

# Practice Note: Measuring Outcomes in Outcomes-Based Financing for Early Childhood Care and Education

Prepared by NORRAG, the Global Education Centre of the Geneva Graduate Institute

October 2024

## Contents

Introduction .....	4
Selecting Targeted Outcomes and Measurement Tools .....	6
Existing Experiences of Measurement in OBF4ECCE Projects .....	11
Conclusion.....	22
References .....	23
Appendix 1. Useful Resources .....	27
Appendix 2. Measurement Tool Descriptions .....	29
Appendix 3. Case Study: The Utah High-Quality Preschool Pay for Success Program: Using a Measurement Tool to Identify High-Risk Children .....	33
Appendix 4. Case Study: Choosing Measurement Tools to Build Access and Quality in Early Childhood Development Centres in Sierra Leone .....	35

## Acknowledgements

This practice note was developed in partnership with the Education Outcomes Fund (EOF), a UNICEF hosted fund, and with the support of the LEGO Foundation. The statements in this report are the views of the authors and do not necessarily reflect the policies or the views of EOF, UNICEF or the LEGO Foundation.

## Suggested Citation

Thorne, G., Adams, L., & Terway, A. (2024). *Practice Note: Measuring Outcomes in Outcomes-Based Financing for Early Childhood Care and Education*. NORRAG, the Global Education Centre of the Geneva Graduate Institute.

# 1 Introduction

Measuring outcomes is a fundamental component of any outcomes-based financing (OBF) initiative as it provides the basis for determining whether payments should be made to the service providers for the project's actual and pre-agreed achievements (Clist & Vershoor, 2014). Early childhood care and education (ECCE) projects present unique challenges for the outcomes measurement process. These challenges stem from the multidimensional nature of child development, which varies significantly across the 0-8 age range, where rapid and dynamic growth occurs. Additionally, the diversity of settings and approaches to ECCE service delivery, coupled with the scarcity of valid, reliable, and standardised measurement tools tested in these diverse contexts further complicate the accurate assessment of outcomes in the sector (Fernald et al., 2009).

This practice note provides some insights into the prominent issues and important considerations that need to be made for measurement in outcomes-based financing for early childhood care and education (OBF4ECCE) projects. The note is supplemented by existing experiences from projects at various stages of development, from design to implementation and completion. This note is primarily intended for OBF4ECCE project designers. It could also be of interest to researchers and other ECCE practitioners in understanding the challenges and status of outcomes measurement in the sector.

## Box 1. Definitions

For this practice note, we define **early childhood care and education (ECCE)** as **learning opportunities for children aged 0-8, including early stimulation, education, guidance, and developmental activities, and incorporating the importance of holistic support at this stage of development, which encompasses additional care domains such as health, nutrition, sanitation, hygiene and protection.** It takes place at home and in the community and is provided through organised services and programmes that target children directly or indirectly (i.e. targeting their parents and other primary caregivers in order to improve their care and education practices vis-à-vis their own children). ECCE programmes also support primary caregivers in enhancing their own and their children's wellbeing through micronutrient supplementation, psychosocial support, parental leave, and childcare. Quality ECCE provision varies across different cultural and country contexts and leverages a variety of resources to meet the specific needs of each child (UNESCO, 2016; UNESCO, 2022; UNESCO, 2021). The information in this note may also refer to early childhood development (ECD) as holistic healthy development of children from birth up to the age of 8, encompassing services in various domains health, nutrition, social protection and stimulation to support motor, cognitive, language, socio-emotional and self-regulatory skills and capacities (UNCIEF, 2023).

**Outcomes-based financing (OBF)** is defined as a financing **mechanism whereby service providers are contracted based on the achievement of outcomes. This can entail tying outcomes into the contract and/or linking payments to the achievement of outcomes.** This definition includes **impact bonds** as a sub-set of OBF. Impact bonds incorporate the use of private funding from investors to cover the upfront capital required for a provider to

set up and deliver a service. The service is set out to achieve measurable outcomes established by the commissioning authority (or outcome payer) and the investor is only repaid if these outcomes are achieved. Impact bonds encompass both social impact bonds and development impact bonds. This definition of OBF intentionally excludes projects that provide incentives to individuals (e.g. conditional cash transfers or individual teacher's pay bonuses) or to whole systems (e.g. debt swaps).

Summary of recommendations on issues to consider when designing the outcomes measurement process for OBF4ECCE:

- **Primary and secondary purposes for measurement:** Determine whether the purpose of the measurement process is only to define the achieved outcomes for making payments to the service provider or whether the measurement process can also help establish baselines, set targets, provide formative assessments to service providers, and generate evidence for the ECCE intervention.
- **Iterative process:** Treat the selection of targeted outcomes and measurement tools as an iterative process. Initial targeted outcomes can be refined once specific measurement tools are selected or more information becomes available, particularly during the early stages of project implementation.
- **Engagement of stakeholders:** Involve relevant actors, including local ECCE experts, psychometricians, teachers, and government representatives, early in the design process to ensure that the targeted outcomes and measurement tools are realistic, achievable, and aligned with local ECCE priorities.
- **Holistic nature of child development:** Determine whether the project should focus on measuring a specific aspect of ECCE or take a multidimensional approach measuring multiple development domains such as cognitive, language, motor, socio-emotional, and executive functions, as these domains are interrelated and essential for assessing holistic outcomes. Also, consider the entire ecosystem of ECCE provision, which includes caregivers, teachers and the community.
- **Contextually relevant measurement tools:** Assess the trade-offs between adapting existing measurement tools to the specific context or developing new tools that are contextually relevant to the target population, considering factors like regional linguistic, cultural aspects, and developmental status. This is particularly important in diverse settings or in low-and middle-income country contexts that do not have well-tested tools for all populations. Evaluate the feasibility of implementing the measurement process, considering the costs of training assessors, administering tests, and ensuring the reliability and validity of the tools. Consider the trade-offs between the ideal measurement approach and what is feasible within the project's timeframe and budget.
- **Risks of perverse incentives:** Be aware of the potential for gaming or cheating due to the financial risks tied to achieving targeted outcomes. Introduce verifiable metrics

and checks to reduce these risks, but also weigh the additional costs and potential distortions this might introduce.

## 2 Selecting Targeted Outcomes and Measurement Tools

**The first few years of a child's life witness rapid brain development.** During these formative years, young children stand to benefit from positive, supportive caregiving and environments, yet are also susceptible to negative impacts from adversity. Investing in ECCE reaps lifelong economic and social benefits, building the foundations for productive workforces and stronger economies. Children from disadvantaged backgrounds are more likely to start primary school behind their higher-income peers, and these gaps often grow over time. **ECCE is one of the most effective ways to address social inequities, enabling all children to enter primary school with higher-level skills that lead to academic advancement, social engagement and, later, success in the workforce** (UNESCO, 2022; Britto et al., 2017).

Measurement in OBF4ECCE projects needs to consider a range of issues due to the holistic nature of **child development and the complex interplay between children's inherent individual differences and variations in practices, expectations, and environments in their households and communities.** High-quality early childhood education does not only build children's early literacy and numeracy skills. More importantly, it supports children's individual, holistic development, including their ability to form relationships with peers and responsive caregivers, manage conflict and frustration, identify and communicate their emotions, strengthen their bodies through good nutrition, healthy practices, and movement, and develop cognitive skills and knowledge through play and exploration (Britto et al., 2017).

Making outcome payments contingent on achieving the pre-defined outcomes incentivises service providers to ensure their efforts improve outcomes for the targeted ECCE population. As such, clearly defined targeted outcomes and an appropriate tool to measure these outcomes need to be determined during the project design phase. **Different targeted outcomes and measurement tools may be more or less appropriate depending on the purpose of measurement, the population being assessed, the intended use of measurement, and the costs in terms of time and money required to gather data.**

**The selection of targeted outcomes and relevant measurement tools is an iterative process.** Project designers may initially choose general targeted outcomes closely aligned with the project objective. Once appropriate measurement tools are identified, targeted outcomes could be refined further with specific definitions (like learning outcomes defined as literacy) and target metrics (typically a numerical goal). Furthermore, designers may also consider multiple **targeted outcomes along the line of theory of change that observes process, short-term, medium-term and long-term outcomes.** Defining targeted outcomes along the intervention theory of change, especially for process, short-term and medium-term outcomes may help in observing changes in the ECCE ecosystem within the local context (Raikes, 2023).

When deciding targeted outcomes, it could be beneficial to **consult relevant local actors,** including but not limited to ECCE experts, psychometricians, teachers and local government representatives, to set achievable and appropriately ambitious targeted outcomes aligned

with local ECCE priorities. However, even with these consultations, determining targeted outcomes that are measurable, affected by the intervention and realistic within the project timeframe can be challenging.

Over the last two decades, global efforts have been made to design and test ECD measurement tools by multi-lateral organisations, non-governmental organisations, private funders and universities. Typically, most tools focus on developmental milestones for children at preschool age. However, not all tools are appropriate for addressing the diverse contextual environment of the children (Raikes, 2023). This is a particular **challenge in low-and middle-income countries** where the supply of ECCE services remains low.

