The LEGO Foundation

Playful Learning Across the Years (PLAY) Measurement Toolkit: Full Report

December 2022

Prepared by
RTI International:[1] Matthew Jukes, Kellie Betts, Margaret (Peggy) Dubeck, Lauren Edwards, Tabitha Nduku, Emily Staskowicz and Jonathan Stern.
NYU-TIES: Hiro Yoshikawa, Kalina Gjicali, Sharon Kim, Tahiya Mahbub, Maylee Montagut, Casey Moran, Krushika Uday Patankar, Sarah Rosenbach, Heba Mohammad Saleh, Kendra Strouf, Vicky Yiran Zhao.

[1] RTI International is a registered trademark and a trade name of Research Triangle Institute.

RTI INTERNATIONAL

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Playful learning has the potential to transform the global learning crisis. In infancy and early childhood, play builds a strong foundation for later learning by improving brain development and growth (Goldstein, 2012). In education systems that lack the capacity to support children effectively, playful learning brings its own powerful engine to drive learning – the joyful, engaged intrinsic motivation of children themselves (Zosh et al., 2017). In this way, play contributes to the holistic development of children, helping prepare them for the challenges of the current and future world. Evidence on playful learning is lacking, particularly in low- and middle-income countries. This is, in part, due to the lack of instruments to assess playful learning. This report describes the development of a toolkit for measuring support for playful learning in home, centre, and school settings across age groups from 0 to 12 years. We set up the rationale and aims of the project in **Section 1.**

In **Section 2**, we describe the conceptual framework guiding the development of the toolkit. We put forward the case that self-sustaining engagement in learning is the key construct underpinning the five characteristics of learning through play: joyful, meaningful, engaging, socially interactive, and iterative. Accordingly, we describe the toolkit as measuring 'how adults support children's self-sustaining engagement in learning, leading to a broad range of learning outcomes.' Based on a review of the literature, we propose that these adult-child interactions encompass six dimensions (we call these 'constructs'), described in **Table ES-1**.

Table ES-1. Constructs of support for children's engagement in learning

Construct	Definition
Support for agency	Adult support for children's ability to influence how and what they learn
Support for connection to experience	Teacher or caregiver support for children's learning that relates new information to already existing knowledge or to real-life experience
Support for exploration	Adult support for children's learning through manipulation, investigation, and acting on the physical or conceptual world
Support for problem solving	Adult support for children's efforts to achieve a learning goal for which they do not have an automatic solution
Support for social connectedness	Adult actions to strengthen, build on, or show the importance of social relationships in the class between teacher and student and among students themselves for the collective good
Positive emotional climate	An environment where interactions between adult(s), child(ren), and peers are warm, respectful, and positive

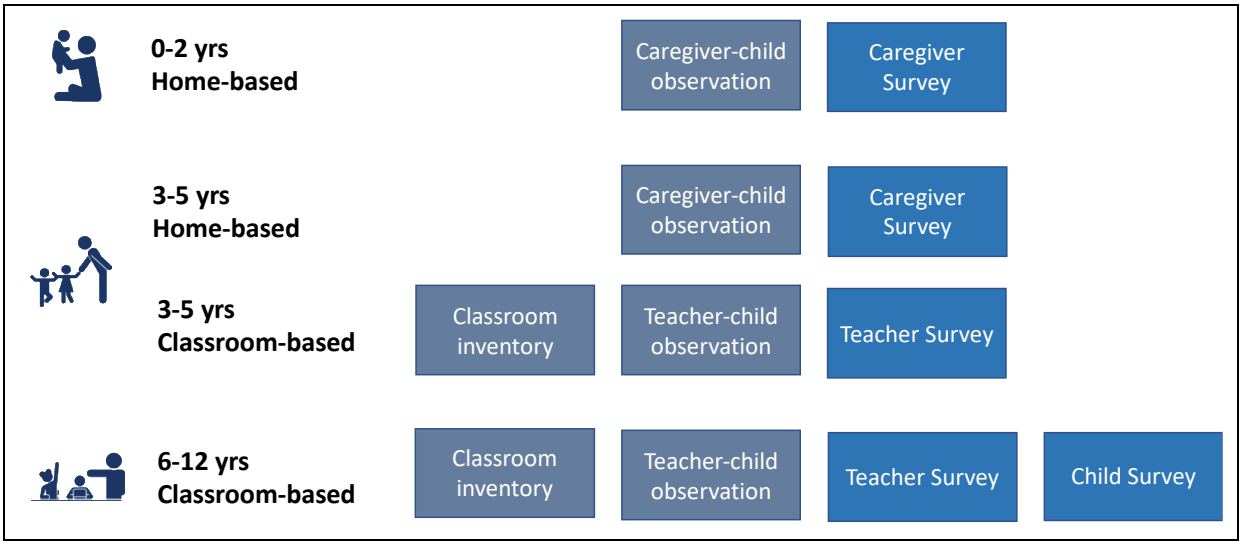
In **Section 3**, we describe our approach to adapting the toolkit across contexts. We aimed to make the toolkit applicable across cultures, whether Western, educated, industrialised, rich, and democratic (WEIRD) settings or rural agricultural communities in low- and middle-income countries. The toolkit was also designed to be used in a range of education systems – from strong education systems with well-trained and well-educated teachers to lower-performing education systems with lower-capacity teachers. Finally, the toolkit is also designed to be used in humanitarian and crisis settings. Section 3 describes the adaptations to the design of the toolkit based on these considerations. Key adaptations involve the inclusion of the social connectedness construct – which is particularly important for cultures where cooperative work towards communal goals is important – and the inclusion of items

sensitive to more subtle markers of support for agency, appropriate for low-capacity contexts and cultures where support for child agency is not the norm.

Based on these frameworks, we developed an initial set of tools for use in multiple age groups across different settings. The tools are not presented in detail as part of this report. Please refer to [link-to-toolkit] for the full toolkit. For the 0- to 2-year-old age group, the toolkit assesses support for children's engagement in the home, largely through interactions between the caregiver and child. In the 3- to 5-year-old age group, tools measure support for engagement in the home and the classroom. Tools for the 6- to 12-year-old age group focus only on the classroom.

Figure 1 shows the types of tools created. For each participant group, there is an observation tool and an interview with an adult (a teacher in classroom settings and a caregiver in home settings) to assess self-rated behaviour. For the 6–12 group, there is also an interview with students to assess self-rated behaviour. In both the 3–5 and 6–12 age groups, there is also a classroom inventory to assess physical aspects of the classroom, such as materials displayed on the walls, which might support self-sustaining engagement in learning. In most cases, the tools were designed to measure all six constructs of support for engagement in learning. One exception was the Caregiver-Child Observation Tool in the 0–2 group, which measured only four constructs; problem solving and connection to experience were not applicable to this age group. The two classroom inventory tools were not systematically designed around the six constructs.

Figure 1. Overview of tools



The PLAY Measurement toolkit is strongly evidence based, having been developed through several phases of data collection in four contexts. The 0–2 tools were adapted and tested in Colombia. The 3–5 tools were adapted and tested in both Jordan and Colombia. The 6–12 tools were adapted and tested in Kenya, Ghana, and Colombia. **Table ES-2** shows the stages of data collection involved in developing and validating the toolkit.

Table ES-2. Overview of tool development methods

Piloting phase		Methods
BUILD	• Local perception of playful learning and engagement • Extend and adapt core constructs • Generate and adapt items for context	Teacher-child interaction observations; teacher/caregiver focus groups; classroom naturalistic observations; point-of view observations and drawing focus groups with children
ADAPT	• Respondent and data collector understanding	Cognitive interviews and piloting with respondents to ascertain their understanding of and response to assessment items
	• Pilot and revise	Medium-scale pilot sample of all tools; data collection to inform replacement and revision of test items
TEST	• Psychometric assessment	Large-sample data collection as basis for factor analyses

In **Section 4,** we describe the methods of the Build phase. The aim of the Build phase was to document local understanding of key terms in order to build the constructs and items in the various tools. The methods focused particularly on the local perception of play, playful learning, and engagement in learning. We used several methods to this end. We conducted observations of teacher-child interactions and ran focus groups with teachers and caregivers. In addition, we conducted naturalistic observations in classrooms. Children were engaged in illuminative drawing exercises that served as the springboard for discussion in focus groups. We also conducted 'point-of-view' observations of children, where the observer records a general sense of what and who the child interacts with using a 'camera on head' approach.

In **Section 5,** we describe the findings of the Build phase. Overall, the qualitative data supported our hypothesis that children's engagement in learning is supported by adults and that the type of support provided is consistent with the six constructs proposed in our conceptual framework. Data led to the extension of some of the constructs. There was evidence that adults engage children in work-like play, which helped expand the characterisation of the 'support for agency' construct and the 'support for connection to experience' construct. The data also led to the expansion of the 'support for social connectedness' construct to include support for prosocial behaviour (such as sharing and friendliness), as well as building a sense of togetherness among children. This sense of togetherness was important in other constructs, too. For example, teachers said that problem solving was often most effective when peers collaborated in the process.

A key challenge to emerge from the data was the apparent contradiction between interview data and teacher practice relating to agency in the two African countries. Although caregivers, teachers, and children emphasised that children were motivated by autonomy in carrying out tasks, several teachers felt that explicit direction from teachers was required for students to be able to act independently. One implication of this finding is that the PLAY observation tools should be sensitive to relatively subtle expressions of child agency in the classroom.

The Adapt phase used quantitative data to adjust the tools before their psychometric properties were assessed in the Test phase. **Section 6** describes the quantitative methods used in these two phases. In all countries, cognitive interviews and small-scale pilots were conducted to test respondent understanding of items (for surveys) and data collector understanding (for observations), as well as face validity of items.

The psychometric properties of the tools were assessed in larger samples. The primary tool was assessed in 75 classrooms in Colombia and 70 schools with approximately 280 classrooms in Kenya and Ghana. The ECE observation tool was assessed in over 100 classrooms, and the caregiver tools were administered to around 150 caregiver-child dyads.

In addition to primary data collection, we conducted secondary data collection using videos of ECE classrooms recorded by Innovations for Poverty Action. Data collectors were trained on the PLAY toolkit and applied it to the video recordings of classroom interactions. In total, 423 classrooms were coded.

Section 7 describes how the conceptual framework and the toolkit evolved through the Build and Adapt phases to form the version of the toolkit that was assessed in the Test phase. This section describes how a literature review and Build phase work resulted in the addition of the 'social connectedness' construct to the conceptual framework. It then describes the development of the primary and ECE tools.

At the primary level, we experimented with adding three constructs: negative agency, negative emotional climate, and participation. We also tried different approaches to coding, applying quality metrics in the observation instrument and to administering the teacher and child surveys.

At the ECE level, we describe how items were developed through interviews with subject-matter experts and how the teacher and caregiver surveys were developed as well as the approach to training on ECE tools developed.

The next step in the process was the analysis of findings. These are extensive and are presented in **Appendix B.** A focus of analyses was on exploring the factor structure of the data from observation tools. This statistical approach identifies groups of items that tend to be observed in the same classrooms, which can then be compared to our hypothesised constructs. Our approach in each age group was the same. We identified a factor analysis model that best described the data in each country. Based on the combination of country-specific models, we developed model with reasonable fit across all contexts. This model formed the basis of our recommended final version of the tools. A similar process was conducted with the survey tools.

Section 8 describes how the analyses presented in Appendix B were used to make recommendations for the PLAY tools. The constructs making up each of the tools were largely derived from those identified in analyses (black crosses in **Table ES-3**). In addition, we propose the addition of some constructs (red crosses in Table ES-3) that were not identified by analyses but we argue would make an important addition to the toolkit. For example, the connection to experience construct was not identified in the 6–12 years observation tool but we include it because it was identified in the 3–5 years observation tool. Similarly, we include the agency construct in the 3–5 years observation tool because it was identified in the 6–12 years tool.

Table ES-3. Final recommended forms of the PLAY Measurement tools

Constructs	Classroom		Caregiver	
	6–12	3–5	0–2	3–5
Observation measures				
Connection to experience	x	x	N/A	
Problem solving			N/A	x
Exploration	x	x	x	

Constructs	Classroom		Caregiver	
	6–12	3–5	0–2	3–5
Observation measures				
Agency	x	x	x	x
Positive emotional climate				
Social connectedness	x	x	x	x

Constructs	Classroom			Caregiver	
	Adults 6–12	Children 6–12	Teachers 3–5	0–2	3–5
Survey measures					
Connection to experience	x		x	x	x
Problem solving			x		
Exploration	x		x		
Agency	x	x	x		
Positive emotional climate		x	x	x	x
Social connectedness	x		x	x	x

X = factors supported by analyses X = additional factors proposed for final tools

1. INTRODUCTION

'Factory' models of education – with an emphasis on solely didactic or rote instruction and memorisation – are still quite prevalent in education systems in much of the world, across low-, middle-, and high-income countries (Banerjee et al., 2017; Hirsh-Pasek et al., 2020; Mehta & Fine, 2019; Rogoff, 2003). Such approaches tend to build extrinsic motivation for school engagement and an emphasis on comparing individuals in performance, rather than intrinsic motivation and fluid collaboration (Rogoff, 2014). The focus on motivation and collaboration is characteristic of recent emphases on 21st-century skills in education, but also of community learning and collaboration norms in some indigenous communities in Central America as well as Mexican-origin children in the United States (Alcalá, Rogoff, & Fraire, 2018). Supporting deeper and broader learning – including application of knowledge to new situations; critical thinking; generation of innovations in problem solving; and the social interactions required to implement these in a variety of home, community, education, and work settings – has been noted to be rare in many large-scale education systems (Christie, 1985, 2008). Integrating culturally specific forms of learning in communities is also often lacking in large-scale education and early childhood systems (Dahlberg, Moss, & Pence, 1999; Ejuu, Apolot, & Serpell, 2019).

Three policy trends are implicated in this lack of attention to deeper and more culturally relevant learning. First, in low- and middle-income countries, and often in crisis or conflict-affected settings as well, the focus of education policy over the last 30 years has been on increasing access. Substantial increases in children's access to primary education occurred, for example, in low-income countries since the Education for All and Millennium Development Goals period. However, neither workforce, pedagogical, parent, and community engagement nor policy and governance systems aligned towards learning beyond access (Pritchett, 2013). In the development of the Sustainable Development Goals, further goal-setting towards universal pre-primary and secondary education occurred (United Nations, 2015). That said, there is scant evidence of large-scale improvement in learning outcomes since the establishment of these goals in 2015 (UNESCO Institute for Statistics, 2020).

Second, a predominant focus on standardised reading and math scores across countries in monitoring academic progress in primary, lower secondary, and upper secondary schooling constrains education systems to often narrowly defined domains of learning based on the worldviews of WEIRD (Western, educated, industrialised, rich, and democratic) countries (Henrich Heine, & Norenzayan, 2010). Domains such as social and emotional learning or skills related to intercultural and civic engagement have until recently been overlooked in many countries' national learning standards, across early childhood through adolescence. The PISA global learning assessment, beginning in 2018, started including a wider range of modules on global competencies related to civic, intercultural, social, and emotional learning and sustainability (OECD, 2018). However, these domains of outcomes have yet to be integrated into most countries' national education monitoring and assessment systems. In addition, national standardised assessment systems often overlook broader approaches to assessing language (e.g. multiple language proficiency and indigenous language proficiency) and quantitative or logical reasoning (Brunette et al., 2019; Piper, Zuilkowski, & Ong'ele, 2016). Assessments ignore the use of language or literacy in its social and cultural context. A broader approach to integrating valued cognitive abilities specific to societies, rather than imposed from elsewhere, can include dimensions ignored in national assessment systems (Nsamenang, 2006; Oppong, 2020). Currently, pedagogical approaches and

curricula are often disconnected from socialisation and learning practices that are important in specific communities and societies (Serpell, 2011; Wadende, Oburu, & Morara, 2016).

Finally, standards for home- and school-based programming often do not consider cultural variation within and across countries in the definitions and dimensions of quality perceived to lead to learning outcomes. For example, standards for early, primary, and secondary education often leave out observed social interactions and instruction in favour of the more easily regulable structural features of infrastructure, lesson plans, materials, teacher formal qualifications, or class size and adult-child ratio. In home-based interventions and parenting programmes, there has been a strong emphasis in the largely WEIRD impact evaluation literature on programmes that encourage parenting sensitivity, responsiveness, and stimulation defined in solely dyadic interactions (Mesman et al., 2017) – a set of parenting styles that a wider literature shows is not universal (Keller et al., 2018). The standards and goals of parenting and of education programmes often do not reflect what community members may value in terms of features of teaching, parenting support, or local socialisation (Apolot, Ejuu, & Lubaale, 2020).

