development and Early Learning. 1 = MELE Learning Activities, 2 = MELE Interactions and Approaches, 3 = MELE Classroom Arrangement, 4 = MELE Facilities and Safety, 5 = MELE Overall Quality, 6 = eHCI Verbal communication, 7 = eHCI Cultural knowledge, 8 = eHCI Social and emotional, 9 = eHCI Perseverance, 10 = eHCI Approaches to learning, 11 = eHCI Numeracy, 12 = eHCI Reading, 13 = eHCI Writing, 14 = eHCI Overall Development, 15 = MODEL Numeracy, 16 = MODEL Literacy, 17 = MODEL Executive Function, 18 = MODEL Overall Development.
*$p < .05$. **$p < .01$.

Supplementary Figure 6.1. Study design of Early Childhood Education Project in Lao PDR



Note. ECE = early childhood education.

Supplementary Figure 6.2. Directed Acyclic Graph of the association between early childhood education quality and child development



Note. ECE = early childhood education, SES = socioeconomic status

Appendix H: Additional published journal article contributing to this thesis

THE LANCET
Child & Adolescent Health

COMMENT | VOLUME 4, ISSUE 5, P352-354, MAY 01, 2020

Overcoming challenges in measuring early childhood development across cultures

Nirmala Rao • Sheena Mirpuri • Alanna Sincovich • Sally A Brinkman

Published: February 11, 2020 • DOI: https://doi.org/10.1016/S2352-4642(20)30026-2 • Check for updates

Early childhood development refers to a period of development from conception to age 8 years, wherein children begin to acquire a range of physical, cognitive, language, social, and emotional skills. In consort with the mounting science evidencing the importance of early childhood development, countries have been increasingly investing in programmes to support children and families with young children. To identify whether these investments are effective, there is a need to develop instruments that can efficiently monitor childhood development at a population level. However, measuring early childhood development is challenging because of its dynamic nature; developmental domains are individually complex and inter-related. As children mature, their rate of development is not constant and they have both growth spurts and periods of slow or stagnant growth. These challenges in measuring early childhood development are exacerbated when making comparisons across cultures. Measurement concerns are particularly pertinent because of Target 4.2 of the UN's Sustainable Development Goals which focuses exclusively on early childhood development.

Population (as opposed to diagnostic) measures have unique methodological and psychometric challenges. These measures should simultaneously capture multiple interdependent developmental domains. At the same time, these measures should be efficient to implement and sensitive enough to detect change in early childhood development caused by policy or service changes. Population-level measures should be culturally neutral to compare children across vastly different contexts and cultures, yet also reflect the values and skills within the local culture or national framework. However, variations in cultural values and norms have implications for the measurement of early childhood development because these differences affect both the interpretation by clinicians and researchers and the response of children to items that assess this development.

We provide some brief examples to highlight these challenges. Created by high-income western countries, a common assessment item is whether children come to school dressed appropriately, which intends to capture disorganised children, but instead captures those from poor families who cannot afford so-called appropriate clothes. Another used assessment item is whether the child stops a quarrel or dispute, which is considered a positive attribute in western cultures; however, in the Australian Aboriginal culture and some Asian cultures, not getting involved in other people's disputes is a sign of respect. Chinese children perform better on tests of early arithmetic than children from other Asian countries,[1] probably because of the simplicity of the Chinese number word system (the absence of teens) and the use of number words in daily life (eg, weekday one and weekday two, uncle one and uncle two, and so on).

It is widely recognised that there is no such thing as a test that is fair for, and measures the same concept in, all cultures. Thus, an overarching goal is to reduce existing cultural biases in tools that measure early childhood development. A common strategy is to have a few assessment items that are considered to be culturally neutral; however, this pruning might result in an instrument that lacks sensitivity to local interventions and policies and does not reflect the important attributes of early childhood development within the local culture.

Although all aspects of reliability and validity are important, we posit that two specific aspects are the most important: predictive validity and sensitivity to change. Early childhood development measures based on milestones and pass or fail outcomes do not have the sensitivity required for the evaluation of early childhood interventions and the range to measure impact. In particular, early stimulation programmes that intend to improve the capability of children will rarely be shown to have impact when evaluated with items based on developmental milestones. A deeper understanding of the childhood capabilities strengthened by these interventions and the causal mechanisms that explain the impact of programmes that focus on health, nutrition, and stimulation is essential to promote further investment.

Most assessments and published literature on early childhood development have been normed on western, educated, industrialised, rich, and democratic populations (also known as WEIRD populations), which presents problems when items do not work in the same way across different contexts. Research with non-WEIRD countries and non-WEIRD subpopulations suggests that culturally valued attributes of a well developed child are not captured by many measures for early childhood development. Psychometric analyses of such items (eg, humility, reciprocity, and independence) show these traits to be important and predictive of later outcomes; however, these attributes are unlikely to be included in a core set of internationally comparable items. However, excluding such items in populations in which these attributes are important is not only disrespectful but hinders science's understanding of how children develop. Therefore, limiting data collection efforts to only internationally comparable items might result in the promotion of poor policy and practice.

Experts in the measurement of early childhood development are challenged by the production of globally comparable data that reflect local culture and context and provide information relevant to local policy and practice.[2] With the probable adoption of the Multiple Indicator Cluster Survey-Early Childhood Development Index (known as MICS-ECDI)[3] as the measure for countries to report data relevant to Target 4.2.1 of the Sustainable Development Goals, policy makers must understand what the resultant data will and will not be able to inform. Limiting the monitoring of early childhood development to a brief, culturally neutral set of items applied within nationally representative surveys will risk poor policy making and hinder the veracity of our science.

Examples of population-level measures for early childhood development that have been specifically created to ensure sensitivity in the contexts of diverse, low-income and middle-income countries include the early Human Capability Index,[4] Global Scale for Early Development,[5] East Asia-Pacific Early Child Development Scales,[6] Regional Project on Child Development Indicators,[7] and the International Development and Early Learning Assessment.[8] The promotion of globally comparable tools does not need to come at the expense of locally relevant instruments. To garner the advantages of both, we suggest the embedment of a culturally neutral set of core items (such as the MICS-ECDI) within an approach that is culturally responsive (such as the aforementioned population-level measures). This approach would advance our understanding of the science behind the strengths and weaknesses of all cultures in the support of early childhood development.

We declare no competing interests.

Editorial note: the *Lancet* Group takes a neutral position with respect to territorial claims in published maps and institutional affiliations.