Furthermore, certain tools and uses of tools may be more **susceptible to inducing gaming, cheating, and perverse incentives**. In contrast with traditional grant-funded projects, in OBF the financial risk of achieving the targets typically lies with the service provider (or the investor in the case of impact bonds). By introducing performance incentives, the service providers (or investors) now stand to make a profit or loss depending on whether the targeted outcomes are met and how much they spend to achieve the outcomes (Terway, Burnett & Dreux-Frotte, 2021). Examples of this could include excluding children from an intervention or an assessment who are likely to score low on a measurement tool, teaching to the test to the detriment of other important dimensions of quality early childhood education, or assessors falsifying data. Project designers could add extra verifiable metrics to create checks on cheating. However, collecting and verifying additional metrics poses costs and may induce service providers to focus on achieving the easiest to attain outcomes (Lee & Medina, 2019).

Borrowing from the guidance developed by Pushparatnam, Seiden and Luna Bazaldua (2022) on choosing the right tools for measuring early childhood outcomes, the remaining section frames **key questions for selecting targeted outcomes and measurement: why, who, what, and how**. Box 2 introduces the framework.

### **Box 2. Practical Guidance for Measuring Outcomes in OBF4ECCE**

#### **Why:**

- Clarify the intended project goal to determine appropriate targeted outcomes.
- Clarify the intended purpose of the measurement tool: determining the disbursement of outcome payments; establishing a baseline and targets to measure; conducting formative assessments for continuous improvement of interventions; and generating an evidence base for the intervention.

#### **Who:**

- Identify the target population for the project and the population sample to be assessed to demonstrate the achievement of outcomes: age; regional, linguistic or cultural aspects; developmental status/ ability (i.e. developmental delays or disabilities).

#### **What:**

- Determine what category of targeted outcomes or developmental domains (multi-domain or domain-specific) will be scored and assessed.

**How:**

- Assess the feasibility of achieving targeted outcomes: if the intervention leads to the outcome and if they are measurable within the project timeframe.
- Assess the feasibility of administering the measurement tool: cost, time and training of assessors; standardisation, reliability and validity of the tool; types of assessment – direct (with trained assessors) or indirect assessments (teacher or parent report).

*Source: Adapted with author analysis of OBF4ECCE from: Pushparatnam et al., 2022, Fernald et al., 2009; Rohacek & Isaacs, 2016.*

## 2.1 Why

OBF4ECCE project designers need to design a robust measurement process that includes clearly defined targeted outcomes and appropriate measurement tools. The primary reason for this is to attribute the effect of service provider interventions on the pre-agreed results. By doing so, the effectiveness of these interventions can be accurately assessed, **ensuring that payments are made based on the actual achievements** of the project, in line with the OBF contract.

Beyond this core function, a well-designed measurement process can serve several additional purposes. First, it can help **screen target beneficiaries**, ensuring that the right population is served. Second, it can help **establish a baseline and set realistic, data-driven targets** for the remainder of the project. In some instances, projects can maintain flexibility by defining these targets as indicative rather than fixed from the outset. This approach allows for adjustments based on data gathered during the early stages of the project, avoiding the potential pitfalls of “moving the goalposts” and losing stakeholder confidence.

Moreover, the measurement process can **provide formative assessments** that enable service providers to modify their interventions as needed. This can affect decisions on the frequency of assessments, whether all beneficiaries are measured, or a representative sample is used. Furthermore, robust measurement tools contribute to **generating a strong evidence base for the intervention**. The evidence generated by measuring OBF interventions could not only support the immediate project but also help contribute to the larger knowledge base on ECCE and strengthen the ECCE system (Raikes, 2023; Rohacek & Isaacs, 2016).

## 2.2 Who

The multidimensional nature of ECCE varies significantly across ages 0-8 with interventions targeting young children both directly and indirectly (targeting caregivers, teachers or centres). **Determining the appropriate sample of the target population that should be assessed** to observe the project results is required when considering which targeted outcomes and measurement tools to use. While the project's overall objective could be **developmental improvements in young children**, the project may also need to observe **improvements in teachers' and caregivers' behaviour** as a targeted outcome.



Evaluating targeted outcomes requires identifying a measurement tool(s) with appropriate **age coverage**, especially if different age ranges are involved, or the intention is to track progress over time (Pushparatnam et al., 2022). This may require different tools for each targeted outcome within a single project. **Regional, linguistic and cultural aspects** should also be taken into consideration. This could include the language of instruction and integrating context-relevant factors to assess. Finally, the relevance of measurement tools should be considered if targeting children with **developmental delays or disabilities** to ensure they accurately capture outcomes (Pushparatnam et al., 2022).

**Measurement tools that capture caregiver, health-care provider, or teacher observations** may need to be used in some instances where the child is too young to directly test for development (Fernald et al., 2009).

### 2.3 What

Young children require multiple types of support to promote their development. ECCE should support children’s holistic development, including access to early learning opportunities, healthcare, good nutrition, protection from violence, abuse and neglect, responsive caregiving, and social safety nets to mitigate the effects of poverty and economic shocks (World Health Organization, 2018). These inputs are all crucial for promoting the different yet interrelated domains of child development: cognitive, language, motor development, social-emotional, and/or executive function/self-regulation (Early Learning Partnership, 2016). These domains are presented in Box 3 below.

#### **Box 3. Developmental domains**

**Cognitive:** Problem-solving, memory, measurement and comparison, early mathematical and pre-literacy skills, and analytical thinking.

**Language:** Understanding and use of words in print and oral form. Ability to tell stories, identify letters and familiarity with books.

**Motor development:** Ability to control movements to perform everyday tasks. Includes gross motor skills (walking, running, jumping, and movement in games) and fine motor skills (drawing, writing, hand-eye coordination and muscle control).

**Socio-emotional:** Awareness of one’s feelings, social interactions with peers and teachers, and behaviour (ability to follow directions, cooperate with requests and control emotions in stressful situations).

**Executive function:** Self-control, persistence, ability to sustain attention and initiate action often when confronted with a novel situation, problem or stimulus.

*Sources: Early Learning Partnership, 2016; Fernald et al., 2009*

Measurement tools **often focus on child developmental outcomes such as the five domains** described above. Multi-domain tests may ask children to perform a variety of tasks or ask their caregivers or teachers if they can perform these tasks or exhibit certain behaviours. These could include motor activities like jumping up and down, using a writing implement to draw a line or asking children to name several types of foods. Children could be asked to count objects or demonstrate that they know the correct orientation to open a book and from which direction the text flows. A test could ask caregivers or teachers if the child seeks comfort when distressed or consoles their peers when they are sad.

Some measurement tools aim to gauge children's readiness for primary school. However, not all of these tools have predictive validity, meaning that they may not be able to predict how well the child would perform in the future in primary school and beyond. Some school readiness tools may not capture or recognise the value of play-based preschool approaches or the importance of supporting children's social-emotional development, both of which have many positive effects beyond academic achievement enduring long after primary and secondary school completion. Similarly, existing measurement tools used in ECCE may not capture the many long-term, positive externalities that high-quality preschool may yield, including better health outcomes, increased maternal labour force participation and household income, and lower crime and higher incomes later in life.

Other tools assess elements of structural and process quality. Typically, structural quality refers to environmental characteristics (such as infrastructure, equipment and materials, play spaces, staff-child ratios and group sizes), caregiver behaviour (caregiver-child interaction and language at home, nutrition and health care support provision) and teacher characteristics (teacher training, qualifications, and professional development opportunities). Process quality refers to the types and quality of interactions between the children, their caregivers, peers, teachers and environments (Fernald et al., 2009; Yoshikawa & Kabay, 2014; Whitebread et al., 2019).

## **2.4 How**

After considering the why, who and what, choosing the measurement tools to assess the achievement of targeted outcomes requires considering the how. This involves assessing costs, weighing trade-offs, and exploring existing standardised tools.

Determining the most appropriate measurement tool requires **assessing costs**, especially relative to total project costs. Some measurement tools are government-produced or public goods and therefore essentially free to use. Others are commercially licensed, and access may be subject to fees, and this should be assessed against the actual payments available for outcomes. Other feasibility considerations are the time and costs needed to train assessors and administer tests. For example, does the tool require particular equipment and assessors need specific training? Does administering the test require renting space? Do children and their families need to travel to the assessment centre? For contexts where tools do not already exist, the development and implementation of the ideal tool may end up costing more, or as much as, the entire outcomes payment. In these situations, designers will need to make some trade-offs.

This may be more difficult in **low-and middle-income countries** than in higher-income countries. The challenge is particularly acute in contexts with emerging ECCE systems where government-developed measurement tools may not yet exist. Alternatively, tools produced in the Global North may not be valid for local populations and contexts. Some tools are designed for global use, based on universal aspects of ECCE. Although these are available in multiple languages and offer degrees of comparability, they may not incorporate contextual nuances (Pushparatnam et al., 2022). Choosing a measurement tool requires balancing the two elements of contextual relevance with global comparability. This is often overcome by **adapting existing tools** such as adding context-specific items (Pushparatnam et al., 2022).

Successfully administering tools relies on the buy-in and confidence of financing partners, implementers, evaluators and in many cases, local and national governments. Using **existing norms-based tools** (a child's performance can be compared to another child of the same age) or **standards-based tools** (scores are compared to levels deemed attainable and desirable by experts) can generate confidence that the measurement tool will capture what is valued in the ECCE system (Fernald et al., 2009). Choosing a tool already in use by governments could facilitate governmental engagement and alignment with the ECCE system.