The PLAY toolkit is designed to support initiatives aimed at addressing the challenges inherent in these three policy trends. The first challenge – the global learning crisis – could potentially be alleviated by a focus on learning through play. In infancy and early childhood, play builds a strong foundation for later learning by improving brain development and growth (Goldstein, 2012). In education systems that lack the capacity to support children effectively, playful learning brings its own powerful engine to drive learning – the joyful, engaged intrinsic motivation of children themselves (Zosh et al., 2017).

Second, the narrow focus on academic skills can potentially be contested by a focus on learning through play. Learning through play has been proposed as a concept and accompanying set of child activities that can lead to a broader and more meaningful set of learning outcomes than those assessed in many standardised assessment systems, with an emphasis on creativity, social engagement, and iterative and innovative problem solving from birth to adolescence (Hirsh-Pasek et al., 2020; Zosh et al., 2017). The PLAY toolkit thus builds evidence around a broader definition of skills.

In responding to these two challenges, the PLAY toolkit needs to be mindful of the third challenge – the lack of culturally relevant school- and home-based programming. To address this, the PLAY toolkit is designed to be applied in – and adapted to – a wide range of cultural contexts and different national policy environments. The tool recognises cultural variation in the interpretation of ‘learning through play’ and that there is variation in the capacity and orientation of education systems to support learning through play.

The role of the PLAY toolkit in addressing these challenges is to measure progress and provide evidence. There is a lack of rigorous, objective tools to measure support for learning through play, and, as a result, the evidence base on playful learning is weak, particularly in low- and middle-income countries. The development of the PLAY toolkit is the first step in strengthening the evidence base. This report describes how the toolkit – a set of culturally responsive instruments to measure support for playful learning in children aged 0–12 – was designed, developed, and tested.

2. CONCEPTUAL FRAMEWORK

2.1 Introduction and Development of the Conceptual Framework

For children, learning through play is experienced as being active, socially engaged, meaningful, iterative, and joyful or attentive (Zosh et al., 2017). In addition, a spectrum of play activities has been put forward as a set of adult practices that might foster these dimensions (Jensen et al., 2019). However, to date, relatively little attention has been paid to defining the dimensions of *observable interactions* or other features of the important activity settings of child and youth development that may be associated with learning through play.

We link dimensions of interactions to the five characteristics of learning through play (Zosh et al., 2017) through the concept of **self-sustaining engagement in learning**. We believe that sustained engagement in activities that constitute learning in home, centre, and school settings is a key underlying dimension of learning through play.

The purpose of this section is to outline a conceptual framework for the measurement of social interactions associated with self-sustaining engagement in learning within the activity settings of home, centre, school, and community across early childhood to early adolescence. We focus on the question, **What kinds of observable interactions across the play facilitation spectrum (and broader features, such as physical aspects or group composition) within homes, centres, and schools support the self-sustaining engagement that underlies learning through play?** This research question focuses primarily on one critical component of settings – interactions with the social or physical environment – in home, centre, and school activity settings, which represent principal contexts of socialisation between early childhood and early adolescence years (birth to age 12) in many societies and are therefore the source of important learning opportunities (Gallimore, Goldenberg, & Weisner, 1993).

We first discuss how we developed this conceptual framework. Then we present definitions of learning through play and how self-sustaining engagement may underlie them; describe six dimensions of observable social interactions (**Figure 2**); describe how these interactions may manifest in both early childhood and middle childhood/early adolescence; consider how they may be associated with the five characteristics of learning through play (Zosh et al., 2017); and finally discuss how other characteristics of activity settings may predict both the interactions and characteristics.

Before turning to the process of development of the framework, we note that learning through play may occur across a variety of unstructured and structured activities and across a variety of levels of peer and adult involvement and guidance. This has been referred to as the spectrum of play (Zosh et al., 2018). Along this spectrum, what is typical of adult or peer involvement in play, initiations of play episodes or actions by children or adults, and the content of play may also vary significantly across cultures (Haight et al., 1999; Lancy 2007; Morelli, Rogoff, & Angelillo, 2003). We cover the entire play facilitation spectrum, which spans completely solitary free play on one end to completely didactic instruction on the other (Jensen et al., 2019; Zosh et al., 2018).

To date, interactions across the play facilitation spectrum have not been distinguished in ways that could inform practice – either in home-based or educational interventions. We consider our dimensions of interactions (the left column in Figure 2) on the one hand as related to the five characteristics of learning through play, but on the other as potential dimensions to include in observational measures of quality – of home-, centre-, community-, and school-based interventions. Some of the dimensions of social interactions that we

propose overlap with constructs used in current observational quality measures (e.g. positive emotional climate in the Teacher Instructional Practices and Process System (TIPPS), the Measure Early Learning Environments (MELE), and the TEACH or Teach Early Childhood Education tools) (Davis & Raikes, 2020; Plata, Hein, & Ponguta, 2020; Wolf et al., 2018; Raikes et al., 2019; Seidman et al., 2018; World Bank, 2019, 2021). We discuss these overlapping instances below. However, most of our dimensions are distinguished from or more specific than these more general constructs and may supplement them for purposes of research, monitoring, or evaluation aimed at facilitating learning through play in local cultural contexts.

Finally, we take a developmental perspective to propose dimensions of interactions within activity settings in two broad age periods – early childhood (birth to roughly age 5) and middle childhood to early adolescence (ages 6–12). These encompass multiple periods of heightened sensitivity of development to environmental influence and, in most countries, the transitions to pre-primary, primary, and lower secondary education.

2.1.1 Process of Development of the Conceptual Framework

The five characteristics of learning through play (Zosh et al., 2017) provided a starting point for identifying constructs. However, our conclusion after reviewing the literature was that the five characteristics may be observable in children's behaviour but may be harder to observe as dimensions of adult support for child behaviour. For example, we may not be able to reliably assess teachers as scoring highly on a scale of 'support for iteration.' Instead, through our review of relevant literature as well as input from the advisory board, we found six areas of adult support for children's engagement – agency, problem solving, exploration, connection to experience, and emotional climate. In addition,

The list of constructs was further revised after a review of literature on the role of culture in human development and education and by data collected through the Build phases of the project, particularly in Kenya and Ghana. This resulted in the addition of social connectedness as a construct. In cultures where social connectedness is valued above psychological autonomy, individuals are motivated by a sense of belonging, by working together towards a common goal, and when given individual responsibility for a task that supports the group. We hypothesised that students would be more engaged in the classroom through these factors.

As we further began operationalising the framework into the caregiver, teacher, and child surveys, as well as the observational tools, we found that not all the constructs were relevant for each tool. For example, some constructs were less suited to the assessment for caregiver-child interactions with 0- to 2-year-olds – that is, the constructs of support for problem solving and connection to experience were less relevant for this age range so were not included in that measure.

Further revisions to the conceptual framework occurred through our data collection and analysis in the Build and Adapt phases. Build data suggested, for example, that true agency – where children make decisions about their learning – was relatively uncommon. Teachers said that children were engaged when teachers gave explicit instruction, because children liked this way of teaching. Although this approach may restrict children's agency in the short term, it was seen by teachers as preparing students to do things on their own in the future. In response to these findings, and based on the work of Jennifer Adair (e.g. Adair and Colegrove, 2014), we included items that captured more limited versions of agency – for example, where teachers do not reprimand students for moving around the classroom. Adapt data in Colombia showed higher levels of agency than in the two African countries. As

a result of these findings across three countries, we retained items capturing both low and high levels of agency.

Based on Adapt phase data in Ghana and feedback from an implementing partner (Play Matters/International Rescue Committee) who tested the tool in Tanzania, we added the construct of 'participation' to the primary school tools. The aim was to capture children's involvement in their own learning – e.g. by practising new skills and being given different ways to respond to questions – that did not qualify as agentic decision-making.

In support for problem solving, the Build data helped us add perspectives such as positive criticism, error-correction, and evaluating children's work were additions from Kenya. For the construct of support for exploration, the importance of fostering a sense of responsibility through home-based activities were found to be important in Colombia.

Build data generated many examples of teachers restricting children's agency and creating a negative emotional climate. Psychometric literature suggests that negatively worded items should constitute separate constructs from positively worded items (Sonderen et al., 2013), so we created new constructs of 'negative agency' and 'negative emotional climate' for the Ghana Adapt phase of the primary tool. However, these negative items were not observed frequently during the Ghana Adapt phase. As a result, we reduced the items, rephrased them with positive wording, and ensured that they were included in the examples listed under the main constructs of agency and positive emotional climate. For example, the item 'teacher limits freedom in student movement and interaction' under negative agency was reworded and incorporated into the item 'teacher allows freedom in student movement' under the agency construct.

Some constructs were less suited to the assessment for caregiver-child interactions with 0- to 2-year-olds – that is, the constructs of support for problem solving and connection to experience were less relevant for this age range so were not included in that measure.

Finally, through our analysis of the factor structure and psychometrics of our observational assessments of centres and classrooms, we found several patterns of note that led to a different balance and representation of constructs in our final measures at the end of this project (**Table 1**). First, items tapping agency were retained in the factor structure of the primary tool, but concerning independence rather than agency in decision-making. In addition, items originally hypothesised to represent agency showed up in both tools as part of the support for exploration and problem-solving construct (as well as scattered among the other constructs, such as connection to experience). Second, support for exploration and problem solving tended to appear together in a combined factor, across both pre-primary and primary assessments. Third, a factor in the area of social connectedness featured in both sets of analyses, but with relatively different emphases. In the pre-primary tool, social connectedness was represented by teachers supporting developing friendships, listening among the children, and social norms of cordiality and affection. In the primary tool, social connectedness was represented by teachers supporting an overall classroom sense of camaraderie and togetherness. This aligns with the developmental tasks of early childhood focusing more on initial friendship formation skills, while for middle childhood and early adolescence there may be more opportunities to develop classroom group identity and togetherness for the class as a whole.

Table 1. Constructs supported by factor analysis across settings for the pre-primary and primary classroom observation tools

Constructs	Classroom observation tool	
	Pre-primary	Primary
Support for agency		X
Support for exploration and problem solving	X	X
Support for connection to experience	X	
Support for social connectedness	X	X
Emotional climate		

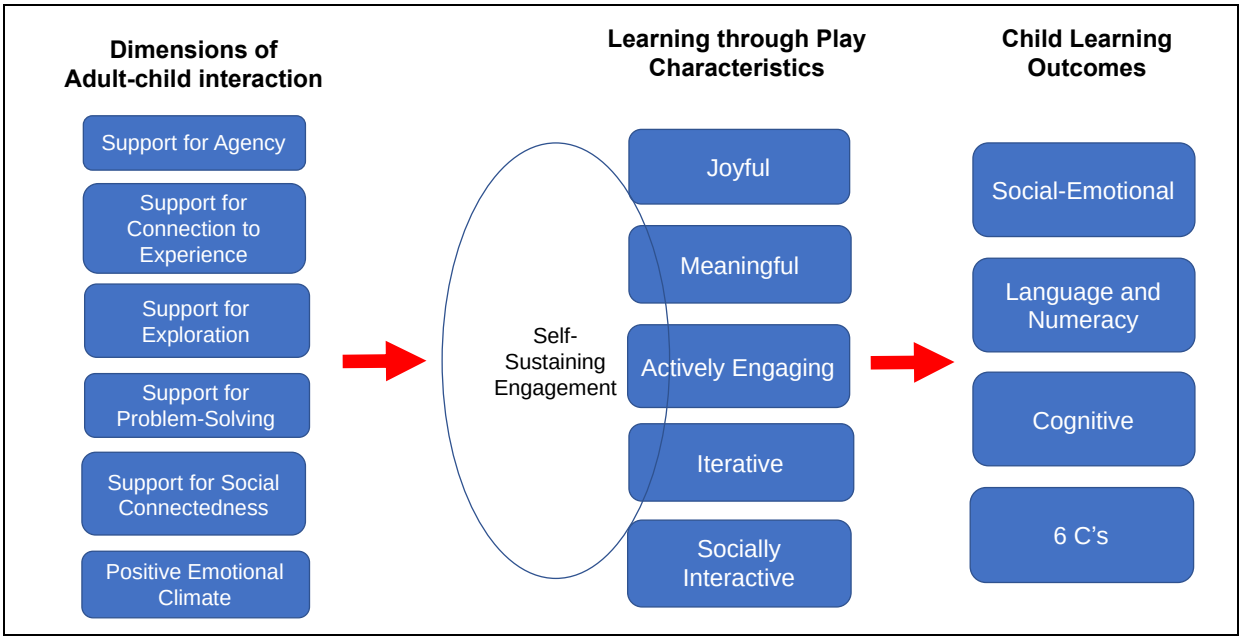
Findings from our Caregiver-Child Observation Tools similarly led to some revisions to our constructs. For the 0–2 age range tools, the original four constructs of support for agency, exploration, social connectedness, and positive emotional climate were revised to three, with social connectedness and positive climate forming a single factor. For the 3–5 age range tool, we included all six constructs; however, we found evidence for only four of the six (agency, problem solving, social connectedness, and positive emotional climate).

How do our obtained factors and revised constructs supplement constructs in the more general observed quality literature? Strikingly, the more general aspects of emotional climate such as teacher positive affect, praise, and behaviour management did not cohere into separate factors in the classroom tools, despite our including them in our instruments. Our construct of social connectedness focuses more on the dyadic and classroom friendships and togetherness that may foster rich experiences of engagement and learning through play. We have a more elaborated construct for support of connection to experience, which includes personal, community and cultural experiences (most existing observational tools have a single item for this construct). And our construct of support for agency is a new contribution to the field, building on recent work by Adair and colleagues (Adair & Sanchez-Colegrove, 2021).

Our caregiver tools extend emphases in existing tools measuring responsive parenting to include support for exploration (0–2) and support for problem solving (3–5). However, we encourage further development of family-based observations of adult-child interactions to include the complex multi-adult, peer, and group interactions that are typical of children's participation in play across cultures (Gaskins, Haight, & Lancy, 2007).

Our observational tools were meant to predict self-sustaining engagement as a blend of motivational and attentional processes that underlie learning through play. Therefore, we place relatively less emphasis on 'time on task' or instructional quality, and exclude topical instruction (e.g. quality of reading, math, or science instruction). Rather, we intend to predict self-sustaining engagement no matter what the activity or instructional content is of a particular curriculum. As a result, our measure can also supplement specific curriculum-implementation measures that assess quality of content instruction.

Figure 2. Conceptual model linking social interaction dimensions to self-sustaining engagement and learning through play



6 Cs: Collaboration, Communication, Content, Critical thinking, Creative innovation, and Confidence

2.2 Defining and Characterising Self-Sustaining Engagement and How It Underlies Learning through Play

Learning through play takes place when a play activity engages activity characteristics that are also known to lead to learning. Based on a review of research in the science of learning, Hirsh-Pasek et al. (2015) identify four characteristics of learning: active ('minds-on'), engaged, socially interactive, and meaningful learning. Zosh et al. (2017) adds the concept of iterative learning to these characteristics. We assume that the characteristic of 'minds-on' learning was subsumed within the concept of 'engaged' learning, but we retain it in our framework because it has potential relevance to learning in non-WEIRD settings. In such settings, children often learn through observation and participation (Rogoff, 2003; Rogoff et al., 2014). At the same time, we observe in many classrooms in low- and middle-income countries that children are behaviourally compliant (e.g. sitting still and facing the teacher) but may not be cognitively engaged with material presented. In such situations, a focus on the concept of 'minds-on' – distinct from 'engagement,' which refers to sustained attention – could distinguish between a child who is learning through observation and one who is simply present and behaviourally compliant. The fifth characteristic of learning through play is 'joyful.' We interpret this characteristic as primarily signalling the child's subjective experience of an activity as playful.