References

1. Rao N • Sun J • Richards B • et al.
Assessing diversity in early childhood development in the East Asia-Pacific.
Child Indic Res. 2019; **12**: 235-254

View in Article   Scopus (25) • Crossref • Google Scholar

2. Gove A • Black MM
Measurement of early childhood development and learning under the sustainable development goals.
J Human Dev Capabil. 2016; **17**: 599-605

View in Article   Scopus (11) • Crossref • Google Scholar

3. Loizillon A • Petrowski N • Britto P • Cappa C
Development of the early childhood development index in MICS Surveys.
https://mics.unicef.org/files?job=W1siZilsljlwMTcvMDkvMTUvMjEvMTUvNDMvMzc4L01JQ1NfTWV0aG9kb2xvZ2ljYWxfU GFwZXJfNi5wZGYiXV0&sha=85c096f0b2c5b0c8
Date: 2017
Date accessed: January 28, 2020

View in Article   Google Scholar

4. Sincovich A • Gregory T • Zanon C • Santos DD • Lynch J • Brinkman SA
Measuring early childhood development in multiple contexts: the internal factor structure and reliability of the early Human Capability Index in seven low and middle income countries.
BMC Pediatr. 2019; **19**: 471

View in Article   Scopus (6) • PubMed • Crossref • Google Scholar

5. Richter L • Black M • Britto P • et al.
Early childhood development: an imperative for action and measurement at scale.
BMJ Glob Health. 2019; 4e001302

View in Article   PubMed • Crossref • Google Scholar

6. Rao N • Sun J • Ng M • et al.
Report on technical support for the validation, finalization and adoption of the East Asia-Pacific Early Child Development Scales (EAP-ECDS).
https://arnec.net/publication/full-report:-east-asia-pacific-early-child-development-scales-(eap-ecds)
Date: 2014
Date accessed: January 18, 2020

View in Article   Google Scholar

7. Verdisco A • Ceuto S • Thompson J • Neuschmidt O
PRIDI. Regional project of child development indicators. Urgency and possibility results: first initiative to comparative data on child development in Latin America.
Inter-American Development Bank, Washington, DC2015

View in Article   Google Scholar

8. Pisani L • Borisova I • Dowd AJ
International development and early learning assessment technical working paper.
https://idela-network.org/wp-content/uploads/2017/06/IDELA-technical-working-paper_Q4-2015.pdf
Date: August, 2015
Date accessed: January 29, 2020

View in Article   Google Scholar

Article Info

Publication History

Published: February 11, 2020

Identification

DOI: https://doi.org/10.1016/S2352-4642(20)30026-2

Copyright

ScienceDirect

Access this article on ScienceDirect

References

1. Irwin LG, Siddiqi A, Hertzman C. *Early Child Development: A Powerful Equalizer. Final Report for the World Health Organization's Commission of the Social Determinants of Health.* Vancouver: Human Early Learning Partnership; 2007.
2. Hertzman C, Boyce T. How Experience Gets Under the Skin to Create Gradients in Developmental Health. *Annual Review of Public Health.* 2010;31:329-347.
3. Lynch JW, Davey-Smith G. A Life Course Approach to Chronic Disease Epidemiology. *Annual Review of Public Health.* 2005;26(1):1-35.
4. Heckman J. Skill formation and the economics of investing in disadvantaged children. *Science.* 2006;312(5782):1900-1902.
5. Britto PR, Lye SJ, Proulx K, et al. Nurturing care: promoting early childhood development. *The Lancet.* 2017;389:91-102.
6. Bradley RH, Corwyn RF. Socioeconomic Status and Child Development. *Annual Review of Psychology.* 2002;53(1):371-399.
7. Hertzman C, Wiens M. Child development and long term outcomes: A population health perspective and summary of successful interventions. *Social Science and Medicine.* 1996;43(7):1083-1093.
8. Feinstein L, Duckwork K. *Development in the early years: Its importance for school performance and adult outcomes.* London: Institute of Education; 2006.
9. Black MM, Walker SP, Fernald LCH, et al. Early childhood development coming of age: science through the life course. *The Lancet.* 2017;389(10064):77-90.
10. Black RE, Liu L, Hartwig FP, et al. Health and development from preconception to 20 years of age and human capital. *The Lancet.* 2022;399(10336):1730-1740.
11. Victora CG, Adair L, Fall C, et al. Maternal and child undernutrition: consequences for adult health and human capital. *The Lancet.* 2008;371(9609):340-357.

12. Victora CG, Hartwig FP, Vidaletti LP, et al. Effects of early-life poverty on health and human capital in children and adolescents: analyses of national surveys and birth cohort studies in LMICs. *The Lancet.* 2022;399(10336):1741-1752.

13. Vaivada T, Lassi ZS, Irfan O, et al. What can work and how? An overview of evidence-based interventions and delivery strategies to support health and human development from before conception to 20 years. *The Lancet.* 2022;399(10337):1810-1829.

14. García JL, Heckman JJ, Leaf DE, Prados MJ. The Life-cycle Benefits of an Influential Early Childhood Program. *National Bureau of Economic Research Working Paper Series.* 2016;No. 22993.

15. Britto PR. *Early moments matter for every child.* New York: UNICEF; 2017.

16. Gertler P, Heckman J, Pinto R, et al. Labor market returns to an early childhood stimulation intervention in Jamaica. *Science.* 2014;344(6187):998-1001.

17. Elango S, Garcia JL, Heckman J, Hojman A. *Early Childhood Education. IZA Discussion Paper No. 9476*: Institute for Study of Labor; 2015.

18. Engle PL, Fernald LCH, Alderman H, et al. Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. *The Lancet.* 2011;378(9799):1339-1353.

19. Lynch JW, Law C, Brinkman S, Chittleborough C, Sawyer M. Inequalities in child healthy development: Some challenges for effective implementation. *Social Science and Medicine.* 2010;71(7):1219-1374.

20. Goldfeld S, Yousafzai A. Monitoring tools for child development: an opportunity for action. *The Lancet Global Health.* 2018;6(3):e232-e233.

21. Brinkman S, Stanley F. Child Indicators in Public Health. In: Ben-Arieh A, Frones I, Casas F, Korbin J, eds. *Handbook of Child Wellbeing*: New York: Springer Science + Business Media Dordrecht; 2012.

22. Gapminder. Child Mortality Rate, under age five. Version 11. Retrieved from: https://www.gapminder.org/data/documentation/gd005/. 2020.

23. United Nations. *The Millennium Development Goals Report 2010.* New York: United Nations; 2010.

24. You D, Hug L, Ejdemyr S, et al. Global, Regional, and National Levels and Trends in Under-5 Mortality between 1990 and 2015, with Scenario-Based Projections to 2030: A systematic analysis by the UN Inter-agency Group for Child Mortality Estimation. *The Lancet.* 2015;386(10010):2275-2286.

25. Sharrow D, Hug L, You D, et al. Global, regional, and national trends in under-5 mortality between 1990 and 2019 with scenario-based projections until 2030: a systematic analysis by the UN Inter-agency Group for Child Mortality Estimation. *The Lancet Global Health.* 2022;10(2):e195 - e206.

26. World Health Organization, United Nations Children's Fund, World Bank Group. *The Nurturing Care Framework for Early Childhood Development: A framework for helping children survive and thrive to transform health and human potential.* Geneva: World Health Organization; 2018.

27. United Nations General Assembly. *Report of the Open Working Group of the General Assembly on Sustainable Development Goals, UNGA a/68/970* 2014.

28. World Bank. *Atlas of Sustainable Development Goals 2018: World Development Indicators.* Washington, DC: World Bank; 2018.

29. Fernald LCH, Prado E, Kariger P, Raikes A. *A Toolkit for Measuring Early Childhood Development in Low- and Middle-Income Countries.* Washington DC: World Bank Group; 2017.