Measurement tools must be reliable and valid. **Reliability** means if measurements are repeated over time by different assessors or enumerators for the same child or group, they will produce similar results. It is often ensured by uniformly administering tests (Fernald et al., 2009; Pushparatnam et al., 2022). **Validity** refers to the accuracy of a measurement tool to assess behaviour or ability and needs to be contextually determined (Fernald et al., 2009).

Tools that have not been widely used in the target population and for a similar intervention may not have yielded adequate data to determine an appropriate target for payment, making setting a target metric extremely difficult. **The greater the evidence base, the easier it is to set a target on a measurement tool.** Consultative processes with relevant actors, including local ECCE experts, psychometricians and teachers, are important to set targets for payment that are achievable and appropriately ambitious.

### 3 Existing Experiences of Measurement in OBF4ECCE Projects

This section discusses the targeted outcomes and measurement tools used in 15 out of 22 OBF4ECCE projects (with available data) identified through a desk review that aimed to surface, analyse and share experiences of OBF in ECCE (Airoldi et al., 2024). The analysed projects were at various stages of development, from design to implementation and completion. Additionally, some were conceptualised but did not launch. This note primarily draws upon publicly available information and, therefore, does not claim to offer exhaustive knowledge on the projects. Supplementary interviews with key stakeholders from two projects were also used to delve deeper into the measurement tools used in the Utah High-Quality Preschool Program and the Sierra Leone ECCE Outcomes Fund. Using the framing questions of why, who, what and how, this section explores existing experiences of targeted outcomes and measurement tools in OBF4ECCE projects.

### 3.1 Why

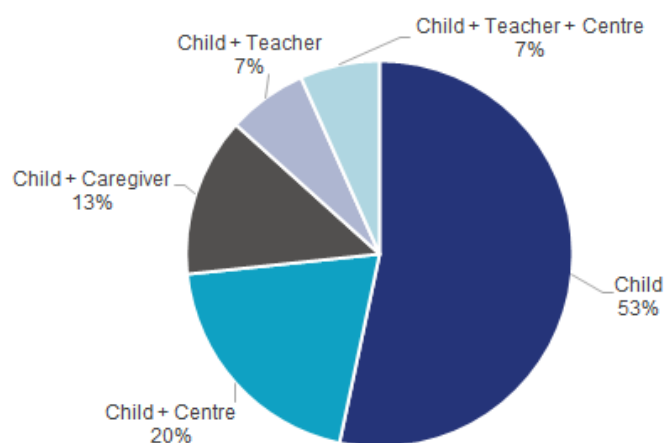
Given that measurement is a fundamental component of OBF contractual arrangements, the primary motivation for all projects in designing the measurement process was for the outcome payer to disburse payments to the service provider based on the achievement of verifiable and pre-defined targeted outcomes. The measurement process would give the stakeholders confidence that the funded intervention yielded the desired outcomes.

In some projects, measurement serves additional purposes. In the Utah High-Quality Preschool Program Pay for Success, the measurement tool (Peabody Picture Vocabulary Test) was used at the beginning of the project to identify low-income preschool-age children at high risk of later receiving special education services. The targeted outcome was established as the number of children from this high-risk group that did not require special education services in later years. In the Impact Bond Innovation Fund in South Africa, there is also some evidence that the service provider used the data from the annual assessments in a formative manner to monitor their progress and modify their approach to service delivery (Rayner & Nkonyeni, 2021).

### 3.2 Who

ECCE projects target children both directly and indirectly (i.e. targeting caregivers, teachers or centres). Therefore, successes in existing OBF4ECCE projects are measured by assessing achievements made by children, caregivers, teachers and/ or centres. These different units of measurement are used in varying combinations in OBF4ECCE projects, as presented in Figure 1 below.

**Figure 1 Unit of Measurement (%)**



The majority of projects use targeted outcomes associated with children. Projects that combine targeted outcomes focusing on children and **centres** include the Rwanda ECCE Outcomes Fund, which aims to improve the structural quality of ECCE centres in addition to improving child development outcomes. **Caregivers** are included as a unit of measurement in both the ParentChild+ and West London Zone (scale-up) projects in the UK for increased parental self-efficacy and giving parental consent for participation, respectively. Programa

Primer Lee in Chile includes both **children and teachers** as a unit of measurement in which it aims to ensure the quality implementation of the educational intervention by teachers to enable improved literacy skills among children. The Child Development Centers project in Colombia measures against the **children, teachers and centres** by tying outcome payments to holistic early childhood assessments, teachers' pedagogy and classroom interactions as well as the physical environment of the ECCE centres. It should be noted that in some cases, there is not a clear delineation between targeted outcomes that concern the centres or teachers, as process quality may include teacher practices within centres.

Within OBF4ECCE projects, the **age range of the beneficiaries at which the outcomes are expected to be realised varies**. For example, the Child-Parent Center Pay For Success Initiative, which provided preschool education to three and four year olds, tracks child outcomes beyond preschool, measuring each year a child does not use special education services up to Grade 12. It also includes targeted outcomes related to kindergarten readiness and third-grade literacy. Considering the age differences at these two moments, two different tools are implemented to measure outcomes. First, the Teaching Strategies Gold Assessment indicated kindergarten readiness by measuring child developmental domains. Second, the Partnership for Assessment of Readiness for College and Careers (PARCC) measurement tool is intended to measure third-grade literacy levels. PARCC is a standardised exam used to calculate the percentage of students who read at or above grade level (Gaylor et al., 2019).

The Utah High-Quality Preschool Program highlights the need to consider linguistic factors within the target population, which included both native Spanish and English speakers. The project used Peabody Picture Vocabulary Test (PPVT) to predict the likelihood of needing special education services in later years. The PPVT is relatively simple and quick to administer. It has been widely used for decades, is normed, has been used in longitudinal studies, and is available in numerous languages and for multiple age groups. The project designers administered the Spanish and English versions of the test to a sample of the Spanish-speaking population and found that these children scored similarly on both versions of the test. This gave the project designers confidence in the validity of the English version of the test for native Spanish-speakers and they continued with only administering the English version to the rest of the target population.

### **3.3 What**

#### *3.3.1 Targeted Outcomes*

OBF4ECCE projects use a wide range of targeted outcomes linked to outcome payments. Table 1 presents the targeted outcomes used in OBF4ECCE projects.

#### *Number of targeted outcomes*

Most OBF4ECCE projects used multiple targeted outcomes (some up to eight). Using multiple targeted outcomes potentially allows the stakeholders to incentivise and observe a holistic change in the ECCE service provision. For example, ParentChild+ in the UK set targets for different groups of beneficiaries accessing the services, as well as differentiated targets and measurement tools for various age groups. Whereas, in some projects, a single targeted outcome would allow stakeholders to focus on the most critical challenge identified for the

beneficiaries, like Cantonese language skills for children who do not speak it as a first language in the Start from the Beginning project in Hong Kong.

### *Categories of targeted outcomes*

The targeted outcomes used in OBF4ECCE projects can be categorised into six groups: learning and developmental outcomes, participation, enrolment, structural quality, process quality and caregiver engagement. While most projects primarily use **learning and developmental outcomes** in their measurement process, including the additional outcomes has the potential to incentivise service providers to develop the supporting ecosystem for providing quality ECCE.

Multiple projects include targeted outcomes related to **structural and process quality**. Support for structural quality could ensure the availability of adequate resources for ECCE provision. It could include improving safety and infrastructure through providing toilets and handwashing stations, reducing class sizes and hiring additional staff. The Rwanda and Sierra Leone ECCE Outcomes Funds both plan to tie outcomes payments to elements of structural quality including basic sanitation, safety, class sizes and instructional materials (EOF, 2023a; 2023b). While process quality, which refers to interactions and relationships between children, teachers and caregivers, may be more difficult to change, the two funds in Rwanda and Sierra Leone also intend to incorporate this in the targeted outcomes (EOF, 2023a; 2023b). The Child Development Centers project in Colombia also included measures on classroom interactions between teachers and children as targeted outcomes (UNICEF et al., 2022).

**Enrolment and participation** targeted outcomes are also common. Although they do not necessarily guarantee learning, they are relatively objective and easy to assess within project timeframes. For OBF4ECCE projects targeting vulnerable and marginalised populations, these categories could demonstrate important steps towards inclusion and subsequent provision of learning opportunities. Nevertheless, it is important to recognise that these targeted outcomes are not used in isolation. All projects using enrolment and participation targeted outcomes also include at least one targeted result relating to learning and development outcomes. For example, the Play2Learn+ project in Australia uses enrolment, attendance and developmental achievements as its targeted outcomes. The project's theory of change illustrates that the strategic activity of targeting disengaged children and families is one of the core components leading to the longer-term outcome of early childhood development results (Moore & Arefadib, 2022).

The final and the least frequently used category of targeted outcomes is **caregiver engagement**, used only in two projects in the UK: the ParentChild+ and West London Zone projects. These examples indicate conscious efforts to ensure holistic child development, given that responsive caregiving is a core component of the Nurturing Care Framework (2018).