What factors may underlie learning through play and, indeed, broader learning? Voluntary activity, intrinsic motivation, and a state of 'flow' have all been linked in relevant studies to learning. We propose the concept of self-sustaining engagement as uniting these concepts. Voluntary engagement has a clear relationship to each of the five characteristics of learning through play. Being 'minds-on' and engaged in an activity is perhaps the definition of voluntary mental engagement in a task. If a child is voluntarily engaged, they are likely to have greater focus, attention and absorption in learning.

Learning through play is also characterised by intrinsic motivation. Self-determination theory (Deci et al., 1991) posits that children are motivated to engage in a task when their needs for competence, autonomy, and relatedness are met. Finally, a state of 'flow' or sustained

absorption in an activity across time has been linked to deeper learning. Work on Csikszentmihalyi's influential theory has shown that a state of flow in an activity is characterised by an optimal balance of challenge and existing skill as well as effortless focus and attention in the activity, with passage of time experienced as occurring quickly (Csikszentmihalyi, 1997).

Our concept of self-sustaining engagement encompasses these three features of voluntary, intrinsically motivated, and having some of the characteristics of 'flow' experiences. With self-sustaining engagement, play activities are more likely to lead to learning in that behavioural (sustained involvement), cognitive (focused involvement), and affective (positive experience of the activity) dimensions of engagement are all present. These three types of engagement have each been linked to learning and achievement in schooling contexts (Wang & Eccles, 2012) and are also present in the concept of 'minds-on' engagement (Zosh et al., 2017).

2.3 Defining and Characterising Dimensions of Social Interaction That Support Self-Sustaining Engagement and Learning through Play

2.3.1 Home, Centre, and School Activity Settings as Central Contexts for Social Interactions That Support Self-Sustaining Engagement and Learning through Play

Homes, centres, and schools are contexts in which children, in many societies, spend the bulk of their time in the first decade of life. In ecological theories, these settings have been termed 'microsystems,' or bounded contexts in which patterned behavioural and social interactions occur with some regularity between children, on the one hand, and adults or peers, on the other (Bronfenbrenner & Morris, 2006; Seidman, 2012). These social interactions and the microsystems in which they are embedded can vary in their support of learning through play. Microsystems in Bronfenbrenner's well-known ecological theory are bounded in time and space; involve regularity in interactions between the developing child and the social, physical, and symbolic aspects of the bounded setting; and influence the child through exposure and involvement over time (i.e. are not limited to one-time events). Home, centre, school, and community environments meet these criteria.

Cultural models of human development conceptualise the setting and the interactions within it with different emphases. Super and Harkness's (1986) notion of the developmental niche encompasses both the physical and social settings of everyday life, the socialisation practices of adults vis-à-vis children, and community-specific cognition and emotion as the central conveyors of cultural beliefs and practices across generations. Gallimore et al. (1993) further extend this notion to the combinations of activities in settings that make up family and community routines of socialisation. Indeed, some scholars conceive of 'activity settings' embedded in family, community, or school contexts as the primary unit of analysis for studying learning environments (O'Donnell, Tharp, & Wilson, 1993). And Eccles (2005) calls for more research on activity settings as engines of human development. This conceptualisation of activity settings in the microsystems of family, school, and community is broad enough to encompass the more child-directed forms of socialisation and play that are characteristic of guided play in Western, rich-country settings, as well as a broad array of child interactions with and involvement in adult and community routines that are less focused on adults directly monitoring children's independence and autonomous activity and more focused on children's autonomous behaviours supporting collective goals. The former has been termed relatedness in the service of autonomy, and the latter autonomy in the service of relatedness (Gaskins, 2015; Keller et al., 2004; Tamis-LeMonda et al., 2008). This

conceptualisation is also broad enough to encompass both cultures in which adults (specifically parents and childcare providers or teachers) are the main caretakers of children, and those in which multiple caretakers beyond the biological parents – siblings, other relatives, and an array of community members – are central to socialisation (Gottlieb, 2004; Weisner, 1996) and may engage more frequently in play activities with children (Roopnarine & Davidson, 2015).

The types of interactions within activity settings that are supported by caregivers vary. While in WEIRD contexts mothers and other adults often act as playmates during early childhood, this occurs less frequently in many other societies (Morelli, Rogoff, & Angelillo, 2003). Among some Mayan communities in Guatemala, for instance, adults rarely play with young children (Bazyk et al., 2003). Concomitantly, both within and outside of play contexts, caregiver sensitivity – or attention and appropriate response to infant cues – may vary both in terms of frequency of occurrence, appearance (e.g. verbal/nonverbal, with/without positive affect), and importance for child well-being and development (Keller et al., 2018; Weisner, 2014). In formal education settings, similarly, the types of pedagogy or instruction vary for both sociocultural and policy reasons. Supporting play, for example, is relatively more prominent in the formal national curricular and pedagogical bases of early childhood education in Colombia and Peru, for example, than in some other countries, such as the United States (Office of Head Start, 2020; Ministerio de Educación Nacional, 2017; Ministerio de Educación de Perú, 2019).

2.3.2 Developmental Changes in Homes, Centres, and Schools as Activity Settings across Early Childhood to Adolescence

We know that the transition from early childhood to middle childhood involves changes in developmental capacities (e.g. self-regulation, executive function, complex reasoning, moral reasoning, perspective-taking, and mentalisation). These are accompanied by changes in social roles and expectations (e.g. the nature of involvement in adult and community activities and expectations in schooling engagement and activities). Therefore, changes in developmental contexts and activity settings in which children spend time occur across these developmental stages, in type, quality, and interactions (Bronfenbrenner & Morris, 2006; Garcia Coll & Szalacha, 2004; Gottlieb, 1991). In many but not all cultural contexts, children have more agency to select the activity settings in which they engage as they proceed through early childhood and middle childhood. For example, pre-primary school settings often (though not always) have a greater variety of learning materials and spaces than childcare centres or home-based care settings for infants and toddlers, due to the greater motor skills and imaginative play, language, and cognitive capacities of preschool-aged children.

Entry into primary schooling increases substantially the assessment and comparison of children on content-based knowledge, abstract reasoning, writing, mathematics, and (potentially) broader holistic skills. However, the emphasis on rote instruction in many education systems often restricts spontaneity and flexibility of learning and the fostering of creative, voluntary learning through play. As part of our discussion below, we highlight the educational approaches – including project-based learning, collaborative learning, arts-based learning, and other modalities of education that facilitate creativity and out-of-school models that build on youths' intrinsic motivation to pursue their own interests – that may foster self-sustaining engagement and learning through play in primary and lower secondary schooling.

2.3.3 Dimensions of Interaction in Homes, Enters, and Schools That May Support Self-Sustaining Engagement and Learning through Play

We conceptualise six dimensions of adult-child interaction associated with self-sustained engagement and learning through play (see **Table 2**). Below, we indicate why each of these dimensions is related to learning through play. We base our dimensions on the literatures on observed quality in educational settings across countries and on the literature on culture, socialisation, and human development. In keeping with our framing in activity settings and ecological models of human development, we conceptualise social interactions as occurring at least at the dyadic level, but also including a variety of classroom or group interactions (e.g. teacher-child or facilitator-child interactions). Interactions with the physical environment, on the other hand, may or may not involve social interactions and thus may be solitary. We begin each discussion with a focus on early childhood and then extend to middle childhood and early adolescence.

Table 2. Constructs of support for children's engagement in learning

Construct	Definition
Support for agency	Adult support for children's ability to influence how and what they learn
Support for connection to experience	Teacher or caregiver support for children's learning that relates new information to already existing knowledge or to real-life experience
Support for exploration	Adult support for children's learning through manipulation, investigation, and acting on the physical or conceptual world
Support for problem solving	Adult support for children's efforts to achieve a learning goal for which they do not have an automatic solution
Support for social connectedness	Adult actions to strengthen, build on, or show the importance of social relationships in the class between teacher and student and among students themselves for the collective good
Positive emotional climate	An environment where interactions between adult(s), child(ren), and peers are warm, respectful, and positive

Support for agency

Agency refers to the extent to which children are able to 'exert their thinking and actions in a social context in which others hold the same rights' (Zosh et al., 2017). Adair and colleagues define it as being able to influence and make decisions about what and how something is learned in order to expand capabilities (Adair, 2014; Adair & Sánchez-Suzuki Colegrove, 2021).

Children's level of agency is high at one end of the spectrum – free play – and may be lower in direct instruction, where children often have limited agency (for example, to initiate an interaction with a peer to share something they are learning, or to get up, move around, or even gesture freely). Having the choice of an activity or material, or being able to choose peers with whom to interact in an activity, are all examples of child agency in centre or school settings. In settings that are more restricted, such as large pre-primary or primary classrooms, agency may be reflected in the degree of freedom of movement, even in a fairly constrained task such as lining up to write on the blackboard. In early caregiving, in addition to child choice in material or movement, support for agency may be reflected in following the child's lead, rather than the child following the caregiver's lead in an interaction or activity.

Support for agency may be restricted in ways that reflect equity or inequity in classrooms. Adair and colleagues note that in the United States, Black and brown children have substantially lower levels of agency in early childhood settings than white children, observing that learning activities with greater child agency, like free play, are more likely to be 'rewards' for sitting still and being silent in other learning activities.

Support for exploration

Support for exploration in the first 12 years of life can foster aspects of self-sustaining engagement such as voluntary and intrinsically motivated interaction with the physical and social environment. Exploration can take different forms depending on the characteristics of the environment the child has access to, how the child's exploration of that environment is facilitated, and the developmental stage of the child. The environment in early childhood includes, for example, whether the child is mostly indoors or outdoors, whether the child is with immediate family or exposed to a wide range of people and surroundings of the household, or as an infant whether the child is on the floor with room to move or being held by caregivers. The physical environment also includes the materials that the child has access to, such as natural objects or everyday household objects.

In addition to providing children with access to a variety of interesting learning environments, the way in which their exploration is facilitated is also critical. As a caregiver, support for exploration could include encouraging children to explore an object in multiple ways. In childcare or preschool environments, support for exploration builds on similar principles. Children's exploration of materials and phenomena in nature – such as sand, water, or plants – can be encouraged in ways that bring about self-sustaining engagement (e.g. by asking questions about the properties or characteristics of a material like sand when wet versus dry). Physical environments can be designed to encourage exploration (e.g. playful learning landscapes that have been integrated into public community space; see Schlesinger & Hirsh-Pasek, 2019).

Support for exploration in the primary schooling context can acknowledge the integration of developmentally specific competencies of middle childhood and early adolescence (Del Giudice, 2014). Teachers may encourage children to extend their understanding of a concept by approaching it using different methods or perspectives. They may also encourage students' independent exploration of a concept or learning material, either solo or in peer groups, in ways that can build self-sustaining engagement.

Support for connection to experience

Support for exploration may be enriched when the new experience is connected to prior experiences. For example, in early education linking exploration with children's home or community experiences may foster deeper learning, including the application of ideas across contexts. An observational item tapping whether new learning was connected to children's prior knowledge and experience constituted part of a central pedagogical quality factor that predicted executive function and emergent language skills in a national study of preschools in Colombia (Maldonado et al., 2021). In a study of Ghanaian preschools, this item formed part of a factor on supporting student expression and predicted growth in early literacy and social-emotional skills (Wolf et al., 2018). An observational item of connecting learning to everyday experiences of students predicted science achievement in secondary schools in Uganda (Seidman et al., 2018). However, existing observational tools limit this construct to one item generally, and do not therefore distinguish between connection to personal experiences, community activities, and cultural traditions.

In primary schooling, project-based learning activities based on children's individual interests can be the focus of school and community activities and support their exploration of the 'universe of alternatives' in any given setting (Sarason, 1996). These may incorporate elaborate play sequences with peers in school or community settings. Providing community-based mentorship for youth to pursue their own arts-based interests and activities in schools, for example, has shown positive impacts on traditional achievement measures and on school attendance in urban El Salvador (Dinarte Diaz & Egana-del Sol, 2019). An arts- and theatre-based approach to job training for youth showed positive impacts on earnings and employment in Rio de Janeiro (Calero et al., 2017).

Caregivers may also engage in connection to experience by reminding children of something they experienced in the past. They may create continuity by stating a link to a child's preferences (e.g. for a toy or activity). Referring to events and experiences that occurred in the past can build children's memory, sense of self, and language ability (Reese, Haden, & Fivush, 1993).

As in the early years, linking exploration to previous experience in community and cultural contexts can be powerful in supporting self-sustaining engagement and learning through play for older children and youth. Project-based pedagogy that incorporates learning as it occurs in nature and non-human systems, as reflected in indigenous cultural ways of knowing, has been used to build systems-based and critical thinking skills in both Native American and non-Native children (Bang, 2020).

Support for problem solving

Support for problem solving in play may be linked to the meaningful, iterative, engaging, and attentive dimensions of learning through play. As children explore their environment with increasingly advanced cognitive skills over the first 12 years of life, they generate, test, and revise 'hypotheses' about the natural and social worlds (Hirsh-Pasek et al., 2020). For example, as they explore which puzzle piece fits in a spot or how much force to use to throw the ball for it to land in the basket, they are problem solving. Problem solving can be an individual or collaborative endeavour. Support for problem solving involves practical as well as emotional scaffolding.

Supporting individual problem solving can happen through guided play, such as when for adults elicit children's hypotheses. For example, a teacher may ask children to speculate how a character in a story might resolve a dilemma by asking what that character might do next. They may ask children to generate explanations for natural or physical phenomena such as different weather patterns. They may encourage multiple approaches to a task such as building something together. This can also involve piquing the child's curiosity and excitement about how to solve the problem so that they are interested in exploring different ways to do so, thus generating sustained engagement with a learning activity. Caregivers may engage in these same kinds of activities with their children, supporting their ability to try an activity in a new way or providing feedback during the activity. Such scaffolding has been incorporated in rich countries into programmes to enhance social-emotional development in young children (Webster-Stratton & Reid, 2018), with some evidence of successful adaptation in Jamaica (Baker-Henningham et al., 2009).

It is also important that children have opportunities to solve problems collaboratively with their peers and siblings. While problem solving individually can be exciting and rewarding by itself, doing so with peers, especially in a relatively unstructured way, can provide the opportunity to learn different types of skills such as joint communication or observational learning, as well as lead to more creative ways of problem solving (Ramani, 2012; Hirsh-

Pasek et al., 2020). Communication and cooperation skills grow over the early childhood years, and supporting them is crucial to expanding children's learning opportunities. Support for collaborative solving can look similar to individual problem solving, with more advanced support for navigating issues such as communication, cooperation, and turn-taking between children.

When applied to a broader and more complex array of activities in middle childhood and adolescence, problem solving may continue to involve the generation of multiple solutions. Teachers may encourage multiple solutions by asking questions in ways that foster multiple answers rather than a single rote answer. Individualised student responses can then result in a range of hypotheses, ideas, solutions, and perspectives on phenomena, all of which may increase motivation in learning activities.

As games and sports increase in complexity across middle childhood, game-based play can encourage cognitive reflection as participants problem solve to move ahead in the game, or generate and revise rules for how to play (Gjicali, Finn, & Hebert, 2020). In person, adults and peers alike can support collaborative problem solving with creative solutions, contributing to disciplinary knowledge and action. The characteristics of intrinsic motivation, attention, and flexibility are all still quite relevant in later development as characteristic of student and classroom engagement in problem solving.

Support for social connectedness

Across early childhood through adolescence, a critical developmental task occurs at the intersection of collaboration, engagement, and learning through play. That is, learning through play occurs in increasingly complex social interactions. At the beginning of life, dyadic interactions with caregivers, peers, siblings, and relatives provide the trust and attachment that allow for exploration and learning about the social and physical world. As group-based interactions become more complex after children start to walk, play begins to occur in peer groups, and not just in dyadic contexts. The well-known transition from parallel play (with children engaged in solo activity with only occasional interaction) to peer interactive play (with children collaborating in pairs or small groups) occurs during the early childhood and pre-primary years, with teachers and caregivers playing a critical role. Caregivers may encourage peer activity and social connections, but it may be difficult to directly observe this in a dyadic caregiver-child observation.