30. Bjorklund DF, Causey KB. *Children's Thinking: Cognitive Development and Individual Differences.* Vol Sixth Edition. Thousand Oaks, CA: SAGE Publications; 2017.

31. Woodward AL, Markman EM. Early Word Learning. *Handbook of Child Psychology.* Vol 2: Cognition, Perception and Language. Fifth ed. New York: John Wiley & Sons; 1998:371–420.

32. Harkness S, Super CM. Why African Children Are So Hard to Test. *Annals of the New York Academy of Sciences.* 1977;285(1):326–331.

33. Zubler JM, Wiggins LD, Macias MM, et al. Evidence-Informed Milestones for Developmental Surveillance Tools. *Pediatrics.* 2022;149(3):e2021052138.

34. National Scientific Council on the Developing Child. *Young children develop in an environment of relationships: Working paper No. 1.* Center on the Developing Child, Harvard University; 2004.

35. National Research Council & Institute of Medicine. *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press; 2000.

36. Walker D, Greenwood C, Hart B, Carta J. Prediction of School Outcomes Based on Early Language Production and Socioeconomic Factors. *Child Development.* 1994;65(2):606-621.

37. Glascoe FP. Screening for Developmental and Behavioral Problems. *Mental Retardation and Developmental Disabilities Research Reviews.* 2005;11(3):173-179.

38. Squires J, Potter L, Bricker D, eds. *The ASQ user's guide for the ages & stages questionnaires: a parent-completed, child-monitoring system*. Baltimore, MD: Paul H Brookes Publishing Co; 1995.

39. Squires J, Bricker D. *Ages & Stages Questionnaires. Third Edition (ASQ-3)*. Baltimore, MD: Brookes Publishing; 2009.

40. Richter J, HJanson H. A validation study of the Norwegian version of the Ages and Stages Questionnaires. *Acta Paediatrica.* 2007;96(5):748–752.

41. Velikonja T, Edbrooke-Childs J, Calderon A, Sleed M, Brown A, Deighton J. The psychometric properties of the Ages & Stages Questionnaires for ages 2-2.5: a systematic review. *Child: Care, Health and Development.* 2016;43:1-17.

42. Small JW, HIx-Small H, Vargas-Baron E, Marks KP. Comparative use of the Ages and Stages Questionnaires in low- and middle-income countries. *Developmental Medicine & Child Neurology.* 2018(61):431-443.

43. Squires J, Potter L, Bricker DD, Lamorey S. Parent-Completed Developmental Questionnaires: Effectiveness with Low and Middle Income Parents. *Early Childhood Research Quarterly.* 1998;13(2):345-354.

44. Marks KP, Sjo NM, Wilson P. Comparative use of the Ages and Stages Questionnaires in the USA and Scandinavia: a systematic review. *Developmental Medicine & Child Neurology.* 2018;61:419-430.

45. Bayley N. *Manual for the Bayley scales of infant development*. New York: Psychological Corporation; 1969.

46. Bayley N. *Bayley Scales of Infant and Toddler Development: Bayley-III.* Vol 7. San Antonio, TX: Harcourt Assessment; 2006.

47. Del Rosario C, Slevin M, Molloy EJ, Quigley J, Nixon E. How to use the Bayley Scales of Infant and Toddler Development. *Archives of Disease in Childhood - Education and Practice* 2021;106:108-112.

48. Ranjitkar S, Kvestad I, Strand TA, et al. Acceptability and Reliability of the Bayley Scales of Infant and Toddler Development-III Among Children in Bhaktapur, Nepal. *Frontiers in Psychology.* 2018-July-24 2018;9.

49. Ballot DE, Ramdin T, Rakotsoane D, et al. Use of the Bayley Scales of Infant and Toddler Development, Third Edition, to Assess Developmental Outcome in Infants and Young Children in an Urban Setting in South Africa. *International Scholarly Research Notices.* 2017:1631760.

50. Azari N, Soleimani F, Vameghi R, et al. A Psychometric Study of the Bayley Scales of Infant and Toddler Development in Persian Language Children. *Iranian journal of child neurology.* 2017;11(1):50-56.

51. Young ME. *Early Child Development: From Measurement to Action.* Washington DC., USA: The World Bank; 2007.

52. Janus M, Brinkman S, Duku E, Hertzman C, Santos R, Sayers M. *The Early Development Instrument: A population-based measure for communities. A handbook on development, properties and use.* Hamilton, Ontario, Canada: Offord Centre for Child Studies; 2007.

53. Janus M, Offord D. Development and psychometric properties of the Early Development Instrument (EDI): A measure of children's school readiness. *Canadian Journal of Behavioural Science.* 2007;39:1-22.

54. Janus M, Brinkman S, Duku E. Validity and psychometric properties of the Early Development Instrument in Canada, Australia, United States and Jamaica. *Social Indicators Research.* 2011;103(2):283-297.

55. Ip P, Li SL, Rao N, Ng SSN, Lau WWS, Chow CB. Validation study of the Chinese Early Development Instrument (CEDI). *BMC Pediatrics.* 2013;13(1):146.

56. Duku E, Janus M, Brinkman S. Investigation of the cross-national equivalence of a measurement of early child development. *Child Indicators Research.* 2014;8(2):471-489.

57. Brinkman SA, Kinnell A, Maika A, Hasan A, Jung H, Pradhan M. Validity and Reliability of the Early Development Instrument in Indonesia. *Child Indicators Research.* 2017;10:331-352.

58. Raikes HA. *Education for people and planet: Creating sustainable futures for all. Measuring of child development and learning. Background paper prepared for the 2016 Global Education Monitoring Report* 2016.

59. Mustard JF. Experience-based Brain Development: Scientific Underpinnings of the Importance of Early Child Development in a Global World. In: Young ME, Richardson LM, eds. *Early Child Development: From Measurement to Action*. Washington DC, USA: The World Bank; 2007:43-84.

60. Brinkman S, Gregory T, Goldfeld S, Lynch J, Hardy M. Data Resource Profile: The Australian Early Development Index (AEDI). *International Journal of Epidemiology.* 2014;43(4):1089-1096.

61. United Nations General Assembly. Transforming Our World: The 2030 Agenda for Sustainable Development. A/RES/70/1. Retrieved from: www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E.

62. Hereward M. Setting the record straight on measuring SDG 4.2.1. *The Lancet Global Health.* 2021;9(7):e910-e911.

63. Olusanya BO, Hadders-Algra M, Breinbauer C, Williams AN, Newton CRJ, Davis AC. The conundrum of a global tool for early childhood development to monitor SDG indicator 4.2.1. *The Lancet Global Health.* 2021;9(3):E586-E587.

64. GSED Team. The global scale for early development (GSED). *Early childhood matters.* 2019;14:80-84.

65. Altafim ER, McCoy DC, Brentani A, Escobar AM, Grisi SJ, Fink G. Measuring early childhood development in Brazil:validation of the Caregiver Reported EarlyDevelopment Instruments (CREDI). *Jornal de Pediatria.* 2018;96(1):66-75.