### *3.3.2 What: Measurement Tools*

A range of measurement tools have been applied to determine the success of OBF4ECCE projects against their pre-determined targeted outcomes. Table 2 lists information on these

tools including the project in which it was used and which developmental domains it assessed. A short description and relevant resources for the tools is included in Appendix 2.

Table 2 identifies 19 measurement tools used in OBF4ECCE projects. The majority of which measure **learning and developmental outcomes**. For example, the Impact Bond Innovation Fund used the Early Learning Outcomes Measure (ELOM) to determine how well the educational intervention prepared children for school. It involved an ECD practitioner administering a test to examine a child's motor development, cognitive and executive functioning and language skills. While the tool was developed for centre-based interventions, the project adapted the tool to be used for home-based interventions.

**Table 1. Targeted outcomes used in OBF4ECCE projects**

Project Name	Location	Targeted Outcomes	Category
Alexandria ECD SIB*	United States	Reduction in special education placements	Learning and developmental outcomes
Child Development Centers	Colombia	Number of dimensions in which the child advances or stays the same measured against a national ECD assessment	Learning and developmental outcomes
		Quality of the ECD centre across two dimensions: 1) physical environment of the centre and classrooms; 2) interactions and pedagogical intention	Structural & process quality
Child-Parent Center Pay for Success Initiative	United States	Improve kindergarten readiness	Learning and developmental outcomes
		Improve third-grade literacy	Learning and developmental outcomes
		Reduce the need for special education services	Learning and developmental outcomes
Impact Bond Innovation Fund	South Africa	Recruitment and retention of children	Enrolment
		Attendance	Participation
		Success against the Early Learning Outcomes Measure	Learning and developmental outcomes
Namibia Early Childhood Development Social Impact Bond*	Namibia	Number of supported ECD centres	Structural quality
		Increase in the number of children attending targeted ECD centres, without exceeding the centre's capacity	Participation
		Increase in scores achieved in an ECD assessment tool	Learning and developmental outcomes
		Additional metric on child development outcomes	Learning and developmental outcomes
ParentChild+/Family Lives [Life Chances Fund]	United Kingdom	Speech and language skills at ASQ level for 94 children within 16 months of starting the project	Learning and developmental outcomes
		Speech and language skills at EYFS level for 77 children within 39 months of starting the project	Learning and developmental outcomes
		Social-personal skills at ASQ level for 94 children within 16 months of starting the project	Learning and developmental outcomes
		Social-personal-emotional skills at EYFS level for 77 children within 39 months of starting the project	Learning and developmental outcomes
		Parental self-efficacy increased by 5% over 15 months of the project compared to baseline assessment for 94 beneficiaries	Caregiver engagement
		150 beneficiaries initially engaging	Enrolment
		98 beneficiaries completing the first cycle of the programme	Participation
94 beneficiaries completing the second cycle of the programme	Participation		
Play2Learn+	Australia	Programme enrolment numbers	Enrolment
		Attendance at the Launching into Learning programme	Participation
		Achievement on the 'Kindergarten Development Check'	Learning and developmental outcomes



Programa Integrado de Promoçãoda Literacia (Integrated Literacy Programme)	Portugal	30% increase in literacy skills	Learning and developmental outcomes
Programa Primero Lee (Read First Programme)	Chile	Quality implementation of Primero LEE by teachers and school leadership teams, measured by activity checklists	Structural quality
		Level of children’s skill in reading comprehension, reading aloud, and writing, measured by tests the children take	Learning and developmental outcomes
Rwanda ECCE Outcomes Fund	Rwanda	Settings and practices meeting quality standards – Structural quality	Structural quality
		Settings and practices meeting quality standards – Process quality	Process quality
		Improved holistic child development outcomes	Learning and developmental outcomes
Sierra Leone ECCE Outcomes Fund	Sierra Leone	Centres are safe for children before they open	Structural quality
		Increased access of children ages 3-5 within teacher-child ratios	Enrolment
		Settings and practices meeting quality standards - structural and process quality	Structural & process quality
		Improved holistic child development outcomes	Learning and developmental outcomes
Start from the Beginning - Chinese Supporting Scheme for Non-Chinese Speaking Students (NCS) in Kindergarten	Hong Kong	Narrow the proficiency gap between Chinese and non-Chinese students by half	Learning and developmental outcomes
Tennessee High-Quality Preschool Program	United States	Consistent attendance	Participation
		Development of early reading skills	Learning and developmental outcomes
		Kindergarten readiness	Learning and developmental outcomes
Utah High Quality Preschool Program	United States	Decrease in the use of special education between grades K-6	Learning and developmental outcomes
West London Zone (scale-up) [Life Chances Fund]	United Kingdom	Children sign up to the West London Zone (WLZ) programme	Enrolment
		Participating children engage with support to a ‘high’ level so that they are more likely to achieve the primary outcome	Participation
		Child achieves at least one outcome on the WLZ rate card to get on track to flourish in adulthood (e.g.) less likely to go on to experience NEET-status and/or crime and/or poor wellbeing by adulthood and more likely to achieve goals for work/college/university	Learning and developmental outcomes
		Child achieves at least two outcomes on the WLZ rate card to get on track to flourish in adulthood (e.g.) less likely to go on to experience NEET-status and/or crime and/or poor wellbeing by adulthood and more likely to achieve goals for work/college/university.	Learning and developmental outcomes
		Child achieves three outcomes on the WLZ rate card to get on track to flourish in adulthood (e.g.) less likely to go on to experience NEET-status and/or crime and/or poor wellbeing by adulthood and more likely to achieve goals for work/college/university.	Learning and developmental outcomes
		Parent consent given to participate	Caregiver engagement

\*Projects that did not launch

Most of the tools used to measure learning and developmental outcomes assess **multiple domains** of development, reflecting the recognition of the holistic nature of child development in early years. For example, the national tool Escala Abreviada de Desarrollo (EAD-3) used in the Child Development Centers project assesses child development by measuring motor development, hearing and language, and personal-social skills (UNICEF et al., 2022). Additionally, the Brigance Inventory of Early Development III Standardized (Brigance IED III) tool used in the Tennessee High-Quality Preschool Program scores children based on their motor skills, language, cognitive and socio-emotional development.

Alternatively, some tools assess **specific domains**. The Start from the Beginning project in Hong Kong uses a language assessment to gauge improvements in Cantonese proficiency. The Peabody Picture Vocabulary Test in the Utah High-Quality Preschool Program assessed language skills.

Other tools identified measure **structural and/ or process quality**. For example, the outcomes funds in Rwanda and Sierra Leone consider using the Brief Early Childhood Quality Inventory (BEQI) to measure process quality (EOF, 2023a; 2023b). A tool based on best practices in ECCE, BEQI is a 50-item checklist based on observing learning environments. It covers play-based learning, learning through conversations and promoting strong relationships (ECD Measure, n.d.).

One tool measures elements of **caregiver engagement**. The ParentChild+ project uses the Parent and Child Together tool to examine parenting skills and parent-child relationships. It uses a 20-point questionnaire to assess the four elements of communication, affection, consistency and responsiveness (Flynn et al., 2022). No specific standardised tools for measuring enrolment and participation were identified for OBF4ECCE projects.

In addition to the tools used in OBF4ECCE projects, other tools were considered but not used. For example, three tools were considered for the Sierra Leone ECCE Outcomes Fund (EOF, 2023b). Teach ECE and MELE were considered for measuring targeted outcomes related to process quality and the Child Progress and Achievement Report was considered for measuring child development outcomes (EOF, 2024). While these tools have not been utilised in the set of OBF4ECCE projects analysed for this practice note, future projects could include them in their assessment.

### **3.4 How**

Assessing the feasibility of measurement tools for a particular project requires considering costs and trade-offs. Generally, OBF4ECCE projects use or adapt existing tools with normative data or existing benchmarks.

Generally, information on costs of measurement process in OBF4ECCE projects is lacking. The cancelled Promoting Early Childhood Development Project in Uzbekistan estimated costs for independent evaluation activities to be approximately 10% of the overall project costs (World Bank, 2018). However, because the project was cancelled, the actual cost of implementation is not available.

Choosing a measurement tool often requires some **trade-offs** between the ideal tool and what is feasible within the project timeframe. Exploring the use of measurement tools in existing OBF4ECCE projects shows a tendency to **use or adapt existing tools** rather than develop new tools. The Play2Learn+ project applies the Kindergarten Development Check tool, first developed in 1994 to assess student development outcomes in Tasmania. The Play2Learn+ project applies 21 markers to assess students' problem-solving, literacy and numeracy, and health and well-being (Department for Education, 2009; Moore & Arefadib, 2022).