Play is not only linked to individual attributes such as individual creativity and agency but could be associated with other socially valued outcomes in communities, such as cooperation. These themselves may be linked to individual-level engagement in play, especially in the peer and multiparty contexts of socialisation in communities (DeLeón, 2011). In many cultures, social connectedness and related areas of social responsibility, including behaviours such as respect, obedience, acknowledging the wisdom of elders, and care for others, are valued alongside individual agency and initiative, as has been observed among the Ewe in Togo and A-Chewa speakers in Zambia, or among Latinx immigrant communities in North America (Ng et al., 2012; Noyau & Gbeto, 2004; Serpell, 1993). In contrast, in WEIRD middle-class contexts, children often do not contribute to community or household activities without being asked (Rogoff, 2014). Consideration of the overall direction of the group in a collaborative play task was higher in Mexican-origin children in the United States compared with European American middle-class children (Alcalá, Rogoff, & Fraire, 2018).

Connectedness may also constitute an important component of children's intrinsic motivation. In collectivist cultures, children's engagement in activities may be motivated by a

desire to strengthen relationships with caregivers or teachers. For example, parents and teachers in Tanzania said that one of the qualities children needed to succeed at school was to 'love' their teacher (Jukes et al., 2018). In this study, both children and teachers described their motivation for most of their actions – from being polite and studying hard to being curious – was for the teacher to love them. The importance of connectedness was mentioned by both children ('I was happy to meet my teacher, I love her deeply from my heart') and teachers ('[the children who learn well are those who] love to be close to teachers, they trust their teachers in a way that nobody else can teach them').

Teachers may support social connectedness in early childhood by first encouraging children's basic interactive skills – their ability to form friendships by initiating conversation; to invite a peer into an activity; and to listen to one another. An emphasis on prosocial behaviours and social norms in the classroom or centre can facilitate such emerging friendships and social connectedness.

In primary education, teachers may continue to support such prosocial behaviours, but there may be increasing focus on group work that is sustained across time. Collaborative group work can be encouraged by teachers and break up predominant patterns of didactic, teacher-led or lecture-based instruction. Working towards a common goal can be made explicit by the teacher. There may be a resulting overall sense of camaraderie and connectedness of the classroom as a whole.

Specific activities may support social responsibility in addition to social connectedness. For example, creative storytelling in early education and creative writing in primary and secondary education can encourage drawing moral and social lessons related to responsibility, as was observed in a Chinese preschool in the well-known Preschool in Three Cultures study (Tobin, Hsueh, & Karasawa, 2009). Traditions of indigenous, oral storytelling in the Chittagong Hills Tracts of Bangladesh have been incorporated along with the interpretation of moral lessons into a curriculum in a primary school intervention, with positive impacts on writing, comprehension, and vocabulary (Nyeu, 2020).

Positive emotional climate

In more general observed quality measures, emotional climate has often been prominent as a predictor of both social-emotional and academic learning outcomes (Wolf et al., 2018). In such studies, this construct has included individualised attention by the teacher, behaviours such as praise and positive reinforcement, and the observed presence of specific positive emotions such as enthusiasm or smiling. In addition, behavioural aspects of pedagogy such as behaviour management (the response to difficult or problem behaviours) or encouragement of prosocial behaviour has been part of emotional climate constructs.

Praise and warmth in caregiver-child interactions, similarly, is central to Western conceptions of responsive and authoritative parenting (Baumrind, 2013). However, positive emotions may be expressed differently in different cultures, in both degree and in types of facial expressions or gestures (Sorkhabi, 2005).

Considered from a broader cross-cultural perspective, learning through play might not occur in contexts where the affective climate is in conflict with support of intrinsic motivation and flow aspects of self-sustaining engagement. Positive emotions can be powerful in motivating learning, whether such emotions are modelled by the teacher or encouraged among students. However, the type and valence of emotion may differ by cultural context. In the WEIRD literature on classroom quality and child learning, for example, emotional climate characteristics such as warmth, praise, and smiling have been predictive of academic and social outcomes in children. They have also been correlated with other aspects of guided

play interactions or learning-through-play characteristics, such as support for exploration and connection to student experiences (e.g. in secondary schools in Uganda, as noted by Seidman et al., 2018; or student engagement in preschools in Ghana, as noted by Wolf et al., 2017).

However, in other cultures, learning through play might not be as strongly associated with warmth and praise. An emotional climate that fosters sustained attention may include individualised attention in centres, classrooms, or homes, but without strong affective display.

2.3.4 Other Dimensions of Activity Settings That May Support Self-Sustaining Engagement and Learning through Play

Additional dimensions of activity settings may influence learning through play through their influence on the preceding dimensions. They constitute more general aspects of the interaction-based theory of educational quality put forward, for example, by Hamre and colleagues (2014). In some cases, they may serve as moderators of the associations of the social interaction dimensions with child-level learning through play.

Beliefs and goals of members of the activity setting and broader systems. The beliefs and goals of teachers and caregivers as well as children and youth may powerfully shape whether and how guided play occurs and whether it influences self-sustaining engagement. These in turn may be influenced by both community norms and formal structures such as goals reflected in curricula or educational policy implementation (Harwood et al., 1996). Systems such as workforce, governance, and training and professional development organisations and institutions may communicate expectations regarding how learning through play is defined and whether learning through play is integral to broader learning goals. In studies of culture and human development, beliefs, goals, and values are critical contexts of intergenerational continuity and change in socialisation (Keller et al., 2006). In our framework, we assess teachers' beliefs about whether particular pedagogical approaches linked to our constructs are important for child learning.

Safety, security, and predictability. Although social interactions that support learning through play could occur in the absence of safety, monitoring for safety may take precedence over the scaffolding of social interactions. The provision of a safe, secure, and predictable environment may be a necessary condition for interactions that support self-sustaining engagement. These may be particularly threatened in cases of armed conflict, disaster, or crisis.

Physical aspects of the setting, including materials. Learning through play can occur in the absence of 'formal' play or learning materials and instead with aspects of the natural environment. Support of exploration is often in the form of supporting physical exploration (e.g. of properties of the natural world, household objects, or work environments such as shops in the community). We assess learning materials and other resources in the classroom in our centre- and school-based assessments.

Instructional content. Supporting learning through play can occur across different forms of instructional content – including language, quantitative reasoning and mathematics, science, exploration of the natural environment, and of course the arts. A challenge in recent work on measurement of quality of education in centre-based and school settings is that of assessing the quality of pedagogy designed to facilitate particular domains of learning (e.g. social-emotional; language and literacy; or numeracy and mathematics) (Burchinal, 2018). We believe that each of these areas of instruction can occur in ways that support self-sustaining engagement. But how the dimensions of social interaction we measure may manifest

differently in different areas of instruction has yet to be determined and will vary based on curricular goals, implementation, and training and professional development systems.

2.3.5 Relationship of Learning through Play to Learning Outcomes

Finally, moving to the relationship between the characteristics of learning through play and learning outcomes (the middle and right columns in Figure 2), we incorporate traditional domains of learning (e.g. social-emotional, cognitive, language, and numeracy) with recent thinking on forms of learning that may be particularly associated with learning through play, as well as specific learning outcomes valued by specific cultural communities in the socialisation of children.

Hirsh-Pasek et al. (2020) and Zosh, Hassinger-Das, and Laurie (2022) summarise evidence to support how learning through play activities can lead to a range of wider and deeper learning outcomes than reading and math skills. The 6 'Cs,' for example – collaboration, communication, content, critical thinking, creative innovation, and confidence – represent hypotheses about the domains of learning outcomes that may be most associated with playful activities. Learning outcomes of play as it has been defined in this typology have emerged from WEIRD contexts and been communicated largely within WEIRD contexts, without sufficient testing in a larger variety of cultural contexts (Serpell, 2021; Zosh, Hassinger-Das, & Laurie, 2022).

In non-WEIRD contexts, a long tradition of research on culture and human development shows the intertwined nature of cognitive development with valued social outcomes such as social responsibility, respect, and engagement in community and household routines (Jukes et al., 2018; Serpell, 2011; Weisner, 2014). Play may then both reflect and provide a means to practice the roles, behaviours, and norms of the local culture in alignment with these developmental values and goals (Roopnarine, 2012).

A culture and human development lens is imperative for understanding whether these forms of play are indeed meaningful in a much wider set of cultural and societal contexts. What Zosh et al. (2017) define as learning through play may vary both in prevalence and in the magnitude of their associations with important learning outcomes. For example, in some settings, the boundary between 'play' and 'work' may be more porous than in others, with children's involvement in household work activities often showing dimensions of voluntary engagement, attention, and creativity and enjoyment that may meet some of the criteria defined here as play (Gaskins, 2015).

The predictive relations of play behaviours with learning and other outcomes may vary across cultures, or within cultures across historical time. One example is the differential impact of behavioural social inhibition on social and psychological adjustment among North American and Chinese children. Specifically, while in North America, children high on behavioural inhibition (a form of 'shyness') often have poorer academic performance and peer relationships, which later impacts their educational attainment as adults (Caspi, Elder, & Bem, 1988; Gazelle & Ladd, 2003), in China, these children are often well accepted by their peers and report high academic performance and psychological well-being (Chen, Rubin, & Sun, 1992). As urban China shifted from planned to market economy and increasingly absorbed Western media influences, the association between shyness and peer, parent, and teacher ratings of overall social competence shifted from positive to negative, over a 15-year period (Chen et al., 2005; Yoshikawa, Way, & Chen, 2012). Learning through play in solitary versus peer groups may therefore also have changed in their associations with developmental outcomes.

2.4 Conclusion

In this section, we have endeavoured to define a typology of the kinds of interactions that may support self-sustaining engagement and learning through play and inform educational and human development interventions. These dimensions of interactions are intended to guide play-based learning programmes and curricular frameworks across the facilitation spectrum, from adult facilitation of free play to inclusion of learning through play in content instructions. We have also endeavoured to describe how these dimensions of guided play and social interactions may vary by cultural context in important ways. They incorporate flexibility in application across cultural and educational contexts that may differ in notions of and balance of autonomy and connectedness; in levels and types of child agency that are encouraged; and in the representation of activities across the play spectrum.

Through their development into an initial set of tools for early childhood and middle childhood home, centre and school settings, most of these constructs have been retained, though with some revisions in emphasis and representation. Future efforts will focus on testing this framework against actual child engagement and learning outcomes in different country and cultural contexts.

3. CONTEXTUALISATION FRAMEWORK

3.1 Introduction

The PLAY Measurement toolkit is designed to be used in different contexts. We considered the context of use in the design of both universal and context-specific aspects of the toolkit. First, universal elements of the toolkit are designed to be applicable across contexts. For example, we aimed to use the same conceptual framework – that is, the same set of constructs – in all contexts. This goal requires an understanding of the full range of contexts in which the toolkit may be used. Second, we considered ways in which the toolkit may need to be adapted to each context in which it is used.

In this section, we define context and outline our approach to contextualisation. We then present the theoretical framework that guided our work. Lastly, we describe the principles of our contextualisation process and provide an overview of the process that is described in subsequent sections of the report.

3.2 A Definition of Context

The key aspects of context are those that have a bearing on the way that children's engagement is supported. We hypothesise the following main drivers of contextual differences in support of children's engagement.

Culture. We define culture as the social learning that influences behaviour (Richerson & Boyd, 2005) based on cultural practices, defined as 'actions that are repeated, shared with others in a social group, and invested with normative expectations and with meanings or significances that go beyond the immediate goals of the action' (Rogoff et al., 2003). Children's development is shaped by the 'developmental niche' (Super & Harkness, 1986), a concept that encompasses the physical and social settings of everyday life, as well as community beliefs and socialisation practices, conveyed across generations. Cultures may vary in their understanding of how children should engage in their learning and how this engagement can be supported.

Cultural practices may be considered adaptive to a particular sociodemographic context (Greenfield, 2009). In this way, sociodemographic factors (e.g. urban/rural, education levels, economic activities, or use of technology) also become important contextual factors.

Education and family policies and system capacity. Current and past policies influence attitudes towards engagement and familiarity with pedagogical and caregiving approaches to supporting learner engagement. The capacity of teachers – that is, their experience and training – in a context can influence how effectively they are able to support learner engagement in the classroom. Similarly, the capacity, values, resources, and knowledge of caregivers influence their caregiving practices.

Crisis and humanitarian settings. There are additional factors for the toolkit to consider in crisis and humanitarian settings. First, it is important that a tool be accompanied by efforts to understand and document both acute and chronic stressors related to crisis and displacement that teachers, caregivers and students may face. Humanitarian contexts may be ones of acute or protracted crisis, and the stressors may be different depending on the longevity or recency of the crisis and its immediate threat to basic shelter, safety, and survival. Second, efforts to adapt the toolkit to situations of displacement should consider the role of the culture and language of the displaced community, as well as host community members, if relevant. Caregivers and teachers from the displaced and host communities may be encountering multilingual and multicultural environments in new ways, particularly if national boundaries have been crossed. Third, there may be a need for educational and family-focused programming to incorporate the assessment of basic and crisis-driven needs (e.g. related to health, nutrition, shelter, or mental health) prior to or in tandem with the assessment of educational supports.

As the toolkit is adapted from one context to another, the level of adaptation required will depend on these three characteristics – namely, culture; education and family policies and system capacity; and presence of humanitarian crises.

3.3 Overview of the Approach to Contextualisation

Our overall approach to contextualisation is informed by an understanding of the different levels at which contextualisation can take place.

First, the **constructs** that make up the tool may be more or less important in different contexts. For example, social connectedness may be a motivating factor for children's engagement in one context but not in another and consequently may be predictive of learning in one context but not in another.

Second, the **items** that constitute a given construct may be constructed differently in different contexts. For example, the behaviours that exemplify social connectedness may differ from one context to another such that items need to be reworded – or new items added – to include contextually relevant behaviours.

Third, similar items may be observed with different **quality metrics** (different frequencies or different levels of effectiveness or participation) across contexts. For example, there may be a high level of support for children's exploration in one context but not in another. Adaptation to items may be required to tune the sensitivity of the tools – for example, to avoid floor and ceiling effects – and to ensure that there is a good balance of high- and low-quality items in a given tool.

3.4 Framework for the Contextualisation of PLAY

3.4.1 Culture

Much of the literature on playful learning and children's engagement comes from WEIRD countries, as is the case for much of psychological science (Arnett, 2016; Henrich et al., 2010; Rad et al., 2018). Childhoods in WEIRD contexts have many similarities (Lancy, 2014). Smaller families and more resources lead to greater investment in individual children – particularly in their cognitive development – to prepare them for life in the globalised knowledge economy. As a result, WEIRD cultures typically value competencies such as independence, self-expression, curiosity, extraversion, uniqueness, competition, and self-esteem.

Our approach to contextualisation considers how this set of cultural values may have shaped previous work on children's engagement in learning through play and learning in general, and how a measurement tool should be adjusted for use in a wider range of contexts. To achieve this aim, we paid particular attention to contexts where a contrasting set of values shapes human development. Namely, in societies currently engaged in (or with a recent history of engaging in) subsistence agriculture, economic productivity depends on cooperation (Greenfield, 2016), and maintaining relationships with peers and elders is of paramount importance. As a result, children's development is shaped by the values of respect, obedience, social responsibility, shyness, cooperation, empathy, and emotional control. At the same time, the majority of these countries are undergoing rapid transitions such that people in some areas of the country (often the capital or other relatively more educated regions) espouse values of individual achievement and creativity, competition, and education- and earnings-based social mobility (Chen et al., 2021; Fong, 2007; Nieto, Leyva, & Yoshikawa, 2019; Jukes, Zuilkowski, & Grigorenko, 2018).

Based on a literature review, we identified the following ways in which these different sets of cultural values have implications for the context-specific approach to the measurement of learner engagement. The literature that we reviewed had a significant focus on learning through play as one approach to promoting learner engagement.

Adult-directed activities. In subsistence communities, hierarchical relationships are emphasised (Keller, 2016). In these cultures, it is common for adults to direct activities and for children to comply. This is evident in the predominance of teacher-led activities in the classroom. It is also evident in the way that adults ask children closed questions seeking a 'correct' answer. Some evidence suggests that adult-directed play is more predictive of positive child outcomes compared to child-directed play (Kärtner, Keller, & Chaudhary, 2010). Such adult direction may not necessarily be a restriction on child autonomy. Children may have autonomy but choose to direct their action towards communal goals (defined by adults) because it confers benefits in terms of a sense of belonging and strengthened relationships.