66. Pushparatnam A, Luna Bazaldua DA, Holla A, Azevedo JP, Clarke M, Devercelli A. Measuring Early Childhood Development Among 4–6 Year Olds: The Identification of Psychometrically Robust Items Across Diverse Contexts. *Frontiers in Public Health.* 2021;9.

67. Janus M. The Early Development Instrument, a tool for monitoring childrens development and readiness for school. In: Young ME, Richardson LM, eds. *Early Child Development - from measurement to action: A priority for growth and equity*. Washington DC., USA: World Bank; 2007:141-155.

68. Offord Centre McMaster University. Early Development Instrument (EDI). Retrieved from: https://edi.offordcentre.com/.

69. United Nations. SDG indicator metada. Retrieved from: https://unstats.un.org/sdgs/metadata/files/Metadata-04-02-01.pdf.

70. UNICEF. *Development of the Early Childhood Development Index in MICS surveys, MICS Methodological papers. Paper No. 6.*: Data and Analytic Section, Division of Data, Research and Policy, UNICEF; 2017.

71. McCoy DC, Peet ED, Ezzati M, et al. Early childhood developmental status in low- and middle-income countries: National, regional, and global prevalence estimates using predictive modeling. *PLOS Medicine.* 2016;13(6):e1002034.

72. Cappa C, de Castro F, Hereward M, Petrowski N, Werntz KM. Beyond block towers: Designing a global measure for early childhood development. UNICEF connect; 2018.

73. Cappa C, Petrowski N, De Castro EF, et al. Identifying and Minimizing Errors in the Measurement of Early Childhood Development: Lessons Learned from the Cognitive Testing of the ECDI2030. *International Journal of Environmental Research and Public Health.* 2021;18(22):12181.

74. Cappa C. *Monitoring SDG 4.2.1: A Briefing from the Inter-Agency and Expert Group on ECD Measurement*: UNICEF; 2019.

75. Petrowski N, Cappa C, Dia E. Two years on: Progress and country uptake of the ECDI2030. *UNICEF Connect.* Vol 2022: UNICEF; 2022.

76. United Nations Children's Fund. *ECDI2030: Instructions for Interviewers.* New York: 2020; 2020.

77. Rao N, Sun J, Ng M, et al. *Validation, Finalization and Adoption of the East Asia-Pacific Early Child Development Scales (EAP-ECDS)*: UNICEF; 2014.

78. Rao N, Sun J, Becher Y. Assessing early development and learning across cultures: The East Asia Pacific-Early Child Development Scales. *Assessment & Development Matters.* 2015;7(2):21-25.

79. Rao N. *Assessing Early Development and Learning: The East Asia-Pacific Early Childhood Development Scales (EAP-ECDS).* Hong Kong: University of Hong Kong; 2017.

80. Rao N, Sun J, Bacon-Shone J, Ip P, Becher Y. *East Asia-Pacific Early Child Development Scales-Short Form: Final Report.* Singapore: Asia Pacific Regional Network for Early Childhood; 2016.

81. Pisani L, Borisova I, Dowd AJ. *International Development and Early Learning Assessment (IDELA). Save the Children Technical Working Paper* 2015.

82. Pisani L, Borisova I, Dowd AJ. Developing and validating the International Development and Early Learning Assessment (IDELA). *International Journal of Educational Research.* 2018;91:1-15.

83. UNESCO, UNICEF, Brookings Institution, World Bank. *MELQO Overview: Measuring Early Learning and Quality Outcomes* 2017.

84. Raikes A, Sayre R, Davis D, et al. The Measuring Early Learning Quality & Outcomes initiative: purpose, process and results. *Early Years.* 2019;39(4):360-375.

85. ECD Measure. *MELQO Planning and Implementation Toolkit. Available: http://ecdmeasure.org/wp-content/uploads/2018/04/MELQO-Implementation-Toolkit_March-2018-1.pdf* 2018.

86. Raikes A, Koziol N, Davis D, Burton A. Measuring quality of preprimary education in sub-Saharan Africa: Evaluation of the Measuring Early Learning Environments scale. *Early Childhood Research QuarterlyMeasuring.* 2020;53(4):571–585.

87. Raikes A, Koziol N, Janus M, et al. Examination of school readiness constructs in Tanzania: Psychometric evaluation of the MELQO scales. *Journal of Applied Developmental Psychology.* 2019;62:122-134.

88. Pushparatnam A, Luna Bazaldua DA. Measuring early childhood outcomes comparably across countries without sacrificing local relevance: can it be done? *World Bank Blogs*; 2021.

89. Bennetts SK, Mensah FK, Westrupp EM, Hackworth NJ, Reilly S. The Agreement between Parent-Reported and Directly Measured Child Language and Parenting Behaviors. *Frontiers in Psychology.* 2016;7(1710):1-18.

90. Snow CE, Van Hemel SB. Early Childhood Assessment: Why, What, and How. Washington, DC: National Academies Press; 2008.

91. Ertem IO, Krishnamurthy V, Mulaudzi MC, et al. Similarities and differences in child development from birth to age 3 years by sex and across four countries: a cross-sectional, observational study. *Lancet Glob Health.* 2017;6:e279-e291.

92. Mustard JF. *Investing in the Early Years: Closing the gap between what we know and what we do.* Adelaide: Department of Premier and Cabinet, Government of South Australia; 2008.

93. Rao N, Mirpuri S, Sincovich A, Brinkman SA. Overcoming conceptual and methodological challenges in measuring early childhood development across cultures. *Lancet Child and Adolescent Health.* 2020;4(5):352-354.

94. Hambleton RK, Merenda PF, Speilberger CD. *Adapting educational and psychological tests for cross cultural assessment.* Mahwah, NJ: Lawrence Erlbaum Associates Inc.; 2005.

95. McCoy DC. Building a model of cultural universality with specificity for global early childhood development. *Child Development Perspectives.* 2022;16(1):27-33.

96. Keller H, Kärtner J. *The cultural solution of universal developmental tasks.* Vol 3. New York, NY: Oxford University Press; 2013.

97. Gove A, Black MM. Measurement of Early Childhood Development and Learning under the Sustainable Development Goals. *Journal of Human Development and Capabilities.* 2016;17(4):599-605.

98. Downing SM, Haladyna TM. *Handbook of Test Development*. Mahwah, NJ: Lawrence Erlbaum Associates; 2006.

99. Brinkman S, Thanh Vu B. *Early Child Development in Tonga*: World Bank. Washington, DC. Retrieved from: http://documents.worldbank.org/curated/en/898671481187892018/pdf/110821-PUB-PUBLIC-pubdate-12-7-2016.pdf; 2017.

100. MacDonald K, Brinkman S, Jarvie W, et al. *Pedagogy versus School Readiness: The Impact of a Randomized Reading Instruction Intervention and Community-Based Playgroup Intervention on Early Grade Reading Outcomes in Tonga*: WPS794: World Bank; 2017.

101. Brinkman SA. Early Human Capability Index (eHCI). *Telethon Kids Institute*. Retrieved from: https://ehci.telethonkids.org.au/.