**Table 2. Measurement tools used in OBF4ECCE projects**

Measurement Tool		Project Background					Child Development Domains Assessed				
Measurement Tool Name	Type of tool	Project Name	Project Location	Age of Project Beneficiaries	Category of Targeted Outcome Measured	Project Modality (home v centre-based)	Cognitive	Executive Function	Language	Motor	Social-Emotional
<b>Ages &amp; Stages Questionnaire (ASQ)</b>	For-profit	ParentChild+/ Family Lives (Life Chances Fund)	United Kingdom	2-3 years	Learning and developmental outcomes	Home-based			X		X
<b>Being a Parent Scale</b>	Non-profit	ParentChild+/ Family Lives (Life Chances Fund)	United Kingdom	2-3 years	Caregiver engagement	Home-based					
<b>Brief Early Childhood Quality Inventory (BEQI)</b>	Non-profit	Rwanda ECCE Outcomes Fund	Rwanda	3-6 years	Process quality	Centre-based					
		Sierra Leone ECCE Outcomes Fund	Sierra Leone	3-5 years	Process quality	Centre-based					
<b>Brigrance Inventory of Early Development III Standardized (Brigrance IED III)</b>	For-profit	Tennessee High-Quality Preschool Program	United States	4 years	Learning and developmental outcomes	Centre-based	X		X	X	X
<b>Cantonese language assessment</b>	Government	Start from the Beginning - Chinese Supporting Scheme for Non-Chinese Speaking Students (NCS) in Kindergarten	Hong Kong	3-6 years	Learning and developmental outcomes	Centre-based			X		
<b>Checklist of teacher activities</b>	Government	Programa Primero Lee (Read First Programme)	Chile	Teachers	Structural quality	Centre-based					
<b>Child Behaviour Traits (CBT)</b>	Non-profit	ParentChild+/ Family Lives (Life Chances Fund)	United Kingdom	2-3 years	Learning and developmental outcomes	Home-based					X
<b>Early Years Foundation Stage (EYFS)</b>	Government	ParentChild+/ Family Lives (Life Chances Fund)	United Kingdom	2-3 years	Learning and developmental outcomes	Home-based	X	X	X		X
<b>Early Learning Outcomes Measure (ELOM)</b>	Government	Impact Bond Innovation Fund	South Africa	3-5 years	Learning and developmental outcomes	Home-based	X	X	X	X	
<b>Escala Abreviada de Desarrollo (EAD-3)</b>	Government	Child Development Centers	Colombia	0-5 years	Learning and developmental outcomes	Centre-based			X	X	X

<b>Instrumento de Medición de la Calidad de la Educación Inicial en Colombia (IMCEIC)</b>	Government	Child Development Centers	Colombia	Centre	Structural & process quality	Centre-based					
<b>International Development and Early Learning Assessment (IDELA)</b>	Non-profit	Rwanda ECCE Outcomes Fund	Rwanda	3-6 years	Learning and developmental outcomes	Centre-based	X		X	X	X
		Sierra Leone ECCE Outcomes Fund	Sierra Leone	3-5 years	Learning and developmental outcomes	Centre-based	X		X	X	X
<b>Istation (only literacy test used)</b>	For-profit	Tennessee High-Quality Preschool Program	United States	4 years	Learning and developmental outcomes	Centre-based	X		X		
<b>Kindergarten Development Check (KDC)</b>	Government	Play2Learn+	Australia	3-4 years	Learning and developmental outcomes	Centre-based	X	X	X	X	X
<b>Literacy assessment (name not available)</b>	N/A	Programa Integrado de Promocao da Literacia (Integrated Literacy Programme)	Portugal	Unknown	Learning and developmental outcomes	Centre-based	X				
<b>Parents and Child Together (PACT)</b>	Non-profit	ParentChild+/ Family Lives (Life Chances Fund)	United Kingdom	2-3 years	Caregiver engagement	Home-based					
<b>Partnership for Assessment of Readiness for College and Careers (PARCC)</b>	Government	Child-Parent Center Pay for Success Initiative	United States	Grade 3	Learning and developmental outcomes	Centre-based					
<b>Peabody Picture Vocabulary Test (PPVT)</b>	For-profit	Utah High Quality Preschool Program	United States	3-4 years	Learning and developmental outcomes	Centre-based			X		
<b>Teaching Strategies Gold Assessment</b>	For-profit	Child-Parent Center Pay for Success Initiative	United States	4 years	Learning and developmental outcomes	Centre-based	X		X	X	X

For the home-based intervention in the Impact Bond Innovation Fund (South Africa), the ELOM tool had recently been developed for use in ECCE centres. On the one hand, it was rooted in South Africa's National Curriculum Framework from Birth to Four and its National Early Learning and Development Standards. It was locally developed and standardised on a sample of South African children from diverse socio-economic and linguistic backgrounds. On the other hand, the tool had not yet been deployed for the target population in Cape Town and was not designed for home-visiting interventions making it difficult to identify a valid comparison group. This was the most appropriate tool available within the context, however, during project implementation, evaluators cautioned that the ELOM tool was not adequately capturing the impacts of a home-based intervention, making the targeted outcome for early learning unrealistic.

The difficulty of choosing measurement tools may be heightened in low and middle-income countries where existing government-produced tools may not yet exist. In Rwanda and Sierra Leone, EOF is considering the IDELA tool for their outcomes funds (EOF, 2023a; 2023b). Developed by Save the Children as an open-source, easily administered and adaptable tool, IDELA is free and has already been used in Rwanda multiple times, demonstrating its potential applicability to the country's context.

Several measurement tools used in OBF4ECCE projects also have **normative data or existing benchmarks**. For example, the Teaching Strategies Gold Assessment used in the Child-Parent Centers PFS initiative compared the children's performance to the national trends of six developmental domains: literacy, language, maths, cognitive development, socio-emotional development and physical health (Gaylor et. al, 2019). This tool was used to predict kindergarten readiness. Other measurement tools are standards-based, meaning that the assessment scores show how well a child has mastered certain skills that experts believe to be attainable at a certain age. Alternatively, if benchmarks are unavailable, it is possible to use the first year of implementation to develop them and begin the OBF modality from Year 2. This can avoid the appearance of "moving the goalposts," which would erode confidence.

## 4 Conclusion

Measurement of ECCE outcomes is a fundamental component of OBF4ECCE projects, serving as the critical mechanism for determining payment disbursement based on verified outcomes. Beyond this primary role, measurement systems can also support additional functions, including identifying target populations, establishing baselines, setting realistic goals, guiding program improvements through formative assessments, and generating evidence to strengthen ECCE systems.

The experience from early OBF4ECCE projects offers valuable insights into designing effective measurement approaches. Key lessons include the need for robust tools that balance global comparability with local relevance, the importance of addressing the holistic and multidimensional nature of child development, and the necessity of engaging stakeholders across all levels to ensure buy-in and alignment with local priorities. By integrating these lessons, measurement approaches can evolve to not only meet the technical demands of OBF but also support evidence-driven investments in ECCE.

## References

- Airoldi, M., Terway, A., Wooldridge, R., Thorne, G., Adams, L., Ronicle, J., Davies, J., & Anastasiu, A. (2024). *Evidence review of outcomes-based finance for early childhood care and education*. Collective Learning Initiative on Outcomes-Based Finance for Early Childhood Care and Education, Government Outcomes Lab, Blavatnik School of Government, University of Oxford.
- Britto, P. R., Lye, S. J., Proulx, K., Yousafzai, A. K., Matthews, S. G., Vaivada, T., Perez-Escamilla, R., Rao, N., Ip, P., Fernald, L. C. H., MacMillan, H., Hanson, M., Wachs, T. D., Yao, H., Yoshikawa, H., Cerezo, A., Leckman, J. F., Bhutta, Z. A., & the Early Childhood Development Interventions Review Group. (2017). Nurturing care: Promoting early childhood development. *The Lancet*, 389(10064), 91–102.
- Clist, P., & Verschoor, A., (2014). *The conceptual basis of payment by results*. DFID. [https://assets.publishing.service.gov.uk/media/57a089bb40f0b64974000230/61214-The\\_Conceptual\\_Basis\\_of\\_Payment\\_by\\_Results\\_FinalReport\\_P1.pdf](https://assets.publishing.service.gov.uk/media/57a089bb40f0b64974000230/61214-The_Conceptual_Basis_of_Payment_by_Results_FinalReport_P1.pdf)
- Department of Education, Tasmania. (2009). *Kindergarten Development Check against the Tasmanian curriculum framework*. <https://nla.gov.au/nla.obj-1382229032/view>
- Department of Social Services, Australia. (2023). *Payment by outcomes trials*. <https://www.dss.gov.au/communities-and-vulnerable-people-programs-services-social-impact-investing/payment-by-outcomes-trials>
- Early Learning Partnership. (2016). *Measuring child development and early learning*. *Early Learning Partnership Guidance Note*. <https://documents1.worldbank.org/curated/en/659701473955877219/pdf/108286-REVISED-PUBLIC-ELP-IB4-MeasuringCD-v7-CEP.pdf>
- ECD Measure. (n.d.). *Brief Early Childhood Quality Inventory (BEQI). Fact Sheet*. <https://static1.squarespace.com/static/65f38c1951a7743d4bc0f56b/t/65fca54ee772a614c88ef626/1711056206324/BEQI-fact-sheet-global.pdf>
- Education Outcomes Fund. (2023a). *Rwanda Early Childhood Education Outcomes Fund. Call for expressions of interest*. <https://www.educationoutcomesfund.org/post/call-for-expressions-of-interest-for-rwanda-early-childhood-care-and-education-programme>
- Education Outcomes Fund. (2023b). *Sierra Leone Early Childhood Education Outcomes Fund. Call for expressions of interest*. <https://www.educationoutcomesfund.org/post/call-for-expressions-of-interest-for-sierra-leone-early-childhood-care-and-education-programme>
- Education Outcomes Fund. (2024, forthcoming). *Assessing and selecting early childhood care and education measurement tools for outcome-based financing. The experience of the Education Outcomes Fund*.