Implications: This observation has implications for items designed to measure support for agency. Agency can be characterised by the degree to which activities are *child initiated* and *child directed.* Our hypothesis is that when the toolkit is used in cultures where hierarchical adult-child relationships are the norm, fewer items measuring support for *child-initiated* activities will show variation, and the degree of child direction of activities will be limited. Conversely, in WEIRD settings, more items measuring support for *child-initiated* activities may be required and show variation. Quantitative pilot data using the tool could indicate the sensitivity of items in the 'support for agency' construct. If more sensitivity is required in a

new context, items could be added or adapted to adjust the level of child initiation and child direction.

In addition, the above hypothesis could be tested. Qualitative work could investigate the extent to which children are engaged by activities that they initiate and direct. Quantitative data could be used to determine the relationship between items in the observation tool and subsequent learning outcomes. Our hypothesis is that high levels of child initiation and child direction would be more predictive of learning outcomes in WEIRD contexts than in subsistence agricultural contexts.

Relatedness and a sense of belonging. As discussed above, children's compliance with adult direction should be seen in the context of benefits to children in terms of a sense of belonging and a strengthening of relationships with adults and the community. Qualitative data from primary schools in Tanzania suggest that a child's 'love' of their teacher (and vice versa) is foremost in their minds – the relationship gives them confidence to study, and they are diligent in schoolwork in order to strengthen the relationship further (Jukes, Gabrieli et al., 2018). Data from Islamic societies in 15 countries supported a 'sense of belonging' construct in school children (Nasser et al., 2019). The presence of role models, a sense of purpose, ownership, and respect may also be important for children's motivation.

Implications: The toolkit includes a social connectedness construct. This construct includes observations of interactions that support engagement by building on children's interpersonal closeness (e.g. between a small group of peers), togetherness and sense of belonging (to the group as a whole), and their pursuit of communal goals. The hypothesis is that this construct will be a more important determinant of learner engagement in contexts closer to the culture of subsistence communities. Data collection in additional contexts can test this hypothesis and help clarify the key components of support for social connectedness. For example, what is the relative importance of the three proposed components of social connectedness (interpersonal closeness, a sense of belonging, and the pursuit of communal goals)? Qualitative work in the contextualisation phase can help investigate this concept further.

Play partners. Adults are more likely to be dyadic play partners in Western middle-class societies. In other cultures, adults are less likely to be involved in play with children, and their involvement may even be seen as limiting the playfulness of an activity. In many cultures, multi-age peer groups scaffold play to a greater degree than adults (Tronick, Morelli, & Ivey, 1992; García-Sánchez, 2016; Nsamenang, 2010).

Implications: This may imply that teachers and caregivers will be less influential in facilitating learning through play in some cultures. It may also imply that observing children's play and engagement in the absence of adults will be more important in some cultures than others. The toolkit at the classroom level includes some items that record opportunities for children to interact with one another without direct teacher supervision or support. Qualitative work in the contextualisation phase could investigate further the extent to which children are engaged in such interactions. Quantitative work could examine whether such interactions are predictive of learning.

Integration of play and work. Children in subsistence communities are more likely to try out work-like activities in a playful manner. The distinction between work and play is less clear (Gaskins, Haight, & Lancy, 2007; Gaskins, 2015; Gosso & Carvalho, 2013). Some of the characteristics of learning through play, such as children being engaged and self-directed while learning, may be evident in work-related activities.

Implications: Our home-based tools incorporate work-like activities in their approaches. The caregiver surveys include a section about household chores, with questions about routines and playing through work-like tasks. In observations of adult-child interactions, the toolkit includes different items related to routines and grooming (e.g. mirror and brush) based on qualitative findings of learning through chores, routines, and work.

Our classroom-based tools also include items related to work-like activities, such as activities where children work together on a task that has a practical purpose. The primary tool includes items where students are given a specific responsibility and where activities encourage students to work towards a common goal.

As with other dimensions of culture discussed above, the importance of work-like activities for children's engagement could be investigated further through qualitative work and tested through quantitative data collection.

Goals of play and learning. Parents in Western middle-class societies engage children in pretend play, in part to develop children's imagination and creativity. In other societies, such play emerges without much parental encouragement, as children imitate household and community routines in their independent peer play (Bolisetty et al., 2022). In general, the implicit and explicit socialisation goals of caregivers and the learning goals of teachers may support different types of play in different cultural contexts. In some subsistence communities, play is seen as a way to occupy children rather than for them to learn (Gaskins, 2015).

Implications: This finding may have particular implications for the observation of caregiver-child interactions. We may expect to see fewer examples of pretend play in the presence of adults in subsistence communities. If teachers share the socialisation goals of parents, it is possible that we will see less pretend play in early childhood and primary classrooms.

The role of observation in learning. Rogoff et al. (2003) describe the important role of observation in learning in subsistence communities. 'Intent participation' is not just passive observation but observation with the intent to participate in an activity.

Implications: We may assume that the learning-through-play characteristic of 'active engagement' requires (immediate) participation on the part of children. Perhaps there is a cultural difference in the extent to which children can be 'engaged' without ostensive participation. This observation was difficult to incorporate into the toolkit. Further qualitative work could examine this phenomenon in adult-child interactions and explore items that capture it.

3.4.2 Education Policies and System Capacity

Support for learner engagement in the classroom can be affected by both policy and system capacity. Some education systems adopt reform efforts to increase the use of pedagogies to support playful learning and children's engagement. For example, in Colombia's education system, the four chief activities in early childhood education include play (Ministerio de Educación Nacional, 2017). Such policies interact with teacher capacity (Akyeampong, 2014). Teachers with experience, training, and support in such pedagogies are more able to implement them in the classroom.

Within the contextualisation approach described here, we anticipate that education policies and system capacity primarily affect the frequency with which items and constructs are observed. That is, we expect to find that items measuring support for learner engagement will be observed more frequently in contexts with education policies and system capacity that support this kind of instruction. Below, we suggest two approaches to adapting the toolkit

between high- and low-capacity environments. The first approach involves selecting items that are more suited to the environment in which the tool is being used. The second approach is to adjust the 'difficulty' of items – for example, by changing the metrics for the item. One way to achieve this is to raise the threshold for 'high quality' of observed items from twice in a lesson to three or four times in a lesson.

3.4.3 Crisis and Humanitarian Contexts

There are several aspects of tool design that need to be considered when working in crisis and humanitarian settings. First, it is important to understand and document the stress facing teachers and caregivers. Second, training should be flexible to allow for potential interruptions. Third, it is important to adapt the toolkit to the culture and language of the displaced people and to seek appropriate expertise. Fourth, the toolkit should exhibit sensitivity to topics such as war-based trauma and mental health. Fifth, training and support for data collectors should acknowledge the practicalities of data collection in crowded households and centres. For an overview of such methods and approaches, see Goodfriend et al. (2022).

3.5 Principles Guiding the Contextualisation Process

Defining constructs. The evidence in this review suggests that certain learning-through-play constructs may be more relevant in some contexts than others. Rather than tailor the set of constructs measured in the tool to each context, which would be time-consuming, we aimed to include all potential constructs in a single comprehensive tool. Not all constructs will be relevant in a given context, but the tool will be capable of capturing any learning-through-play constructs that are relevant. In this approach, **the aim of the work to understand contextual and cultural variation in learning through play is to map out the 'universe of alternatives'** (Sarason, 1982, cited in Tseng and Siedman, 2007) for such constructs, with the aim of including all alternatives in a single tool. It is unlikely that we were able to map out the entire 'universe of alternatives' in this project, but we retain this as an aim for the toolkit as it undergoes further development in the next phase of work (PLAY 2.0).

Interpreting constructs. While we aim to have the same set of constructs across contexts, we expect that each construct will manifest differently across contexts. Contextualisation work will aim to understand the way that the same construct may be assessed by different behaviours across contexts.

Which constructs lead to learning? The behaviours that are adaptive vary across contexts. Different behaviours lead to development and learning across contexts. For example, a longitudinal study of young children in Delhi and Berlin (Schroder et al., 2012) found that different behaviours predicted positive outcomes at age 3. In Berlin, maternal support for toddlers' self-expression during free play at 19 months predicted positive outcomes (children's memory elaborations) at 3 years. In Delhi, toddlers' willingness to carry out their mothers' requests at 19 months predicted positive outcomes at 3 years. The findings of this study suggest that child-initiated play may not be more beneficial than adult-initiated play in all contexts. The implication is that the same behaviour may lead to different outcomes in different contexts. The relationship between the PLAY Measurement tools and learning outcomes will be assessed in PLAY 2.0.

Changing constructs. Similarly, the outcomes of a specific behaviour may change over time in one context. Chen et al. (2005) studied three cohorts of Chinese schoolchildren and found changing associations with shyness. Whereas shyness was associated with social and academic achievement in the 1990 cohort, the associations became weaker or

nonsignificant in the 1998 cohort. Furthermore, shyness was associated with peer rejection, school problems, and depression in the 2002 cohort. This highlights the need to continue collecting data on the relationship between the PLAY tools and learning beyond PLAY 2.0.

3.6 Overview of the Contextualisation Process

The contextualisation process involved three steps. The first step was to investigate local participants' understanding of key constructs in this project: 'play,' 'learning,' 'learning through play,' and 'engagement in learning.' The second step was to document how dimensions of support for learner engagement – those proposed in the conceptual framework as well as potential additions to the framework – arise in discussion with participants. The third step was to identify the behaviours that exemplify the constructs in each context. These three steps were operationalised in the Build phase of the project, which is described in Section 4. The findings of the Build phase are described in Section 5, and Section 7 documents how the conceptual framework and the tools were developed based on these findings.

4. BUILD PHASE DATA COLLECTION METHODS

Qualitative methods were employed in the Build phase to develop an understanding of project-relevant terms in a localised manner. Methods employed varied by country and focal age group (i.e. early childhood or primary). **Table 3** provides a brief description of each method.

Table 3. Overview of Build phase methods by country

	Colombia early childhood + primary	Ghana primary	Jordan	Kenya primary
Build phase data collection	Sept. 27–Oct. 8, 2021	Oct. 12–22, 2021	Mar. 20–28, 2021	June 25–July 28, 2021
Child level	• Point-of-view observations • Drawing focus groups			• Drawing focus groups
Adult level	• Teacher-child interaction observations • Focus groups • Surveys	• Teacher-child interaction observations • Focus groups • Surveys	• Teacher-child interaction observations • Focus groups	• Teacher-child interaction observations • Focus groups • Surveys
Setting level	• Classroom naturalistic observations	• Classroom naturalistic observations	• Classroom naturalistic observations	• Classroom naturalistic observations

Note: For details on sample sizes per country, see Appendix A.

4.1 Methods

4.1.1 Child Level

Point-of-view observations, or 'child pov' observations, are five-minute observations that result in a detailed narrative summary of the child's point of view. These data were collected in the family environment and community centre settings. For each child pov observation, there was one focal child whom the observer was meant to follow around and take notes on.

The observer recorded notes to provide a general sense of what and whom the child interacts with using a 'camera on head' approach – that is, recording notes from the child's point of view as much as possible. Observers were tasked with focusing on the focal child's reaction, gestures, and vocalisations, paying particular attention to concrete behaviours, actions, words, and observed emotions. We suggested the following prompts for observers to consider when conducting child pov observations: *What is the child doing? With whom is the child interacting (or not)? What objects or people does the child fixate their gaze upon? What noises prompt the child to turn their head? What is said to the child, and how do they respond?* The purpose of this method was to facilitate our understanding of what play might look like in a child's natural context by allowing us to gauge with whom the child primarily socialises and to effectively 'see' the world through the child's eyes (Smith et al., 2015; Kim, 2020). These data allowed us to think about whether our conceptualised dimensions of play applied to the respective setting. In addition, we generated culturally relevant examples for the rubric in the observation tools based on findings.

Illuminative drawing + focus group is an arts-based technique of data collection with children (Coyne & Carter, 2018). We asked children to draw two pictures, one per prompt: (1) *Draw anything you can think of when you hear the word 'play.' For example, you could draw you and your friends playing.* (2) *Draw anything you can think of when you hear the words 'learning something.' For example, can you draw yourself learning something in the classroom or outside (anywhere)?* The drawings served as a springboard for discussion on the children's experiences, thoughts, and beliefs in the focus group setting after they completed their drawings. Once children finished drawing, data collectors asked each child about both pictures and recorded notes carefully. For example, *Tell me more about your drawing – what is happening here? What/who did you draw and why? What are the characters/people in your drawing feeling, and why?* Then, data collectors asked children to share their images by laying them in the middle of the table or on the floor. Data collectors provided guidelines for children (e.g. 'Let's not interrupt each other') and encouraged a free-flowing conversation with questions to help facilitate the discussion, especially for quieter groups. Specifically, children were prompted to compare and contrast their drawings, noticing colours, the characters or people portrayed, settings, and feelings. In this way, children were the experts on their own experiences, and they made their own meaning of play and learning through drawing.

4.1.2 Adult Level

Teacher-child interaction observations took place in the classroom during a 'typical' day and focused on teachers and how they interact with children. Peer interactions were not a focus of this observation. Each observation lasted approximately 30–45 minutes, or if slightly outside of this window of time, followed the natural course of the lesson. Observers kept a detailed record of what occurred in an open-ended manner, free from any particular coding scheme. The goal was to capture exchanges and conversations between the teacher and any child(ren). The protocol included several prompts to guide observers in their notetaking related to each construct, such as *How are children in this setting supported in exploring their surroundings? Elaborate events or situations they are describing. How are children encouraged to elaborate solutions or come up with new ideas to address a problem or to resolve a situation (e.g. figuring out how something works)? How are children in this setting encouraged to link an idea or object being discussed in the classroom to an aspect of their lives outside the classroom or in the past?* Observers were asked to use pseudonyms in their notes and not participants' names. For sequences of interactions, we guided observers to focus particularly on (1) gestures; (2) dialogue, noting tone and vocal inflection; (3) facial expressions; (4) affect (e.g. smile, frown); and (5) situations that encourage imagination or

coming up with new ideas and solutions (e.g. brainstorming 'what if…,' 'should we…,' 'what would you do if…').

Focus groups were conducted with teacher and caregiver participants separately. Focus groups ranged from four to six adult participants each. Facilitators conducted focus groups in informal settings, usually at the schools or community centres. They took notes on responses during the sessions. Where possible, facilitators audio recorded focus group discussions and listened to them again later in order to add details to their notes. Facilitators had protocols with prompts and open-ended questions in order to keep the group on topic but were otherwise non-directive; the facilitator was tasked with letting the group explore each subject and question. The protocols varied slightly by country and age group but largely aimed to cover the same topics. For each topic, the protocol stated the aim or goal of the question, the primary question, and possible probes the facilitator could use if the group was quiet or just to elicit more information. Some examples of aims and goals of questions in the protocol include the following: to understand what it is about activities that keep children engaged and what children experience when they are completely engaged; to understand what participants do to allow children to be completely engaged in an activity; to understand participants' opinions about what might motivate children to enjoy some activities over others; and to understand how participants help create environments or opportunities for children to make their own choices and decisions. Caregiver focus groups, in particular for the younger age group (i.e. 0–2 years), homed in on child–family member interactions and the child's routine at home. In addition, facilitators recorded basic demographic information about the group.

Surveys were semi-structured and conducted with teacher and caregiver participants separately. Interviewers did not abide strictly by a list of formalised questions but rather asked open-ended questions, allowing for a discussion with the interviewee rather than a straightforward question-and-answer format. In addition, the protocols had suggested probes to help interviewees elaborate. Interviewees took notes and, where possible, audio recorded the conversations and listened to them again later in order to make their notes more detailed. Survey protocols included similar questions and prompts as described for the focus group discussions. Interviewers also recorded basic demographic information on each participant.