102. Brinkman S, Sincovich A, Danchev P. *The Status of Early Childhood Health and Development in Northern Lao PDR: Baseline Results from the Early Childhood Education Project.* Washington, D.C: World Bank Group. Retrieved from: http://documents.worldbank.org/curated/en/228221488446175113/The-status-of-early-childhood-health-and-development-in-northern-Lao-PDR 2016.

103. Brinkman S, Sincovich A, Than Vu B. *Early Childhood Development in Samoa: Baseline results from the Samoan Early Human Capability Index.* Washington, D.C; World Bank Group. Retrieved from: http://documents.worldbank.org/curated/en/363761563248716398/pdf/Main-Report.pdf 2017.

104. Brinkman S, Sincovich A, Thanh Vu B. *How are Tuvalu's children developing? Evidence-based policy recommendations for better early childhood development in Tuvalu.* Washington, D.C, World Bank Group. Retrieved from: http://documents.worldbank.org/curated/en/373411563250298102/How-are-Tuvalu-s-children-developing-Evidence-based-policy-recommendations-for-better-early-childhood-development-of-Tuvaluan-children 2017.

105. Zhao J, Brinkman SA, Zhang Y, et al. Measuring early childhood development with The Early Human Capability Index (eHCI): a reliability and validity study in China. *BMC Pediatrics.* 2020;20(323):2-10.

106. Brinkman S, Sincovich A, Thanh Vu B. *The Status of Early Childhood Health and Development in Kiribati: Results from a population wide census.* Washington, D.C; World Bank Group. Retrieved from: http://documents.worldbank.org/curated/en/853791563251577440/pdf/The-Status-of-Early-Childhood-Health-and-Development-in-Kiribati-Results-from-a-Population-Wide-Census.pdf 2017.

107. Santos DD, Zanon C. *The Brazilian experience with the eHCI: the adaptation process, reduction proceedings, and validation study*: Laboratório de Estudos e Pesquisas em Economia Social; 2017.

108. Thanh Vu B, Beemer S, Brinkman S, et al. *Operational Guide. Implementing and Evaluating Interventions to Improve School Readiness and Early Literacy: Experience from the Pacific Early Age Readiness and Learning Program.* Washington DC: World Bank Group; 2019.

109. Andrich D, Marais I. *A Course in Rasch Measurement Theory. Measuring in the Educational, Social and Health Sciences*. Melbourne: Springer; 2019.

110. Traub RE, Rowley GL. Understanding Reliability. *Educational Measurement: Issues and Practice.* 1991;10(1):37-45.

111. Frisbie DA. Reliability of Scores From Teacher-Made Tests. *Educational Measurement: Issues and Practice.* 1988;7(1):25-35.

112. Guttman L. A basis for analyzing test-retest reliability. *Psychometrika.* 1945;10(4):255-282.

113. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. *BMJ.* 1997;314(7080):572.

114. Downing SM. Validity: on the meaningful interpretation of assessment data. *Medical Education.* 2003;37(9):830-837.

115. Messick S. Meaning and Values in Test Validation: The Science and Ethics of Assessment. *Educational Researcher.* 1989;18(2):5-11.

116. Mosier CI. A critical examination of the concepts of face validity. *Educational and Psychological Measurement.* 1947;7(2):191-205.

117. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychological Bulletin.* 1955;52(4):281-302.

118. Nunnally JC, Bernstein IH. *Psychometric theory (3rd ed.)*. 3rd edition ed. New York: McGraw-Hill; 1994.

119. Whitley BE. *Principles of Research in Behavioral Science (4th ed.)*. Mountain View, CA: Mayfield Publishing; 1996.

120. Cohen RJ, Swerdilk ME. *Psychological testing and assessment: an introduction to tests and measurement.* 6th ed. New York: McGraw-Hill; 2005.

121. World Bank. World Bank Open Data. Retrieved from https://data.worldbank.org/.

122. UNESCO. *Policy Review Report: Early Childhood Care and Education in Brazil.* Paris, France: The Division for the Promotion of Basic Education. UNESCO Education Sector; 2006.

123. Telethon Kids Institute. School-based surveys: Helping to stop Brazilian children from falling behind. Retrieved from: https://ehci.telethonkids.org.au/resources/brazil-case-study/. 2022.

124. Chen S, Zhao C, Chen C, Wu Z, Snow CE, Lu M. Does One More Year Matter? Dosage Effect of the One-Village-One-Preschool Intervention in Rural China. *Journal of Research on Educational Effectiveness.* 2022(2):217-242.

125. Telethon Kids Institute. Investing in childhood development with confidence thanks to a free and simple tool. Retrieved from: https://ehci.telethonkids.org.au/resources/china-case-study/. 2022.

126. Warr P, Rasphone S, Menon J. *Two Decades of Rising Inequality and Declining Poverty in the Lao People's Democratic Republic.* Manila, Philippines: Asian Development Bank; 2015.

127. World Bank Group. *Lao People's Democratic Republic - Early Childhood Education Project (English).* Washington, D.C.: World Bank Group; 2014.

128. Brinkman S, Sincovich A, Lam B, Beteille T. *The Early Childhood Education Project in Lao PDR: Impact Evaluation Report.* Washington D.C.: World Bank Group; 2022.

129. United Nations Children's Fund. *Situation Analysis of Children in Kiribati.* Suva, Fiji: UNICEF; 2017.

130. Schuster S. Early Childhood Education in Samoa: Preservice Curriculum and Concerns Regarding Play. In: Antala B, Demirhan G, Carraro A, Oktar C, Oz H, Kaplanova A, eds. *Physical Education in Early Childhood Education and Care. Researches – Best Practices - Situation*. Bratislava: Slovak Scientific Society for Physical Education and Sport and FIEP; 2019:409-420.

131. World Bank. Migration and Remittances Data. Retrieved from: https://www.worldbank.org/en/topic/migrationremittancesdiasporaissues/brief/migration-remittances-data.

132. UNESCO. Tonga Non-State Actors in Education. Retrieved from: https://education-profiles.org/oceania/tonga/~non-state-actors-in-education#Early%20childhood%20care%20and%20education.

133. United Nations. Least Developed Country Category: Tuvalu. *United Nations Department of Economic and Social Affairs*. Retrieved from: https://www.un.org/development/desa/dpad/least-developed-country-category-tuvalu.html.

134. WHO Multicentre Growth Reference Study Group. *WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development.* Geneva: World Health Organization; 2006.

135. UNICEF. *MICS6 TOOLS.* Multiple Indicator Cluster Surveys (MICS). Retrieved from http://mics.unicef.org/tools 2019.

136. Dubeck MM, Gove A, Alexander K. *School-based assessments: What and how to assess reading. Understanding What Works in Oral Reading Assessments*: UNESCO, Institute for Statistics (UIS), Montreal, pp. 41-57; 2016.

137. Visu-Petra L, Stanciu O, Benga O, Miclea M, Cheie L. Longitudinal and concurrent links between memory span, anxiety symptoms, and subsequent executive functioning in young children. *Frontiers in Psychology.* 2014;5(443).

138. McClelland MM, Cameron CE, Duncan R, et al. Predictors of early growth in academic achievement: the head-toes-knees-shoulders task. *Frontiers in psychology.* 2014;5:599.