- Fernald, L. C. H., Prado, E., Kariger, P., & Raikes, A. (2009). *A toolkit for measuring early childhood development in low and middle-income countries*. World Bank Group.  
<https://documents.worldbank.org/en/publication/documents-reports/documentdetail/384681513101293811/a-toolkit-for-measuring-early-childhood-development-in-low-and-middle-income-countries>
- Flynn, S., Sutherland, D., & Centre for Educational Development, Appraisal and Research (CEDAR), University of Warwick. (2022). *Evaluation of Parent Child+ in England, 2019-2025*. [https://golab.bsg.ox.ac.uk/documents/PC\\_Warwick\\_Interim\\_Report\\_-\\_Final.pdf](https://golab.bsg.ox.ac.uk/documents/PC_Warwick_Interim_Report_-_Final.pdf)
- Garcia, J. L., Heckman, J. J., Leaf, D. E., & Pradon, M. J. (2019). *Quantifying the life-cycle benefits of a prototypical early childhood program*. (NBER Working Paper No. 23479). National Bureau of Economic Research. <https://www.nber.org/papers/w23479>
- Gaylor, E., Ferguson, K., McCracken, M., Wei, X., & Spiker, D. (2019). *Evaluation of child outcomes in nine Child-Parent Centers: Report for 2017-18*. SRI International.  
[https://iff.org/wp-content/uploads/2018/04/SIB-CPC\\_Report\\_Year4\\_FINAL\\_wAppx.pdf](https://iff.org/wp-content/uploads/2018/04/SIB-CPC_Report_Year4_FINAL_wAppx.pdf)
- Lee, J. D., & Medina, O. (2019). *Results-based financing in education: Learning from what works*. Results in Education for All Children (REACH). World Bank.  
<https://openknowledge.worldbank.org/handle/10986/31250>
- Moore, T., & Arefadib, N. (2022). *Tasmanian Play2Learn+ Trial: Evidence Review*. Centre for Community Child Health, Murdoch Children's Research Institute.  
[https://www.rch.org.au/uploadedFiles/Main/Content/ccchdev/2212\\_CCCH\\_TAS-Play2Learn.pdf](https://www.rch.org.au/uploadedFiles/Main/Content/ccchdev/2212_CCCH_TAS-Play2Learn.pdf)
- Nurturing Care for Early Childhood Development (2018). *A closer look at the nurturing care components*. [https://nurturing-care.org/wp-content/uploads/2020/12/closer\\_look\\_nov.pdf](https://nurturing-care.org/wp-content/uploads/2020/12/closer_look_nov.pdf)
- Pushparatnam, A., Seiden, J., & Luna-Bazaldua, D. (2022). Guiding questions for choosing the right tools to measure early childhood outcomes: Why, what, who, and how. The World Bank.  
<https://documents1.worldbank.org/curated/en/210101645165216405/pdf/Guiding-Questions-for-Choosing-the-Right-Tools-to-Measure-Early-Childhood-Outcomes-Why-What-Who-and-How.pdf>
- Raikes, A. (2023). Using early childhood data and measurement to leverage change: Are we making progress? *Journal of Research in Childhood Education*, 38(sup1), S180–S190.  
<https://www.tandfonline.com/doi/full/10.1080/02568543.2023.2248223>
- Rayner, C., & Nkonyeni, N. (2021). *Financing early childhood development: The Impact Bond Innovation Fund, South Africa*. NORRAG.  
<https://resources.norrag.org/resource/688/financing-early-childhood-development-the-impact-bond-innovation-fund-south-africa>



- Rohacek, M., & Isaacs, J. (2016). *PFS + ECE: Outcomes measurement and pricing: Pay for success early childhood education toolkit report #3*. Urban Institute  
<https://www.urban.org/sites/default/files/publication/82801/2000869-PFS-ECE-Outcomes-Measurement-and-Pricing.pdf>
- Terway, A., Burnett, N., & Dreux Frotte, M. (2021). *Results-based financing in education for sub-national government and school administrators—A Conceptual Framework and Practical Recommendations*. (Working Paper No. 12). NORRAG.  
<https://resources.norrag.org/resource/689/results-based-financing-in-education-for-sub-national-government-and-school-administrators-a-conceptual-framework-and-practical-recommendations>
- UNICEF. (2023). *The Early Childhood Development Index 2030. A new measure of early childhood development*. <https://data.unicef.org/resources/early-childhood-development-index-2030-ecdi2030/>
- UNICEF, Bienestar Familiar, & Instiglio. (2022). *Results-based financing (RBF) mechanism: An opportunity to improve results in early childhood development (ECD) in Colombia*.
- UNESCO. (2016). *New horizons: a review of early childhood care and education in Asia and the Pacific*. <https://unesdoc.unesco.org/ark:/48223/pf0000245728>
- UNESCO. (2021). *Inclusive early childhood care and education: from commitment to action*. <https://unesdoc.unesco.org/ark:/48223/pf0000378076>
- UNESCO. (2022). *Education starts early: Progress, challenges and opportunities; conference background report*. <https://unesdoc.unesco.org/ark:/48223/pf0000383668>
- Whitebread, D., Kvalja, M., O'Connor, A., & Qatar Academy. (2019). *Quality in early childhood education: An international review and guide for policy makers*. University of Cambridge. [https://www.wise-qatar.org/app/uploads/2019/04/wise-research-7-cambridge-11\\_17.pdf](https://www.wise-qatar.org/app/uploads/2019/04/wise-research-7-cambridge-11_17.pdf)
- World Bank. (2018). *Uzbekistan Promoting Early Childhood Development Project*. <https://documents1.worldbank.org/curated/en/472601559440909823/pdf/Uzbekistan-Promoting-Early-Childhood-Development-Project.pdf>
- World Health Organization. (2018). *Nurturing care for early childhood development. A framework for helping children survive and thrive to transform health and human potential*. <https://iris.who.int/bitstream/handle/10665/272603/9789241514064-eng.pdf>
- Yoshikawa, H., & Kabay, S. (2014). *The evidence Base on early childhood care and education in global contexts*. In Education for All Global Monitoring Report, UNESCO, & New York University, Steinhardt School of Culture, Education and Human Development, *Background Paper for the UNESCO 2015 Education for All Global Monitoring Report*. <https://repositorio.minedu.gob.pe/bitstream/handle/20.500.12799/4187/The%20Evidence%20Base%20on%20Early%20Childhood%20Care%20and%20Education%20in%20Global%20Contexts.pdf>

[20Evidence%20Base%20on%20Early%20Childhood%20Care%20and%20Education%20in%20Global%20Contexts.pdf?sequence=1&isAllowed=y](#)

## Appendix 1. Useful Resources

- Cavallera, V., Lancaster, G., Gladstone, M., Black, M. M., McCray, G., Nizar, A., Ahmed, S., Dutta, A., Anago, R. K. E., Brentani, A., Jiang, F., Schönbeck, Y., McCoy, D. C., Kariger, P., Weber, A. M., Raikes, A., Waldman, M., van Buuren, S., Kaur, R., & Janus, M. (2023). Protocol for validation of the Global Scales for Early Development (GSED) for children under 3 years of age in seven countries. *BMJ Open*, *13*(1), e062562. <https://doi.org/10.1136/bmjopen-2022-062562>
- Early Learning Partnership, Bassett, L., Devercelli, A., Solano Rocha, A., Raikes, A., Anderson, K., Manuelyan Atinc, T., Britto, P., Nieto, A., & Lara, F. (2016). *Measuring the quality of early learning programs*. Guidance Note <https://documents1.worldbank.org/curated/en/474431473958525937/pdf/108285-REVISED-PUBLIC-ELP-GN-MeasuringQuality-CEP.pdf>
- Education Outcomes Fund. (2024, forthcoming). *Assessing and selecting early childhood care and education measurement tools for outcome-based financing. The experience of the Education Outcomes Fund*.
- Fernald, L. C., Kariger, P., Engle, P., & Raikes, A. (2009). *Examining early childhood development in low-income countries: A toolkit for the assessment of children in the first five years of life*. World Bank. <https://documents1.worldbank.org/curated/en/474431473958525937/pdf/108285-revised-public-elp-gn-measuringquality-cep.pdf>
- Halpin, P. F., Wolf, S., Yoshikawa, H., Rojas, N., Kabay, S., Pisani, L., & Dowd, A. J. (2019). Measuring early learning and development across cultures: Invariance of the IDELA across five countries. *Developmental Psychology*, *55*(1), 23–37. <https://doi.org/10.1037/dev0000626>
- Pushparatnam, A., Luna Bazaldua, D. A., Holla, A., Azevedo, J. P., Clarke, M., & Devercelli, A. (2021). Measuring early childhood development among 4–6 year olds: The identification of psychometrically robust items across diverse contexts. *Frontiers in Public Health*, *9*. <https://doi.org/10.3389/fpubh.2021.569448>
- Pushparatnam, A., Seiden, J., & Luna-Bazaldua, D. (2022). Guiding questions for choosing the right tools to measure early childhood outcomes: Why, what, who, and how. The World Bank. <https://documents1.worldbank.org/curated/en/210101645165216405/pdf/Guiding-Questions-for-Choosing-the-Right-Tools-to-Measure-Early-Childhood-Outcomes-Why-What-Who-and-How.pdf>
- Raikes, A. (2023). Using early childhood data and measurement to leverage change: Are we making progress? *Journal of Research in Childhood Education*, *38*(sup1), S180–S190. <https://www.tandfonline.com/doi/full/10.1080/02568543.2023.2248223>