4.1.3 Setting Level

Classroom naturalistic observations took place in the classroom during a 'typical' day and, unlike the observations described above, focused on the classroom environment as a whole. Each observation lasted approximately 30-45 minutes, or if slightly outside of this window of time, followed the natural course of the lesson. Observers assumed the role of passive observer (Mack 2005; Zieman, 2012), taking detailed notes on the participants in their environment, free of judgment, noting relationships among participants and with their environment, as well as observing their behaviours. One goal for the naturalistic observation in classrooms was to provide a framework through which to understand the larger context (e.g. physical, social, and cultural) in which participants live and operate (Kawulich, 2005; Mack, 2005; Zieman, 2012). The focus was not on the teacher's activities only but on the overall essence of the classroom, capturing snapshots of what everyone was doing, including children in peer groups and on their own. The protocol included a section for observers to describe the setting (e.g. physical space, materials or resources available, and size) and the participants being observed (e.g. ages and number of children and adults). The protocol provided two focal questions for the observer to consider: (1) When are children engaged in a self-sustained way? and (2) How do adults and others support or prompt that

engagement? Then, the protocol provided a chart for open-ended notetaking according to three-minute time intervals. This breakdown was provided in order for notes to be more organised and coherent later on. After the naturalistic observation was complete, observers completed a post-observation reflection and summary notes. This included several prompts to describe the setting and participants – for example, *describe what was happening in the classroom/playground today; describe what the teacher(s) were doing; highlight one interesting occurrence where students were most engaged and explain what the children and teachers were doing at this time, and how you could tell that the children were engaged.*

Using both inductive and deductive analytical approaches, we generated a variety of qualitative outputs to aid in tool revision during the Adapt phase.

5. OVERVIEW OF BUILD PHASE FINDINGS

In this section, we present qualitative findings from classroom observations and from caregiver and teacher surveys and focus groups, which were utilised to refine the six PLAY constructs and also helped us generate items for the tools. The implications of the Build phase findings differed by construct. Some revisions were subtle; others were significant. Note that the Build phase marked the beginning of the process of adaptation and tool development; significant work was conducted in later stages. In this section, we first describe the findings for each construct and their impact on the tools. We then offer additional insights and explore certain themes via a country-level overview of the Build phase data.

5.1 Analysis by Construct

Build phase data show a general endorsement of the six constructs across countries and illustrate variation in how they manifested by country as well as age or grade level. These data were compiled as positive examples, or 'endorsement,' of the constructs. In some cases, findings disagreed with the constructs. These data were compiled as 'negative support' and would indicate a reverse-coded item if used in the tools. Negative support was more visible in the primary data than in the early childhood data.

This section is organised according to the six constructs. Within each construct, we share findings from each country, as well as feedback from caregiver and teacher participants.

5.1.1 Support for Agency

Across primary settings, adult support for agentic behaviour can be seen in the data. The Ghana dataset shows that children have the opportunity and resources to exercise decision-making in learning and feel free to use one another as resources in the classroom. Teachers are present for agentic behaviour and remain 'hands-off' rather than involve themselves. For example, observation data show that 'students use different methods to come up with the answer/solution in doing math problems; the teacher taught multiple methods so students can choose whatever method is easiest for and makes the most sense to them.' In Kenya, data indicate that certain activities begin with adult initiation and proceed to become child led, while other activities show adults stepping back, fully allowing child choice and autonomy. This is reflected by the caregiver survey excerpt noting that they support their child 'by respecting her decision and allowing her to proceed with selected activity.' Teacher survey data similarly reflect agency: 'They are always happy and organise themselves during play activities; they follow instructions given by their peers.' As noted by one teacher who was interviewed, 'During group activities I give them leeway to choose whatever they

want to discuss within their groups.' Further, teacher survey data from Colombia reflect strong support for agency: 'The girl took her classmates and sat them down and began to read. She was imitating what we did before, organising them on their mats and reading them a story. She was reading and was directing the children. I believe that what she did was also imitating her parents because they are always very aware of the activities she does and at home they read stories to her and that is why she is attracted by it.'

In early childhood settings, agency is also valued and encouraged. Teachers seek to offer opportunity and create an environment for child agentic behaviour, while monitoring children to ensure that they have the support and resources they need. In a teacher focus group in Jordan, a teacher reported, 'The teacher gives the child the freedom to choose and supports him in his choices to make the child feel comfort and safe.' This was echoed in a teacher focus group in Colombia: 'They take ownership of things more when they are autonomous. When you allow them to be autonomous, they don't just say I'm going to do something, they go and do it, you go and ask them what they are doing and they tell you, they explain it to you. And they feel much more involved when they are autonomous.' In addition, data collected from caregivers in Colombia show that caregivers facilitate agency and child independence. One caregiver in a focus group stated, 'I tell him the name of the object he is asking for so that he can learn what he is asking for.'

Adult support for agency in its most limited sense may simply mean giving permission to the child to carry out a task or activity they would like to do, thereby transferring agency to the child; negative examples – a lack of support for agency – may be evidenced by adult control. This was evident in the primary datasets. For example, the data from Ghana show that teachers exert control over how children learn, including how children participate. This is seen in a teacher survey: 'The teacher has to monitor so that the students will do the activity right/well; the teacher has to clarify and demonstrate; especially, for students who are not doing well, the teacher has to give instructions again and go back to the board and point to the picture so the students know how to follow the instructions and get the activity correct.' Similarly, a teacher in Kenya emphasised the importance of instruction in encouraging student participation: 'Students will be more active if they have instructions on how to do something; it will be easy for students to follow; the students will stick to what they know and were instructed/taught.'

Teachers in Kenya also pointed to school-level factors that restrict their ability to facilitate agency in the classroom – for example, 'the curriculum is a hindrance because it doesn't offer much freedom of choice and is very rigid,' and 'syllabus is directed and limits opportunities for learners to make choices.'

Overall, most of the Build data therefore supported the idea that children are motivated by autonomy in decision-making. There is little support for the hypothesis in the literature that children may be equally motivated by tasks given to them by someone in their group (to which they felt a sense of belonging). Data from Kenya and Ghana suggest an apparent contradiction in support for agency in that teachers believed children needed explicit instructions to have the confidence to conduct a task by themselves. One implication is that agency may be expressed by children in more limited or subtle ways. This led to the inclusion of items in all age groups concerning children's choice on how to engage in a task that had been set for them by an adult. For example, one item in the 3–5 Caregiver-Child Observation Tool is 'Caregiver permits the child to choose how to engage with material/activity.' A similar item in the primary toolkit is 'Teacher allows freedom in approaching an academic task. Students decided what or how.'

5.1.2 Support for Connection to Experience

In primary settings, we observed and teachers reported connecting learning to children's background knowledge, experiences, and personal lives; integrating local languages where appropriate; and encouraging children to reflect on their families and communities in the classroom. For example, according to observation data in Ghana, 'Students give creative and interesting answers as they apply the story and the meaning of the story and lessons to themselves.' This construct is similarly encouraged in Kenya, where a teacher reflected in a survey, 'It was about the cultural activities that are performed in their area/region so the students showcased dancing moves and costumes and instruments mostly from dominant regions; the students participated though there was a team leader who demonstrated and gave guidance.' And in a focus group, a participant stated that 'learners are able to internalise better when the activities allow them to learn through doing or apply the knowledge in real life situations.' Caregivers in a focus group in Kenya reflected on how support for connection to experience can manifest as intent participation (Rogoff, 2014): 'practising for future desires; the child is practising or pretending to sell the produce in a way to develop skills for future entrepreneurship/businessman because he has seen from his dad and is learning the same skills.' Observation data from a primary classroom in Colombia similarly show that teachers link children's learning to their lives: 'The teacher begins the activity by asking if they remember what they have worked on in their review, and they say "the numbers with the paint." She then asks, "Did you go fishing?" One child responds by recounting an experience with his father fishing in the river.'

Data from early childhood settings also reveal a clear intentionality among teachers and caregivers to invoke past and personal experiences, including topics of interest to children and their home environment. A teacher in Colombia highlighted this in a focus group: 'It is about integrating the experiences that they have on a daily basis and thus, integration.' Caregivers in Colombia facilitate connections for children in the home environment, and children copy or model others. An interview excerpt reflects this: 'For example in the kitchen, when she feeds me [in pretend play], I tell her that we have to pick up the dishes [in the game], and when she eats [at feeding time], when she finishes, she does the same, she picks up the plate and takes it to the kitchen.' Classroom observation data in Jordan show this as well: 'The teacher asked what's the benefits of chamomile? One student answered that when her stomach hurts her mom prepares a chamomile drink for her.'

Negative support for connection to experience appears occasionally in the data and refers to times when children are discouraged from connecting their interests to class lessons. In a primary setting in Ghana, a teacher expressed in a survey, 'Sometimes, I can see where the child's interest is but they still have to do the right thing and if they have other interests then they should do it in their own time and not part of the class assignments.'

Overall, Build data reinforced the importance of connection to experience, and adults regularly and intentionally referred to children's experiences when learning new concepts. Most teachers we spoke to said they used this strategy in their teaching, and most caregivers also practised this type of interrogation at home. Build data gave us examples to include in the tools, such as 'Teacher connects new concepts to real-life or everyday experiences (i.e. showing relevance of lessons outside the classroom or to real life)' in the ECE Classroom Observation Tool. In the Primary Classroom Observation Tool, a similar item was included: 'Teacher orally connects concepts in the lesson to the students' interests, background, or life outside the classroom.'

5.1.3 Support for Exploration

Primary teachers present opportunities for exploration through learning outdoors, art, and probing questions during discussions. In a focus group in Ghana, one teacher noted, 'Nature walks during science lesson[s]: the students go outside and walk around to observe and then the students come back in the classroom to discuss living and non-living.' Data from Kenya also show that children want to learn more about different topics and that adults facilitate their creativity and invite children to express, contribute, and share. For example, a teacher noted in a focus group, '[I] encourage the learners to be creative – such as exploring the environment to find tools and materials they need.' In primary settings in Colombia, teachers also facilitated class discussions to encourage child exploration. Observation data there included this example: 'The block game was an event where the children were actively manipulating materials, trying new ways to build, they were motivated to build higher and higher towers. They stayed focused and interacted with each other.'

In early childhood settings, data show that children explore their environment and specific items within it and engage in role-playing to understand new concepts and extend learning. In a focus group in Colombia, one teacher said, 'I also think that children at this age are learning and picking things up, so if it is a new story or a new activity, they get involved all at once.' Caregivers in Colombia also support child exploration. In a survey, one caregiver reflected on an activity that she and her two-year-old child did together: 'She loves the foam, she plays with it like exploring it with her hands and applies it on her body.' Exploration is similarly facilitated by ECE teachers in Jordan, as shown in classroom observation notes: 'The teacher brought a frozen water bottle so that the children could see the snow and gave them the bottle so that they could touch it'; and in a focus group, one teacher reported, 'Students enjoy going out to the garden and seeing the plants, removing them from their original place, seeing them in reality, holding them, feeling their parts, redrawing and colouring them according to each student and the colours he/she loves and collected.'

As seen with support for agency, teachers sometimes restrict children in the classroom, limiting their opportunities to explore. This was noted in primary settings. For example, a teacher interviewed in Ghana said, 'The students feel happy when they follow what the teacher is doing and when they all have materials to do what the teacher is doing; they absolutely like it when they are given instructions; if students are given materials/assignment but with little instruction or asked to try/explore on their own, they won't and they will come to the teacher with questions about how to do it or ask the teacher to show; there may be a few who can/would try on their own.' A similar sentiment was noted by a teacher in a focus group in Kenya: 'They might try to play/mess around instead of staying on task so teachers should provide structure and discipline.'

Overall, therefore, support for exploration across the ages is extended through opportunities that encourage children to be curious, inquisitive, tactile, and sensory. This led us to generate the item 'Teacher models curiosity to lead children to ask questions and/or gather new information' in the ECE Classroom Observation Tool. In the Primary Classroom Observation Tool, we included the item 'Teacher gives explicit statements to encourage students to continue to explore the concept.'

5.1.4 Support for Problem Solving

In primary settings, adults probe children to consider their approaches to problem solving, and children use one another as resources when they are trying to find a solution to a problem. Observation notes from a primary classroom in Ghana show teachers supporting problem solving: 'The teacher supports students to revise their sentence[s] (orally) when

they [make] a mistake. The teacher did not give the correct answer but prompted the students to think about what they said and how to fix it.' In classrooms in Kenya, data show that adults facilitate space for children to turn to one another when they cannot solve a problem on their own. This strategy was mentioned by a teacher in a focus group: 'They like learning together and from one another; when they see several people doing one thing then it is easier for them to pick it up or learn the concept; even if the teacher first explains and then they learn it in groups or as a class then they are more comfortable with the concept because of the shared learning.' A caregiver in Kenya similarly noted in a survey, 'When they do peer learning and revise what was taught at school, she will help on what her friends do not understand.' An observation of a primary classroom in Colombia shows how teachers use strategies to support problem solving skills: 'In some cases she [the teacher] lets them count and perform the operation alone, in other cases she counts with the children and guides them step by step through the exercise, even bending down to their level or height to interact with them in those moments and manipulate the adding machine.'

Academic problem solving was less visible in the early childhood data compared to the primary data. An example shared by a Colombian teacher in a focus group illustrates how teachers are more hands-on when it comes to problem solving: 'You start doing the activity and get involved yourself. If you have to get down on the floor, spin around, you do it first and then the children. If the child could not do it, I go back and repeat it with him. You get very involved with them.' This approach to problem solving is also apparent in the Jordan dataset. For example, observation notes from an ECE classroom report, 'The teacher began to correct the activity for each student by sitting next to him on his seat, praising him and helping those who did not know or made a mistake solving it.'

Negative support for problem solving surfaced in the primary datasets. For example, support for problem solving is impeded by teachers when children arrive at an incorrect answer and teachers dismiss students' learning process. According to observation notes of a primary classroom in Ghana, 'Student attempts to give a response or try to read the card; the student is scolded in front of the class when [student] gives the wrong answer. Teacher gives negative remarks when the student gives [the] wrong answer.' In addition, a lack of problem-solving support from adults might result in child disengagement. For example, a teacher interviewed in Kenya reported, 'Bored for those who didn't get the concept or a chance to participate; enjoyable for those who did; bored because they feel neglected.'

Overall, Build phase data helped us understand different strategies for supporting problem solving. In some cases, adults provided direct support for problem solving, guiding children through the process. In other cases, children were given problems to solve with limited support. Teachers also emphasised the importance of peer support in problem solving. In all cases, it was important for children to have some space to problem solve, even in scenarios where the teacher needed to intervene. Often, both teachers and caregivers provided support only after the child had an opportunity to find a solution without help. These findings informed items such as 'Caregiver allows the child to figure out how to do something by themselves when stuck by a challenge' and 'Teacher asks children to make predictions' in the ECE Classroom Observation Tool. In the Primary Observation Tool, we included the item 'Students try different solutions (iteration).'

5.1.5 Support for Social Connectedness

Teachers foster peer connection and social connectedness through dynamic and engaging classroom activities. Notes from a classroom observation in Ghana reflect this: 'Using an interactive activity ... the students can come to the front and move around; students were learning vocabulary words as part of the lesson but were able to be involved in learning the

words by matching the cards and then working to figure out the pronunciation together.' In Kenya, data show that peer learning is facilitated via group work and that caregivers and teachers are responsive to children and engage in communication with them, building positive rapport within the adult-child relationship. For example, a teacher reported in a survey, 'I give learners an opportunity to learn something new from their friends and peers,' demonstrating the role of the teacher in fostering connectedness and relatedness in learning activities. Similarly, observation data from a primary classroom in Colombia show teachers' efforts to support social connectedness: 'The teacher runs and crosses the room and takes the hand of a child who has not participated with the others in the activities, in the other hand she takes a chair and places him at the table where there were only three children and gives him one of the children's drawings.'

In both Kenya and Ghana, teachers said that students were motivated by being part of a group. One teacher in Kenya said, 'There is a sense of belonging that makes them feel good.' This may be one way in which students are motivated by working together on tasks – it helps reinforce a sense of togetherness.

In early childhood settings, adults build positive relationships with children and facilitate prosocial behaviours and interactions among children and between children and other adults. In a focus group in Colombia, a teacher commented on a child's relationship with the teacher and classmates: 'When we are all together, we play together. Then we are ready to play freely; and even they have like, for example, we both like soccer so they go to play soccer; affinities; so that's how it happens. There are those who like to assemble, there are those who like to prepare. So there is also like that taste, and in the same way they understand each other and share more with each other.' An observation from an ECE classroom in Jordan shows how teachers facilitate prosocial behaviours and interactions during lessons. For example, the observation notes indicate, 'The teacher distributed cardboard sheets that [had] two letters on it to each student and made them communicate with each other to form a meaningful word.'