139. Su Y, Rao N, Sun J, Zhang L. Preschool quality and child development in China. *Early Childhood Research Quarterly.* \ 2021;56:15-26.

140. Brinkman S, Sincovich A, Beteille T, Danchev P. *Lao PDR Early Childhood Education Project Midline Results. Snapshot Four: Instruments to Support Child Development.* Xieng Ngeun Village: World Bank Group; in press.

141. *Mplus* [computer program]. Version 7.31. Los Angeles, CA: Muthen & Muthen; 2012.

142. Flora DB, LaBrish C, Chalmers RP. Old and new ideas for data screening and assumption testing for exploratory and confirmatory factor analysis. *Frontiers in Psychology.* 2012;3(55):1-21.

143. Holgado–Tello FP, Chacón–Moscoso S, Barbero–García I, Vila–Abad E. Polychoric versus pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity*. 2010;44(1):153-166.

144. Brown TA. *Confirmatory factor analysis for applied research*. New York: Guildford Press; 2006.

145. Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal.* 1999;6(1):1-55.

146. Yu C. *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. PhD Thesis.* University of California; 2002.

147. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research.* 2006;99(6):323-338.

148. Costello AB, Osborne JW. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation.* 2005;10(7):1-9.

149. Pett MA, Lackey NR, Sullivan JJ. *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Thousand Oaks, California: Sage Publications; 2003.

150. Gadermann AM, Guhn M, Zumbo BD. Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation.* 2012;17(3):1-13.

151. *psych: Procedures for Psychological, Psychometric, and Personality Research.* [computer program]. R package version 1.7.8. Evanston, Illinois, USA: Northwestern University; 2017.

152. *IBM SPSS Statistics for Windows.* [computer program]. Version 25.0. Armonk, NY: IBM Corp; 2017.

153. *R: A Language and Environment for Statistical Computing* [computer program]. Version 3.6.0. Vienna, Austria: R Foundation for Statistical Computing; 2014.

154. Hauke J, Kossowski T. Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones Geographicae.* 2011;30(2):88-93.

155. Alderman H, Fernald L. The Nexus Between Nutrition and Early Childhood Development. *Annual Review of Nutrition.* 2017;37(1):447-476.

156. Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. *British Medical Journal (Clinical research ed.).* 1986;292(6522):746-750.

157. Brinkman S, Gregory T, Harris J, Hart B, Blackmore S, Janus M. Associations between the Early Development Instrument at Age 5 and Reading and Numeracy Skills at Ages 8, 10 and 12: A Prospective Linked Data Study. *Child Indicators Research.* 2013;6(4):695-708.

158. Pisani L, Seiden J, Wolf S. Longitudinal evidence on the predictive validity of the International Development and Early Learning Assessment (IDELA). *Educational Assessment, Evaluation and Accountability.* 2022;34:173-194.

159. Altman DG, Bland JM. Statistics Notes: Diagnostic tests 3: receiver operating characteristic plots. *BMJ.* July 16, 1994 1994;309(6948):188.

160. Akobeng AK. Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta Pædiatrica.* 2007;96:644-647.

161. Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med.* Jul 2003;29(7):1043-1051.

162. *IBM SPSS Statistics for Windows* [computer program]. Version 27.0. Armonk, NY; 2020.

163. Britto PR, Yoshikawa H, Boller K. Quality of early childhood development programs in global contexts: Rationale for investment, conceptual framework and implications for equity. *Social Policy Report.* 2011;25(2).

164. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology.* 1999;10:37-48.

165. Vyas S, Kumaranayake L. Constructing socio-economic status indeces: how to use principal component analysis. *Health Policy Plan.* 2006;21(6):459-468.

166. *Stata Statistical Software: Release 17* [computer program]. Version. TX: StataCorp LLC: College Station; 2019.

167. Hughes RA, Heron J, Sterne JAC, Tilling K. Accounting for missing data in statistical analyses: multiple imputation is not always the answer. *International Journal of Epidemiology.* 2019;48(4):1294-1304.

168. Little RJA, Rubin DB. *Statistical Analysis with Missing Data. 2nd edition.* NJ: Wiley; 2002.

169. Bartlett JW, Harel O, Carpenter JR. Asymptotically Unbiased Estimation of Exposure Odds Ratios in Complete Records Logistic Regression. *American Journal of Epidemiology.* 2015;182(8):730-736.

170. White I, Carlin J. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. *Statistics in Medicine.* 2010;29(28):2920-2931.

171. Sincovich A, Gregory T, Zanon C, Santos DD, Lynch J, Brinkman S. Measuring early childhood development in multiple contexts: The factor structure and reliability of the

early Human Capability Index in seven low and middle income countries. *BMC Pediatrics.* 2019;19(471):1-14.

172. Raikes A, Yoshikawa H, Britto PR, Iruka I. Children, Youth and Developmental Science in the 2015–2030 Global Sustainable Development Goals. *Society for Research in Child Development Social Policy Report.* 2017;30(3).

173. Rao N, Sun J, Richards B, et al. Assessing Diversity in Early Childhood Development in the East Asia-Pacific. *Child Indicators Research.* 2018;12:235-254.

174. International Test Commission. *The ITC Guidelines for Translating and Adapting Tests (Second Edition)* 2017.

175. Byrne BM. *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York: Taylor and Francis Group; 2012.

176. Yapa HM, Bärnighausen T. Implementation science in resource-poor countries and communities. *Implementation Science.* 2018;13(154):1-13.

177. Duncan G, Claessens A, Huston A, et al. School readiness and later achievement. *Developmental Psychology.* 2007;43(6):1428-1446.

178. Mesquita B. Emotions in Collectivist and Individualist Contexts. *Journal of Personality and Social Psychology.* 2001;80(1):68-74.

179. He J, van de Vijver F. Bias and Equivalence in Cross-Cultural Research. *Online Readings in Psychology and Culture.* 2012;2(2):1-19.

180. Harzing AW. Response styles in cross-national survey research: A 26-country study. *International Journal of Cross Cultural Management.* 2006;6:243-266.

181. Swain SD, Weathers D, Niedrich RW. Assessing Three Sources of Misresponse to Reversed Likert Items. *Journal of Marketing Research.* 2008;45(1):116-131.

182. Netemeyer RG, Bearden WO, Sharma S. *Scaling Procedures: Issues and Applications*. Newbury Park, CA: Sage Publications, Inc; 2003.

183. Sincovich A, Gregory T, Zanon C, Santos DD, Lynch J, Brinkman SA. Measuring early child development in low and middle income countries: Investigating the validity of the early Human Capability Index. *SSM - Population Health.* 2020/08/01/ 2020;11:100613.

184. Sincovich A, Gregory T, Zanon C, Santos DD, Lynch J, Brinkman S. Measuring early childhood development in multiple contexts: The factor structure and reliability of the early Human Capability Index in seven low and middle income countries. *BMC Pediatrics.* 2019;19(471).

185. Bennetts SK, Mensah FK, Westrupp EM, Hackworth NJ, Reilly S. The Agreement between Parent-Reported and Directly Measured Child Language and Parenting Behaviors. *Frontiers in Psychology.* 2016;7(1710).