Rohacek, M., & Isaacs, J. (2016). PFS + ECE: Outcomes measurement and pricing: Pay for success early childhood education toolkit report #3. Urban Institute  
<https://www.urban.org/sites/default/files/publication/82801/2000869-PFS-ECE-Outcomes-Measurement-and-Pricing.pdf>

UNESCO. (2017). *Overview: MELQO: Measuring early learning quality and outcomes*. UNESCO Publishing. <https://unesdoc.unesco.org/ark:/48223/pf0000248053>

Zuilkowski, S. S., McCoy, D. C., Serpell, R., Matafwali, B., & Fink, G. (2016). Dimensionality and the development of cognitive assessments for children in Sub-Saharan Africa. *Journal of Cross-Cultural Psychology*, 47(3), 341–354. <https://doi.org/10.1177/0022022115624155>

## Appendix 2. Measurement Tool Descriptions

### Ages & Stages Questionnaire (ASQ)

A questionnaire to be completed by parents for the developmental and social-emotional screening of children up to the age of six. Taking approximately 10-15 minutes to complete, the questionnaire can be used to celebrate milestones and identify developmental delays early (Ages & Stages Questionnaire, n.d.-a; n.d.-b).

### Being a Parent Scale

Assesses parents' sense of competency by measuring three factors: parenting satisfaction, efficacy and interest (Lindsay & Totsika, 2017). The measure consists of 17 items against which parents are asked to indicate their agreement or disagreement (Flynn et al., 2022).

### Brief Early Childhood Quality Inventory (BEQI)

BEQI is applied across different early childhood settings, including home-based, centre-based and community-based programmes. It has two components: the BEQI observation tool and the BEQI educator questionnaire. The observational tool, administered by trained enumerators, assesses children aged 0-5 over the course of 90 minutes. It measures a variety of areas of evidence-based practices and early childhood setting characteristics: learning through conversation, fostering child autonomy, stimulating materials and space, facilitating learning through play, promoting strong relationships, safe and healthy environment. The BEQI educator questionnaire, either self-administered or conducted as a 10-15 interview, is used to capture background information on educators, their knowledge of evidence-based practices, and perceptions of strengths and areas for improvement (ECD Measure, n.d.).

### Brigance Inventory of Early Development III Standardized (Brigance IED III)

The Brigance IED III Standardized is a selection of 55 assessments for comparing a child's performance to a nationally representative sample (SpEd, n.d.). It scores children against five domains: physical development, language development, cognitive development, adaptive behaviour and social-emotional development, and is used for standardised reporting and benchmarking for children aged 0-7 (Brigance Early Childhood, n.d.).

### Child Behaviour Traits (CBT)

A 20-item questionnaire that a home visitor uses to capture a child's behaviour. It measures five factors: independence, social cooperation, task orientation, cognitive ability, and emotional stability (Flynn et al., 2022).

### Early Years Foundation Stage (EYFS)

The EYFS sets standards for children's learning, development and care from birth to five years old. Using observational assessments, children's progress is reviewed twice. First, between the ages of two and three years, assessed by an early years practitioner or health visitor and second, class teachers assess progress at the end of the school year when a child turns five (GOV.UK, n.d.).

### Early Learning Outcomes Measure (ELOM)

Developed and validated in South Africa for children aged four to six years, ELOM establishes whether an early childhood development programme is effective in preparing children for

school entry. It is currently available in nine official South African languages (Innovation Edge, n.d.).

#### Escala Abreviada de Desarrollo (EAD-3)

A Colombian instrument used to measure child development of children up to the age of seven aiming to result in early identification of developmental delays (Gobierno de Colombia, 2023). The tool measures gross motor development, adaptive-fine motor skills, hearing and language and personal-social development (UNICEF et al., 2022).

#### Instrumento de Medición de la Calidad de la Educación Inicial en Colombia (IMCEIC)

IMCEIC measures the quality level of ECD centres, ensuring that they align with the national context. It uses an observational guide, undertaken by external assessors, to assess structural and process quality as well as a teacher survey administered by trained data collectors (Maldonado-Carreño et al., 2022). The observation guide includes 45 items to assess safety, resources and learning materials, quality of teacher-child interactions and developmental activities (Maldonado-Carreño et al., 2022).

#### International Development and Early Learning Assessment (IDELA)

Produced by Save the Children, IDELA is a global tool to measure the early learning and development of children aged three and a half to six years. It measures four core developmental domains: motor, emergent literacy, emergent numeracy, and socio-emotional development (IDELA, n.d.).

#### Istation

Offering reading, maths and Spanish literacy assessments for schools, Istation is used to inform decision-making, implement interventions and monitor student progress (Istation, n.d.).

#### Kindergarten Development Check (KDC)

A standardised state-wide diagnostic tool assessing kindergarten students against 21 developmental markers across three areas (thinking and problem-solving, literacy and numeracy, and health and wellbeing) (Moore & Arefadib, 2022).

#### Parent and Child Together (PACT)

A 20-item questionnaire that measures child-parent interactions, encompassing four areas: communication, affection, consistency and responsiveness (Flynn et al., 2022).

#### Partnership for Assessment of Readiness for College and Careers (PARCC)

An Illinois state-wide exam to assess English language and mathematics. Administered between grades three to 11 in public schools, PARCC intends to provide evidence to determine whether students are on track for college and to provide data to help inform classroom instruction (Illinois State Board of Education, n.d.; PARCC, 2018).

#### Peabody Picture Vocabulary Test (PPVT)

An assessment for evaluating receptive vocabulary (Pearson, n.d.). To do so, an examiner says a word and the examinee selects one of four pictures that best illustrates the word's meaning. The test is administered to people aged two and a half up to 90+ years (Eigsti, 2021). It is an

untimed assessment that requires no reading, writing or expressive verbal language (Eigsti, 2021).

### Teaching Strategies Gold Assessment

An observation-based assessment for measuring the knowledge, skills and behaviour most predictive of school success. It is used from birth through kindergarten and is organised around various areas of development: socio-emotional, physical, language, cognitive, literacy, mathematics, science and technology and arts (Teaching Strategies, 2011).

## References

Ages & Stages Questionnaires. (n.d.-a). *How ASQ works*. <https://agesandstages.com/about-asq/how-asq-works/>

Ages & Stages Questionnaires. (n.d.-b). *Why screening matters*. <https://agesandstages.com/about-asq/why-screening-matters/>

Brigance Early Childhood. (n.d.). *Product sampler*. <https://www.curriculumassociates.com/-/media/mainsite/files/brigance/brigance-early-childhood-sampler-brochure.pdf>

ECD Measure. (n.d.). *Brief Early Childhood Quality Inventory (BEQI) Together in Practice Toolkit*. <https://www.ecdmeasure.org/beqi>

Eigsti, I. M. (2021). Peabody picture vocabulary test. In F.R. Volkmar (Ed.), *Encyclopedia of Autism Spectrum Disorders* (pp. 3357-3360). Springer International Publishing. [https://link.springer.com/referenceworkentry/10.1007/978-1-4419-1698-3\\_531](https://link.springer.com/referenceworkentry/10.1007/978-1-4419-1698-3_531)

Flynn, S., Sutherland, D., & Centre for Educational Development, Appraisal and Research (CEDAR), University of Warwick. (2022). *Evaluation of Parent Child+ in England, 2019-2025*. [https://golab.bsg.ox.ac.uk/documents/PC\\_Warwick\\_Interim\\_Report\\_-\\_Final.pdf](https://golab.bsg.ox.ac.uk/documents/PC_Warwick_Interim_Report_-_Final.pdf)

Gobierno de Colombia. (2023). *Escala Abreviada de Desarrollo -3*. <https://www.slideshare.net/slideshow/escalaabreviadadedesarrollo3pdf-255905086/255905086#1>

GOV.UK. (n.d.). *Early years foundation stage*. <https://www.gov.uk/early-years-foundation-stage>

IDELA. (n.d.). *About IDELA*. <https://idela-network.org/about/>

Illinois State Board of Education. (n.d.). *Assessment PARCC*. <https://www.isbe.net/parcc>

Innovation Edge. (n.d.). *Early learning outcomes measure (ELOM)*. <https://innovationedge.org.za/project/elom/>

Istation. (n.d.). *Digital Literacy and Math Assessment Solution*.  
<https://www.istation.com/assessment>

Lindsay, G., & Totsika, V. (2017). The effectiveness of universal parenting programmes: the CANparent trial. *BMC Psychology*, 5(35). <https://doi.org/10.1186/s40359-017-0204-1>