At times, however, teachers discourage participation, disrupting prosocial behaviour and hindering social connectedness in the classroom. As noted in a primary observation in a classroom in Ghana, 'The teacher tells the students to put their heads on the table while she writes the assignment,' perhaps suggesting that social connectedness among students is being blocked by the teacher in order to maintain order or control.

Overall, prosocial behaviour was therefore prevalent in non-WEIRD settings, and both teachers and caregivers wanted to foster peer-to-peer connection. Engagement of children with other children was a priority at home and at school and only seen as a minor distraction. Children were generally allowed to participate in activities and playfulness to create bonds and connections with those around them. These findings informed the tool design: in the ECE Classroom Observation Tool, we included the item 'Teacher encourages prosocial behaviour among children (e.g. sharing, friendliness, and affection).' In primary schools in Kenya and Ghana, it was important for students to feel a sense of belonging, which motivated them to work together with peers on tasks in the classroom. This observation is reflected in the item in the Primary Classroom Observation Tool titled 'Teacher discusses or otherwise creates a sense of student/class togetherness.'

5.1.6 Support for Positive Emotional Climate

In primary settings, teachers facilitate an encouraging environment, and adults are attentive to children. In a focus group in Ghana, one teacher reflected on how they foster a positive emotional climate: 'Encourage children even when it is wrong; teacher gives

recommendations to the students about their performance and if it was good, average, fair and how to improve; teacher prevents [criticism] from other students.' In Kenya, data show that adults are attentive to children and have deep care towards students. A teacher reported in a survey that they 'build confidence of the weak learners,' demonstrating their interest in children as learners. Focus group data from caregivers in Kenya reflect adult responsivity as well. For example, one caregiver highlighted the following: 'Parental involvement; when the parent is involved for example the dad will sit down when the daughter asks to draw him and she feels as if the dad is more involved in the activity.' Data from a teacher focus group in Colombia show teacher attunement and responsiveness as well. One teacher said, 'I always talk to the children a lot. When a child doesn't want to work something is going on, so I call them to find out what is going on. If the child has been working and suddenly doesn't do it anymore, it is important to know what is going on, to talk to them and for them to have the confidence that they can talk to the teacher, that motivates them and makes the class to their liking.'

There is an attentiveness to taking a holistic approach in ECE classrooms; teachers consider the child's whole environment. A teacher in Colombia shared the following reflection in a survey: 'The teacher's motivation, attitude, is important. If one is sad, that is transmitted to them. It must convey joy, enthusiasm for wanting to do things. And love, I try to be very affectionate with them, sometimes one must be firm, serious, with some.' A caregiver of a 12-month-old interviewed in Colombia echoed this warmth: 'I have to hug her and sit with her, and then I have to bathe with her.' Similarly, data from Jordan reveal a high level of warmth in terms of tone of voice, how they treat children, and enthusiasm. A teacher in a focus group noted, 'The child feels the importance of himself and his self-esteem through the attention and care given to him from the teacher.'

Examples were found in the data where teachers contributed to a negative environment (e.g. humiliation and shame) in response to students' errors and were unresponsive to students' requests or needs. An observation in a primary classroom in Ghana illustrates this: 'The teacher asks the class to recite an answer in unison; one student give[s] the wrong answer and is heard by the teacher; the teacher requests the entire class to shame (shout 'shame, shame') to the student who gets it wrong.'

Overall, adults allowed for a positive emotional climate where they displayed care, support, love, and respect for children's needs. This display was quite explicit and easily informed items such as 'The caregiver maintains a positive stable attitude/emotion throughout the interaction and tolerates the varying emotional state of the child' in the 0–2 Caregiver-Child Observation Tool. In the Primary Classroom Observation Tool, we included the item 'Teacher provides a risk-free environment for participation.'

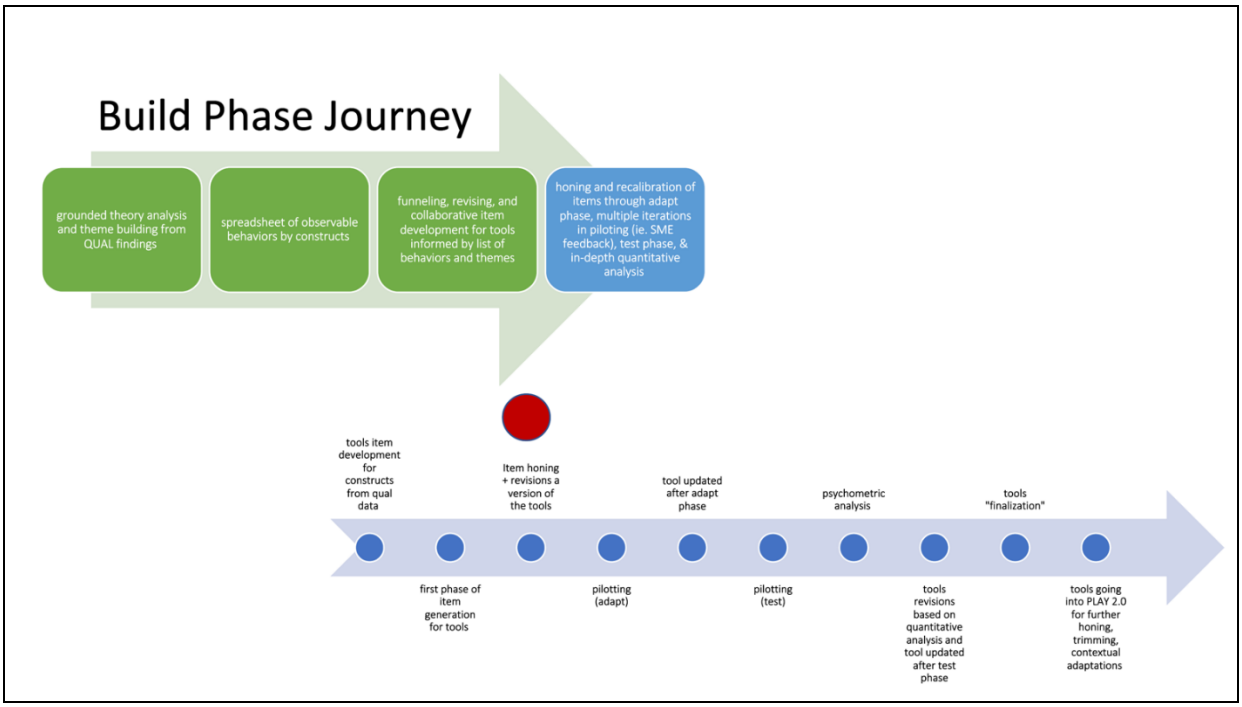
5.1.7 Summary of Analysis by Construct

Overall, the qualitative data supported our hypothesis that children's engagement in learning is supported by adults and that the type of support provided is consistent with the six constructs proposed in our conceptual framework. Data led to the extension of some of our constructs. There was evidence that adults engage children in work-like play – as explained in Section 5.1.8 – which helped expand the characterisation of the 'support for agency' construct and the 'support for connection to experience' construct. The data also led to the expansion of the 'support of social connectedness' construct to include support for prosocial behaviour (such as sharing and friendliness), as well as building a sense of togetherness among children. This sense of togetherness was important in other constructs, too. For example, teachers said that problem solving was often most effective when peers collaborated in the process.

A key challenge to emerge from the data was presented by the apparent contradiction between interview data and teacher practice in the two African countries. Although caregivers, teachers, and children emphasised that children were motivated by autonomy in carrying out tasks, several teachers felt that explicit direction from teachers was required for students to be able to act independently. One implication of this finding is that the PLAY observation tools should be sensitive to relatively subtle expressions of child agency in the classroom.

It is important to note that the qualitative Build phase made up only a portion of the journey of tool development, so we cannot credit the qualitative phase alone for the development of the constructs and items and how they have manifested in the final toolkit. **Figure 3** provides a visual illustration of the journey of Build data and the incorporation of relevant findings into subsequent phases of this project. The red dot indicates the connection and subsequent transition from the Build phase into other phases.

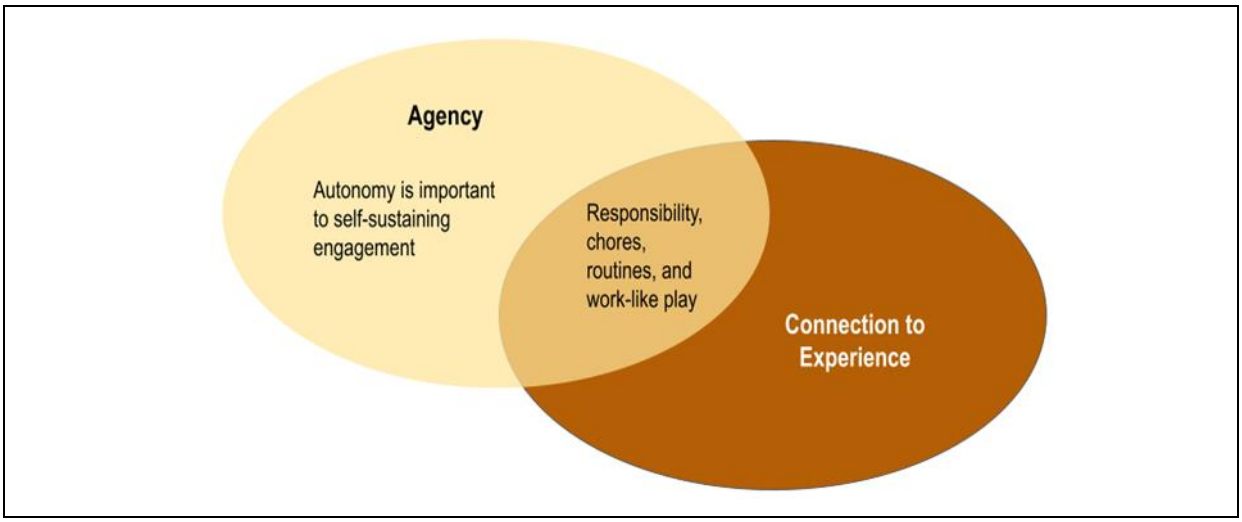
Figure 3. The role of the Build phase in the overall process of tool development



5.1.8 Constructs across Contexts

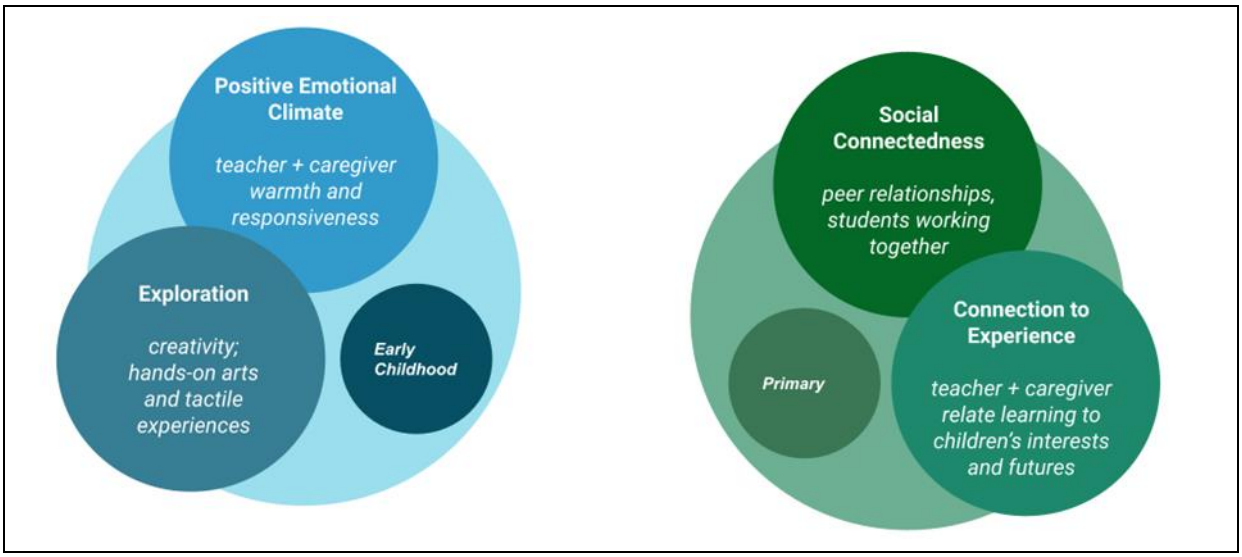
Across the countries, overlapping themes surfaced in the data in relation to the constructs. Specifically, agency came out in the Build phase findings in each country. While agency was conceptualised and observed differently in each country, data suggest that allowing children some sense of autonomy (however that may be defined in that specific context) is important for self-sustaining engagement. Another theme illuminated by findings is that responsibility, chores, routines, and work-like play are important ways that adults engage children at home and at school. This theme was found in Kenya, Ghana, and Colombia and may relate to agency as well as connection to experience, as shown in **Figure 4**.

Figure 4. Agency and connection to experience



In addition, we found that some themes are specific to early childhood or primary contexts (**Figure 5**). Specifically, in Colombia's and Jordan's early childhood contexts, we saw more of a focus on warmth and adult responsiveness in relation to the construct *positive emotional climate*. Another early childhood-specific theme to come out in analyses was a focus on caregivers and teachers to support creativity and to foster *exploration* through the arts and hands-on experiences. In primary contexts in Kenya, Ghana, and Colombia, data reveal an emphasis on peer relationships and working together, either formally coordinated by adults or organically through child-initiated activities, perhaps related to the construct *social connectedness*. In relation to the construct *connection to experience*, adult support at the primary-age level focuses on making learning interesting and relatable to students – either by connecting it to their current interests or by showing a topic's applicability to their future.

Figure 5. Key themes emerging by age group



5.2 Country-Level Overview: Additional Themes

Insights from qualitative data collection also suggest several place-specific themes. We do not propose that these themes be adopted as new constructs; rather, we suggest that

consideration of these themes and how they relate to the dimensions of the conceptual framework are important for providing a full picture of the findings within each context.

In **Figure 6**, we provide a summary of findings and additional themes found across contexts as well as in each country.

Figure 6. Additional themes emerging in each country

5.2.1 Kenya

Children are engaged in activities **related to their interests** and intrinsic motivation is reflected in tasks that children enjoy. Activities that children are passionate about and have an expressed interest in may be important in fostering self-sustainment.

'They feel interested when it is an activity that is more practical; when they are involved or active in the learning then it doesn't feel like abstract knowledge.' —teacher focus group

With regard to positive emotional climate and support for social connectedness, '**encouragement/praise**' and '**recognise**,' more specific than praise, provide an opportunity for students to demonstrate knowledge to others and is captured by adults appreciating, acknowledging, and valuing the child.

'I bestow upon them leadership roles on a rotating basis [so that] they feel appreciated and motivated.' —caregiver survey

Adults **provide materials** and ensure that resources are available to children so they can engage in playing and learning activities.

'Providing materials, ensuring materials are enough for every learner.' —teacher survey

Adults **allow space and time** for children engaged in a particular activity; because of this, children have the opportunity to do a task or activity or to explore an interest.

'Allow or create time without being biased.' —teacher focus group

Adults try to create an **equitable environment—**one that is accessible to all children and inclusive.

‘Providing [an] enabling environment and responding to their needs.’ —teacher survey
Teachers create a sense of **healthy competition** and use a **reward system** as a way to foster engagement.

‘A sense of accomplishment, everything is a competition so when they do it the best they are recognised or praised then they feel accomplished.’ —teacher focus group

5.2.2 Colombia

Adults **value agency and enable children to be leaders in their own learning**. In addition, early childhood data show that adults are ‘hands-off’ but attentive to children.

‘One has to support them in whatever they want and it is not always to invite them or tell them what to play. It is important that they create their roles.’ —ECE teacher survey

In line with the ECE findings, primary teacher data show that children have a say in what and how they learn.

‘That is what is important, to start from the activity, what the child likes, what is the context or situation of the child and how he/she wants to learn.’ —primary teacher focus group

Children express creativity in play.

‘But there are also moments and space for play, so that they can enjoy and have their own game, they can create their own experiences.’ —ECE teacher survey

Primary teacher data demonstrate that children make creative connections in learning.

‘For example, I work with the children putting words together, so I work a lot on the syllables, I put the words together with them and that word they told me, they write it in the notebook. One child told me, ah, that's like a puzzle. The children themselves show us which path to follow.’ —primary teacher focus group

Caregiving is not limited to the mother; **extended families** – including grandparents, aunts and uncles, older siblings, and cousins – **are involved in caregiving**.