186. McCoy DC, Peet ED, Ezzati M, et al. Early childhood developmental status in low- and middle-income countries: National, regional, and global prevalence estimates using predictive modeling. *PLOS Medicine.* 2016;13(6).

187. Janus M, Harrison LJ, Goldfeld S, Guhn M, Brinkman S. International research utilizing the Early Development Instrument (EDI) as a measure of early child development: Introduction to the Special Issue. *Early Childhood Research Quarterly.* 2016;35:1-5.

188. Rao N, Sun J, Richards B, et al. Assessing Diversity in Early Childhood Development in the East Asia-Pacific. *Child Indicators Research.* February 10 2018.

189. McCoy DC, Waldman M, Team CF, Fink G. Measuring early childhood development at a global scale: Evidencefrom the Caregiver-Reported Early Development Instruments. *Early Childhood Research Quarterly.* 2018;45:58-68.

190. Rao N, Richards B, Sun J, Weber A, Sincovich A. Early childhood education and child development in four countries inEast Asia and the Pacific. *Early Childhood Research Quarterly.* 2018;47(2019):169-181.

191. Wolf S, Halpin P, Yoshikawa H, Dowd AJ, Pisani L, Borisova I. Measuring school readiness globally: Assessing the construct validity and measurement invariance of the International Development and Early Learning Assessment (IDELA) in Ethiopia. *Early Childhood Research Quarterly.* 2017;41:21-36.

192. Richter L, Black M, Britto P, et al. Early childhood development: an imperative for action and measurement at scale. *BMJ Global Health.* 2019;5(i154-i160).

193. Sabanathan S, Wills B, Gladstone M. Child development assessment tools in low-income and middle-income countries: how can we use them more appropriately? *Archives of Disease in Childhood.* 2015;100:482-488.

194. Upadhyay RP, Naik G, Choudhary TS, et al. Cognitive and motor outcomes in children born low birth weight: a systematic review and meta-analysis of studies from South Asia. *BMC Pediatrics.* 2019;19(35).

195. Perkins JM, Kim R, Krishna A, McGovern M, Aguayo VM, Subramanian SV. Understanding the association between stunting and child development in low- and middle-income countries: Next steps for research and intervention. *Social Science & Medicine.* 2017.

196. Rao N, Richards B, Lau C, et al. Associations Among Early Stimulation, Stunting, and Child Development in Four Countries in the East Asia–Pacific. *International Journal of Early Childhood.* 2020/08/01 2020;52(2):175-193.

197. Sirin SR. Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research.* 2005;75(3):417-453.

198. Woessmann L. The economic case for education. *Education Economics.* 2016;24(1):3-32.

199. Yoshikawa H, Kabay S. *The evidence base on early childhood care and education in global contexts.* . Paris: UNESCO; 2014.

200. Bustamante AS, Dearing E, Zachrisson HD, Vandell DL. Adult outcomes of sustained high-quality early child care and education: Do they vary by family income? *Child Development.* 2021;00:1-22.

201. McCoy DC, Yoshikawa H, Ziol-Guest KM, et al. Impacts of Early Childhood Education on Medium-and Long-Term Educational Outcomes. *Educational Researcher.* 2017;46(8):474–487.

202. Sylva K, Melhuish EC, Sammons P, Siraj-Blatchford I, Taggart B. *The Effective Provision of Pre-School Education (EPPE) Project final report : Technical paper 12 - The final report: Effective pre-school education.* London: DfES/Institute of Education, University of London; 2004.

203. Campbell FA, Ramey CT, Pungello E, Sparling J, Miller-Johnson S. Early Childhood Education: Young Adult Outcomes From the Abecedarian Project. *Applied Developmental Science.* 2002;6(1):42-57.

204. UNESCO. *UNESCO Institute of Statistics Global Database. Education. Sustainable Development Goals (full dataset)*: Retrieved from: http://data.uis.unesco.org/; 2020.

205. Rao N, Sun J, Zhou J, Zhang L. Early achievement in rural China: The role of preschool experience. *Early Childhood Research Quarterly.* 2012;27(1):66-76.

206. van Huizen T, Plantenga J. Do children benefit from universal early childhood education and care? A meta-analysis of evidence from natural experiments. *Economics of Education Review.* 2018;66:206-222.

207. Burchinal M. Measuring Early Care and Education Quality. *Child Development Perspectives.* 2018;12(1):3-9.

208. Anderson K, Raikes A, Kosaraj S, Solano A. *National early care and education monitoring systems*: Brookings Institution, UNICEF, The World Bank.; 2017.

209. Perlman M, Falenchuk O, Fletcher B, McMullen E, Beyene J, Shah PS. A Systematic Review and Meta-Analysis of a Measure of Staff/Child Interaction Quality (the Classroom Assessment Scoring System) in Early Childhood Education and Care Settings and Child Outcomes. *PLoS One.* 2016;11(12):e0167660.

210. Mashburn AJ, Pianta RC, Hamre BK, et al. Measures of Classroom Quality in Prekindergarten and Children's Development of Academic, Language, and Social Skills. *Child Development.* 2008;79(3):732-749.

211. Melhuish E, Ereky-Stevens K, Petrogiannis K, et al. *A review of research on the effects of early childhood education and care (ECEC) upon child development*: University of Oxford; 2015.

212. Vandell DL, Belsky J, Burchinal M, Steinberg L, Vandergrift N. Do Effects of Early Child Care Extend to Age 15 Years? Results From the NICHD Study of Early Child Care and Youth Development. *Child Developemnt.* 2010;81(3):737-756.

213. Wolf S, Raza M, Kim S, Aber J, Behrman J, Seidman E. Measuring and predicting process quality in Ghanaian pre-primary classrooms using the Teacher Instructional Practices and Processes System (TIPPS). *Early Childhood Research Quarterly.* 05/17 2018;45.

214. UNICEF. *Addressing the learning crisis: An urgent need to better finance education for the poorest children.* New York, NY: UNICEF Education Section; 2020.

215. Harms T, Clifford R. *The early childhood environment rating scale (ECERS)*. New York, NY: Teachers College Record; 1980.

216. Harms T, Clifford R, Cryer D. Early Childhood Environment Rating Scale. 3rd Edn. 3rd ed. New York, NY: Teachers College Press; 2015.

217. Aboud FE. Evaluation of an early childhood preschool program in rural Bangladesh. *Early Childhood Research Quarterly.* 2006;21(1):46-60.

218. Li K, Ying Hu Y, Pan Y, Qin J, Fan X. Quarterly Chinese Early childhood environment rating scale (trial) (CECERS): A validity study. *Early Childhood Research Quarterly.* 2014;29(3):268–282.

219. Rao N, Pearson E. *An Evaluation of Early Childhood Care and Education Programmes in Cambodia*: UNICEF; 2007.

220. Betancur L, Maldonado-Carreño C, Votruba-Drzal E, Bernal R. Measuring preschool quality in low-and middle-income countries: validity of the ECERS-R in Colombia. *Early Childhood Research Quarterly.* 2021;54(1):86–98.