Maldonado-Carreño, C., Yoshikawa, H., Escallón, E., Ponguta, L. A., Nieto, A. M., Kagan, S. L., Rey-Guerra, C., Cristancho, J. C., Mateus, A., Caro, L. A., Aragon, C. A., Rodríguez, A. M., & Motta, A. (2022). Measuring the quality of early childhood education: Associations with children's development from a national study with the IMCEIC tool in Colombia. *Child Development*, 93, 254–268. <https://doi.org/10.1111/cdev.13665>

Moore, T., & Arefadib, N. (2022). *Tasmanian Play2Learn+ Trial: Evidence Review*. Centre for Community Child Health, Murdoch Children's Research Institute.  
[https://www.rch.org.au/uploadedFiles/Main/Content/ccchdev/2212\\_CCCH\\_TAS-Play2Learn.pdf](https://www.rch.org.au/uploadedFiles/Main/Content/ccchdev/2212_CCCH_TAS-Play2Learn.pdf)

PARCC. (2018). *Final Technical Report for 2017 Administration*.  
<https://www.isbe.net/Documents/PARCC-2017-Technical-Report.pdf>

Pearson. (n.d.). *Peabody Picture Vocabulary Test. Fourth edition*.  
<https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Academic-Learning/Peabody-Picture-Vocabulary-Test-%7C-Fourth-Edition/p/100000501.html?tab=overview>

SpEd. (n.d.). *BRIGANCE Inventory of Early Development III (IED III)*.  
[https://www.spedsg.com/assessment\\_tools/Brigance/IEDIII.htm](https://www.spedsg.com/assessment_tools/Brigance/IEDIII.htm)

Teaching Strategies. (2011). *Teaching Strategies GOLD assessment system. Technical summary*. <https://teachingstrategies.com/wp-content/uploads/2017/03/GOLD-Tech-Summary-8-18-2011.pdf>

UNICEF, Bienestar Familiar, & Instiglio. (2022). *Results-based financing (RBF) mechanism: An opportunity to improve results in early childhood development (ECD) in Colombia*.



## Appendix 3. Case Study: The Utah High-Quality Preschool Pay for Success Program: Using a Measurement Tool to Identify High-Risk Children<sup>1</sup>

Launched in 2013, the Utah High-Quality Preschool Pay for Success (PFS) Program aimed to increase access to preschool for children from low-income families and raise public investment in ECCE. In this outcomes-based financing project, the state of Utah paid investors of the preschool project the principal invested plus a financial return if the children identified as potentially eligible for government-funded special education avoided receiving these services. The project used one targeted outcome against which outcome payments would be disbursed: A decrease in the use of special education between grades K-6. This was chosen to address the challenge of children from low-income households being more likely to require special education services than their peers of higher-income backgrounds. The choice was also supported by the availability of cost data, which showed how much special education services cost the public system and how much money avoiding special education would save for the state of Utah.

### **Identifying the target population using the Peabody Picture Vocabulary Test**

In contrast to many other OBF4ECCE projects, the Utah High-Quality Preschool Pay for Success (PFS) Program used a measurement tool to identify the children whose outcomes would be tied to outcome payments. The project used the Peabody Picture Vocabulary Test (PPVT) to identify low-income preschool-age children at high risk of later receiving special education services. The PPVT assesses receptive language skills which correlate well with other child development and school readiness measures, and—most importantly—predicts the likelihood of later needing special education services. At the time, Utah did not have a standard school readiness measure in use. Had there been such a measure, it is possible that it would have been used to identify the high-risk group. Lacking an existing standardised tool, the PPVT was considered the best option for the project's purpose given its large research base, including longitudinal studies, existing use among the target population and established predictive value for special education.

The project designers defined preschool children at high risk of receiving special education services in primary school as those who scored at or below two standard deviations of the mean on the PPVT. The analysis of existing research on PPVT by the project designers revealed that 20-25% of children in the targeted low-income population scored below two standard deviations compared to only 3% of children in the general population. Based on data available at the time, it was predicted that these children were highly likely to experience academic failure and be placed in special education services in primary and secondary school.

### *Challenge of using PPVT*

The project designers were concerned about the use of the American English PPVT for children who did not speak English at home (primarily Spanish speakers). The project designers administered the Spanish and English language PPVT to a sample of these children.

---

<sup>1</sup> With thanks to Mark Innocenti of Utah State University for sharing his experiences and insights on measurement tools in this project.

Their scores were similarly low on the Spanish version as on the English version, so they believed it was valid to use the English version in the project among children who primarily speak Spanish at home.

*Mitigating perverse incentives using the PPVT*

The measurement tool and targeted outcome were designed to resist unintentional gaming and perverse incentives in the OBF approach. The enumerators who administered the PPVT to the children to identify the high-risk group had no stake in the repayment. Teachers and school administrators also had no stake in the repayment, as it did not affect their pay or funding levels. No one at the school, community, or school district level knew who the children in the high-risk group were, including the children and their families. The project researcher was the only person with access to the data, and any identifying information was removed from the data set. The high-risk children received the same preschool intervention, in the same classrooms, as all the other children from low-income households. The decision to place a child in special education services was based on numerous factors and actors, without any information on whether they were identified as high-risk through the PPVT in the PFS Program.

## Appendix 4. Case Study: Choosing Measurement Tools to Build Access and Quality in Early Childhood Development Centres in Sierra Leone<sup>2</sup>

In 2024, EOF, in partnership with the Government of Sierra Leone, started designing the Sierra Leone ECCE Outcomes Fund. The project aimed to expand access to high-quality early learning opportunities for children aged three to five years in remote areas of the country. By establishing community-based ECD centres, the outcomes fund aspired to ensure access and attendance, provide nutrition services and raise quality standards, ultimately improving child development outcomes. The project was designed to disburse outcome payments against verified successes associated with structural and process quality elements as well as child development outcomes. For this, the identification of measurement tools was required.

To determine which measurement tools to use in the project, EOF first reviewed internationally recognised tools previously used in low- and middle-income countries with an expert measurement advisory board. This resulted in a shortlist of validated tools that measure what they considered most important in child development and ECCE quality. Then, EOF considered tools already in use in the country to examine whether they would be suitable for an OBF project and compared them to the shortlisted tools, ensuring that selecting an international tool would be additional and complementary to the local context and national ECCE policies. For the Sierra Leone ECCE Outcomes Fund, EOF and the Government of Sierra Leone jointly agreed upon the following measurement tools, in collaboration with measurement experts.

### **Measuring process quality: Brief Early Childhood Quality Inventory (BEQI)**

In the absence of a government-produced tool for measuring process quality, the project planned to use BEQI as the tool to measure process quality. The tool involves a trained enumerator completing a 50-item checklist on a yes/no scale based on observations of classroom practices and classroom learning environments, covering play-based learning, learning through conversations, and promoting strong relationships. Project designers chose BEQI based on the ease of training enumerators, the mix of easy and harder-to-achieve items within the assessment tool, and the ability to adapt it to national standards and priorities.

### **Measuring structural quality: Minimum Standards Checklist**

For structural quality, the project planned to use the Government of Sierra Leone's Minimum Standards Checklist supplemented with specific structural quality elements from the international measurement tool, BEQI. The checklist covers aspects of basic sanitation, safety and hazards, class size and staff characteristics, and space and materials. It uses a four-tiered rating system, with specific requirements for each. Using a "building block" approach, where all standards from one level must be achieved before the next level can be met, the intention was to mitigate perverse incentives and help service providers know specifically what they need to do to improve.

---

<sup>2</sup> With thanks to Adriana Balducci of the Education Outcomes Fund for sharing her insights on measurement tools in this project.

### **Measuring child development outcomes: International Development and Early Learning Assessment (IDELA)**

The project planned to use IDELA to measure child development outcomes. IDELA has been previously applied in several West African countries and many LMICs, meaning there is a lot of existing data available to facilitate target-setting. The assessment contains 101 items, with 24 subtasks, covering four developmental domains: emergent numeracy, emergent literacy, gross and fine motor skills, and social-emotional learning. It is administered by trained enumerators using locally sourced materials.

EOF engaged in consultations for piloting IDELA in Sierra Leone to avoid floor (whereby the vast majority of participants answer the item incorrectly) and ceiling effects (the vast majority of participants answer an item correctly), which would yield little useful information about child development. Piloting is essential in building implementers' faith in the tool. Local ECCE educators and enumerators were engaged in deciding if any edits were required and if any items were too easy or too difficult. They would also ensure that the questions reflect the local context, an important factor because IDELA does not fully align with Sierra Leone's Early Learning Standards that serve as national quality benchmarks. Most of the content and domains of the Standards align with IDELA, but several were not included in the tool.

### **Reflections on selecting and adapting measurement tools to the local context**

The Sierra Leone experience highlights the importance of ensuring alignment between the measurement tool and the local context. Following are some of the reflections made by EOF based on the experience in Sierra Leone:

- It is essential to work with governments to align the measurement tools and processes with government policies, standards, and existing government tools.
- Conduct rigorous consultation, adaptation and validation processes to ensure measurement tools align with the local context and perverse incentives for gaming the measurement are mitigated.
- Utilising a robust and transparent process of selecting and validating measurement tools is crucial for building the confidence of implementing partners in the project and the achievements it could reach.
- Utilise existing performance data to set targets for the expected outcomes.
- Consider trade-offs between the validity of the measurement process with the time and financial resources available.