221. Chen S, Wolf S. Measuring the Quality of Early Childhood Education in Low- and Middle-Income Countries. *Frontiers in Psychology.* 2021;12(774740).

222. Hu BY, Dieker L, Yang Y, Yang N. The quality of classroom experiences in Chinese kindergarten classrooms across settings and learning activities: implications for teacher preparation. *Teaching and Teacher Education.* 2016;57:39–50.

223. Raikes A. Measuring child development and learning. *European Journal of Education.* 2017;52(4):511–522.

224. World Bank. The World Bank In Lao PDR. Retrieved from: https://www.worldbank.org/en/country/lao/overview#1.

225. Lao PDR Ministry of Education. *Education for All (EFA) National Plan of Action 2003-2015.* Bangkok, Thailand: Lao PDR Ministry of Education & UNESCO Bangkok; 2005.

226. World Bank. *Lao People's Democratic Republic Second Global Partnership for Education Project (P149130) - Project Appraisal Document* 2015.

227. Brinkman S, Sincovich A, Beteille T, Danchev P. *Lao PDR Early Childhood Education Project Midline Results. Snapshot Two: Early Childhood Development.* Xieng Ngeun Village: World Bank Group; in press.

228. Tabachnick BG, Fidell LS. *Using multivariate statistics*. 6th ed. Upper Saddle River, N.J.: Pearson Education; 2013.

229. Hatfield BE, Burchinal MR, Pianta RC, Sideris J. Thresholds in thea ssociation between quality of teacher-child interactions and preschoolchildren's school readiness skills. *Early Childhood Research Quarterly.* 2016;36:561–571.

230. Burchinal M, Xue Y, Auger A, et al. Testing for quality thresholds and features in early care and education. *Monographs of the Society for Research in Child Development.* 2016;81:46-63.

231. Molina E, Fatima SF, Ho AD, Melo C, Wilichowski TM, Pushparatnam A. Measuring the quality of teaching practices in primary schools: Assessing the validity of the Teach observation tool in Punjab, Pakistan. *Teaching and Teacher Education.* 2020;96:103171.

232. The World Bank. Teach: Helping Countries Track and Improve Teaching Quality. Retrieved from: https://www.worldbank.org/en/topic/education/brief/teach-helping-countries-track-and-improve-teaching-quality#ref1.

233. Pushparatnam A, Ding E, Lee K, Rolla A, Wilinkski B. *Effective Teaching Practices in Early Childhood Education (ECE): The Evidence Base for the Teach ECE Classroom Observation Tool.* Washington, DC: World Bank; 2021.

234. Fuligni AS, Howes C, Huang Y, Hong SS, Lara-Cinisomo S. Activity Settings and Daily Routines in Preschool Classrooms: Diverse Experiences in Early Learning Settings for Low-Income Children. *Early Childhood Research Quarterly.* 2012;1(27):198-209.

235. McCoy DC, Wolf S. Changes in Classroom Quality Predict Ghanaian Preschoolers' Gains in Academic and Social-Emotional Skills. *Developmental psychology.* 2018;54(8):1582-1599.

236. Maier MF, M.P. M, Xia S, et al. Content-rich instruction and cognitive demand in prek: using systematic observations to predict child gains. *Early Childhood Research Quarterly*. 2022;60:96-109.

237. Araujo MC, Carneiro P, Cruz-Aguayo Y, Schady N. Teacher Quality and Learning Outcomes in Kindergarten. *The Quarterly Journal of Economics*. 2016;131(3):1415–1453.

238. Offord Centre. Early Development Instrument Validity and Reliability. Retrieved from: https://edi.offordcentre.com/researchers/validity-and-reliability/.

239. Forget-Dubois N, Lemelin JP, Boivin M, et al. Predicting early school achievement with the EDI: A longitudinal population-based study. *Early Education & Development*. 2007;18(3):405-426.

240. D'Angiulli A, Warburton W, Dahinten S, Hertzman C. Population-Level Associations between Preschool Vulnerability and Grade-Four Basic Skills. *PLOS ONE*. 2009;4(11):e7692.

241. Glascoe FP, Sandler H. Value of parents' estimates of children's developmental ages. *The Journal of Pediatrics*. 1995;127(5):831-835.

242. Heyman M, Poulakos A, Upshur C, Wenz-Gross M. Discrepancies in parent and teacher ratings of low-income preschooler's social skills. *Early Child Development and Care*. 2018;188(6):759-773.

243. Achenbach TM. Commentary: Definitely More Than Measurement Error: But How Should We Understand and Deal With Informant Discrepancies? *Journal of Clinical Child & Adolescent Psychology*. 2011;40(1):80-86.

244. WHO Multicentre Growth Reference Study Group. WHO Motor Development Study: windows of achievement for six gross motor development milestones. *Acta Paediatr*. 2006;450:86-95.

245. Trommsdorff G, Friedlmeier W, Mayer B. Sympathy, distress, and prosocial behavior of preschool children in four cultures. *International Journal of Behavioral Development.* 2007;31:284–293.

246. Callaghan T, Corbit J. Early prosocial development across cultures. *Current Opinion in Psychology.* 2018;20:102-106.

247. Bornstein M. From Measurement to Meaning in Caregiving and Culture, Current Challenges and Future prospects. In: Wortham C, Plotsky PM, Schechter DS, Cummings CA, eds. *Formative Experiences. The Interaction of Caregiving, Culture and Developmental Psychobiology.* Vol 1. New York: Cambridge University Press; 2010:36-50.

248. Bhutta ZA, Boerma T, Black MM, Victora CG, Kruk ME, Black RE. Optimising child and adolescent health and development in the post-pandemic world. *The Lancet.* 2022;399(10337):1759-1761.

249. Vandenberg RJ, Lance CE. A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods.* 2000;3(1):4-70.

250. Tol WA, Greene MC, Lasater ME, et al. Impact of maternal mental health interventions on child-related outcomes in low- and middle-income countries: a systematic review and meta-analysis. *Epidemiol Psychiatr Sci.* Oct 19 2020;29:e174.

251. WHO. *Improving early childhood development: WHO guideline.* Geneva: World Health Organization; 2020.

252. Knerr W, Gardner F, Cluver L. Improving Positive Parenting Skills and Reducing Harsh and Abusive Parenting in Low- and Middle-Income Countries: A Systematic Review. *Prevention Science.* 2013;14(4):352-363.

253. Taylor RD, Oberle E, Durlak JA, Weissberg RP. Promoting Positive Youth Development Through School-Based Social and Emotional Learning Interventions: A Meta-Analysis of Follow-Up Effects. *Child Development.* 2017;88(4):1156-1171.

254. Franklin C, Kim JS, Beretvas TS, et al. The Effectiveness of Psychosocial Interventions Delivered by Teachers in Schools: A Systematic Review and Meta-Analysis. *Clinical child and family psychology review.* 2017;20(3):333-350.

255. Kruk ME, Lewis TP, Arsenault C, et al. Improving health and social systems for all children in LMICs: structural innovations to deliver high-quality services. *The Lancet.* 2022;399(10337):1830-1844.