‘My oldest son takes care of her when I have to go out. He studies virtually. He gives her dolls, gives her food. She enjoys everything with him. My oldest son plays a lot with her, although he is rough, for example they play with their fists, then she fights him (laughs). Or he teaches her things like pretending to be asleep.’ —caregiver survey (24 months old)

Household work, chores, and routines incorporate play and learning. Data show that adults use chore-like tasks to teach responsibility and how to care for oneself and others.

‘In the kitchen, when she feeds me, I tell her that we have to pick up the dishes [in the game], and when she eats [at feeding time], when she finishes, she does the same, she picks up the plate and takes it to the kitchen; sometimes she doesn’t have enough and leaves it in a mezzanine downstairs, when she leaves food she calls me and I receive it. Playing kitchen is very useful for her because she learns to be tidy. With the babies, with the dolls she has learned to be more delicate. I tell her soft, soft, it hurts and she is more careful, because before she was very rough. Now she sees a baby and is more gentle.’ —caregiver survey (18 months old)

Teachers have a **clear intent on creating a positive learning environment** in which children can thrive. Teachers use different strategies and multiple methods to engage children in learning.

‘Thinking about children has to do with the fact that what is planned must be for the child to understand, so that it can be useful for his life; thinking about children has to do with the fact

that if we teach them to add and the child does not understand, we must look for another method. We cannot limit ourselves to only one method.' —primary teacher focus group

Teachers are thoughtful and devoted to engaging children in learning and play. ECE teachers in particular also change the scenery as a strategy to re-engage children.

'Several children run and some chase each other laughing. It is a time where there is no order, even some chaos is perceived. "I'm going! the teacher told them, then she asked them, "Let's go, [all in] a row. Let's go for a walk, if you behave, we go to the park." The children immediately made an orderly line in front of the door.' —ECE classroom observation

Teachers frequently refer to the use of **tactile materials and physical manipulation of objects** when describing activities they do in the classroom. The importance of children learning through physically exploring, and specifically using their hands with various materials, relates to support for connection to experience in ECE and primary settings, as well as support for agency in primary settings.

'It is important that the activities are meaningful, that they can manipulate things, that they can touch things. I did a fruit salad activity. So I let them taste what is acidic, what is bitter. [When] they can manipulate things, that makes it very meaningful for them.' —ECE teacher survey

'During the game with the tiles, they look for different letters, manipulate the tiles in different ways and change the order or organisation of the tiles on the table.' —primary classroom observation

5.2.3 Ghana

While data show evidence of support for exploration, **teachers generally endorse providing explicit instructions to students before trying something new**. Overall, teachers report that it is necessary and beneficial for students to understand a concept before attempting to learn about it on their own. This may reflect a perspective on the importance of arriving at a particular answer instead of placing value on the learning process.

'If students try without the teacher's help, they can get it wrong; if the teacher doesn't give instructions, the students feel that the teacher is not helping them or showing them.' —teacher focus group

Students **utilise one another as learning resources** in the classroom; they turn to one another to elicit, offer, and follow through in helping one another, demonstrating agency in the classroom. This learning strategy is both student and teacher led.

'The group members will be able to correct each other and so the students feel like others have their back but if it is an individual assignment then you don't feel as free to try and do what you want.' —teacher survey

Teachers make **lessons relatable and applicable to students** frequently and organically in classrooms.

'Teacher gives space for students to decide on examples that are most meaningful/interesting to them.' —teacher survey

Teachers identify **motivation as being at the root of their decision-making** in the classroom. Specifically, teachers take child motivation into consideration when thinking about how or why they do certain things in the classroom, and they use child motivation as a reason to justify their methods of teaching.

‘Children are motivated when they love the way/methods the teacher teaches them; the students are eager and more motivated to participate if the teacher uses teaching methods the students like; using a variety of methods, especially if the teacher tries one method but the children don't understand and then try another one to help the students better (they will be more motivated); when the teaching/lesson involves [active] activities and allow for a lot of participation then the children show more enthusiasm.’ —teacher survey

Based on teachers’ explanations of their own teaching processes, it may be important to **rethink and redefine the construct ‘support for agency’ in this context**. This stems from the finding above (teachers favour providing instructions over open-ended exploration). Teachers support children through explicit instruction in order to equip children to do something on their own. Therefore, support for agency in this context may not present itself as anticipated.

‘Children would feel happy to participate as she instructs them; they would feel competent or as if they now have the knowledge on how to do it because the teacher has showed them how to do it.’ —teacher survey

5.2.4 Jordan

Teachers have a **strong bond** with their children. Caregivers use multiple methods to keep children aware and engaged in the educational process.

‘The teacher mentioned that the parents communicate with her and tell her about what they are facing with their children at home, where the teacher in her own way tells the children in the school about it and the students do what the teacher asks.’ —teacher focus group

Teachers value agency and encourage children's autonomous learning and children taking control of their choices.

‘The teacher gives the child the freedom to choose and supports him in his choices to make the child feel comfortable and safe.’ —teacher focus group

Teachers use **storytelling** as a way teachers to teach children and keep them engaged in the learning process in a fun way.

‘Then she started telling a story entitled ‘Who Heals Trees’ by modelling enthusiasm as she changed her tone of voice according to the characters of the story and imitated their voices to get children's attention and to increase their enthusiasm.’ —observation data

The use of **drawings and art objects** during activities are frequently mentioned in the data as ways teachers engage children to facilitate learning and playing.

‘Colours, forming letters and colouring them with modelling clay, emptying letters and colouring them with colours, mosaic painting, cutting and colouring the flag.’ —teacher focus group

5.3 Build Phase Conclusion

In summary, the Build phase illustrated the ways in which playful learning was experienced and understood in the respective country contexts. Broadly, this phase ended with two significant take-aways encompassing what was found at the construct level, country level, and age level, whether similarly, differently, or interestingly.

The first take-away centres on informing existing constructs to reflect the field. The second take-away concerns the conceptual reorientation of how we thought about our framework.

The operationalisation of what that specifically means for some constructs and the overall conceptual framework is detailed in Section 7.2.

Overall, it is important to remember that while the findings varied across age range and context and some constructs were informed differently than others, the Build phase significantly endorsed the applicability of constructs across the range of countries. In this way, the Build phase helped us add new dimensions to the existing constructs or expand them to better reflect the various contexts, but no new constructs were added to the existing six after the end of the Build phase.

6. QUANTITATIVE DATA COLLECTION METHODS

6.1 Overview

Quantitative data collection activities were conducted in the Adapt and Test phases of this project. These efforts began with small-scale pilot activities (used to better understand administration procedures and generate basic descriptives for item responses). Findings from small-scale pilots were then used to revise instruments before conducting larger-scale 'final' data collections (used to assess the reliability and validity of the tools, as well as the factor structure of the instruments).

For each participant group, observation and survey instruments were designed to measure the quality of adults' support for learning through play at the setting level (i.e. homes, centres, or schools). An overview of the tools by age and participant group is displayed in **Table 4**.

Table 4. PLAY tools by age and participant group

Age group	Participant group	PLAY tools	Country
0–2 years	Caregiver-child	• Caregiver-Child Observation Tool • Caregiver Survey	Colombia
3–5 years	Caregiver-child	• Caregiver-Child Observation Tool • Caregiver Survey	Colombia
	Classrooms	• ECE Classroom Observation Tool • ECE Teacher Survey • ECE Classroom Inventory*	Colombia, Jordan, Ghana**
6–12 years	Classrooms	• Primary Classroom Observation Tool Primary Teacher Survey • Primary Student Survey • Primary Classroom Inventory*	Colombia, Ghana, Kenya

* This classroom inventory is taken within what we refer to as 'supplementary items' in the early childhood observation tools.
** The ECE data collection in Ghana focused on the classroom observation only.

Since the purpose of this activity was to develop a final version of each instrument for use across contexts, data collections were staggered to allow for iterative adaptation of the instruments (with findings of each administration leading to revisions prior to the ensuing administration within and across countries). These data collection activities occurred from December 2021 to July 2022 in Colombia, Ghana, Jordan, and Kenya, as shown in **Table 5.**

Table 5. Quantitative data collection activities by age, participant group, and stage

Data collection stage	Age (participant group)			
	Caregiver-child		Classrooms	
	0–2 years	3–5 years	3–5 years	6–12 years
Pilot	February 2022 (Colombia)	February 2022 (Colombia)	• February 2022 (Colombia) • March 2022 (Jordan)	• December 2021 (Kenya) • March 2022 (Ghana)
Full scale	March–May 2022 (Colombia)	March–May 2022 (Colombia)	• April 2022 (Jordan) • March–May 2022 (Colombia) • June 2022 (Ghana)	• March–May 2022 (Colombia) • May 2022 (Kenya) • July 2022 (Ghana)

6.2 Sample Size Estimates

Small-pilot sample data were not intended to undergo rigorous quantitative analyses. Therefore, pilot sample sizes were estimated based on best practices for the intended purpose of identifying major issues that may exist with items and administration procedures. Proposed minimum sample sizes for piloting are displayed in **Table 6** (final sample sizes are included in the Pilot Data Collection subsection, below).

Table 6. Proposed minimum sample sizes for small-scale piloting

Method	Minimum sample size ranges per age group		
	Parents / caregivers	Teachers	Students
Surveys	25–50†	25–50	50–100‡
Observations	J=25–50†	J=25–50	N/A

† Parent/caregiver samples are applicable only for early childhood age groups.
‡ Student survey samples are applicable only for primary school age groups.
N/A = Not applicable
J = Sample size at the setting level; observations will occur at the classroom level

Larger, full-scale data collection, on the other hand, required sufficient sample sizes for psychometric assessment of the tools. Power analyses were conducted based on initial draft tools, consisting of six constructs, with six items per construct. The additional parameters were as follows: RMSEA values for the null hypothesis = .05; RMSEA values for alternative hypothesis = .08; type I error rate = .05.

This led to final sample size estimates shown in **Table 7.** In this table, overall power refers to the probability of correctly rejecting a confirmatory factor analysis (CFA) model in which six constructs explain variation on the 36 items. Power per construct refers to the probability of correctly rejecting a one-factor CFA model for each set of six items. While this shows that a sample size of 150 would be sufficiently powered to test whether the data fit our conceptual model, this sample size allows for considerably less power per construct (than it does overall). Our analytic approach therefore included pooling across grades (within the primary sample), as well as across countries (e.g. Kenya and Ghana) in order to produce more reliable estimates. Final sample sizes are provided in the data section below.

Table 7. Sample sizes and power calculations for full-scale data collection

	Pooling		
Single age group for a single country	**Two age groups within a country**	**Across two countries within an age group**	**Across two age groups and across two countries**
N = 150	N = 300	N = 300	N = 600
Power per construct = .21	Power per construct = .39	Power per construct = .39	Power per construct = .65
Power overall > .99	Power overall > .99	Power overall > .99	Power overall > .99

6.3 Tool Administration

As previously noted, there is both an observation and a survey measure for each participant group and setting (with only a survey measure for primary school aged children). All tools were administered by trained data collectors during home or school visits. **Table 8** provides an overview of the tool administration for each instrument.

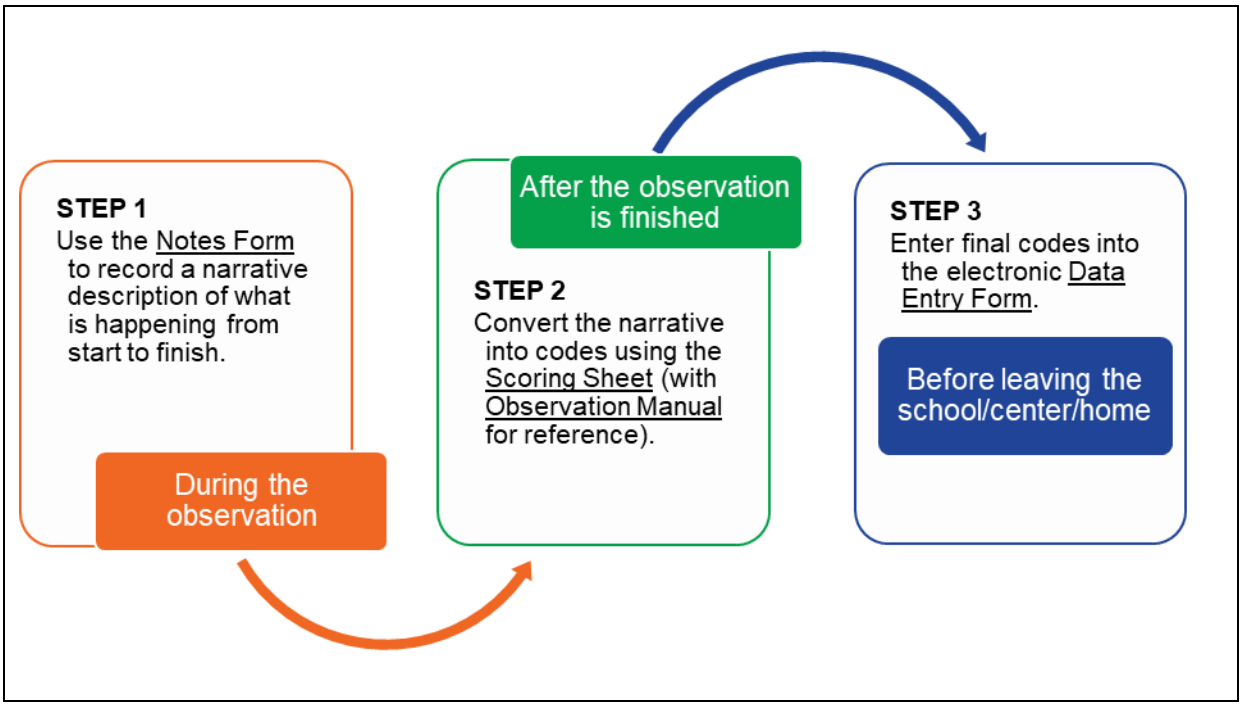
In order to improve data quality, as well as time and cost efficiency, all PLAY data were recorded electronically, using Tangerine or Kobo. Electronic data collection simplifies the preparation and implementation of fieldwork by reducing assessment times, mitigating measurement and data entry errors, and eliminating manual data entry from paper forms to the database for analysis, thus making results more readily available.

Table 8. Tool administration by setting, participant group, and instrument type

	Caregiver		Classroom			
			Observation Tool		Primary Survey	
	Child Observation Tool (0–2 & 3–5 years)	**Survey**	**ECE (3–5 years)**	**Primary (6–12 years)**	**ECE & Teacher (3–5 & 6–12 years)**	**Student (6–12 years)**
Time	15 minutes	20–25 minutes	3 hours (30–45 minutes if video)	45 minutes	30 minutes	20–25 minutes
Setting	Community centre or home	Community centre or home	Classroom in a centre or preschool	Classroom in a primary school	ECE centre, preschool, or primary school	Classroom in a primary school
Materials	Observation rubric, contextually appropriate toys, pen + paper	Survey (tablet)	Observation rubric, pen + paper	Observation rubric, pen + paper	Survey (tablet)	Survey (tablet)

As shown in Table 8, surveys were administered via tablets, while observations were directly scored on paper-based observation rubrics (during the observation) before being entered electronically. The general steps used for conducting observations (shown in **Figure 7**) were similar across the ECE and Primary Classroom Observation Tools.

Figure 7. Steps for conducting observations



In order to support data collectors in the field, each tool was accompanied by a user protocol that included guidance and a short script for administering that tool. The protocols were designed for data collectors to refer to and use during data collection, guiding them step by step through the process of administering each tool.

Each protocol contained (1) the tool in question; (2) a list of all of the accompanying data that the data collector must enter that is supplemental to the main sections of the tool (such as data collector name and ID; school or household ID; child ID; and date and time); (3) instructions for introducing themselves and the project; (4) a guide to confidentiality procedures; (5) details of the data collection process and the specific tool being administered; (6) overall instructions for administering the tool; and (7) general guidelines to take into account throughout.

6.4 Pilot Data Collection

During the Adapt phase, pilot data collection was completed for primary-level instruments in Ghana and Kenya.

In Ghana, this activity focused on the piloting of instruments for the 6- to 12-year-old age group (in Ghanaian primary schools). The data were collected over a five-day period from February 28 to March 8, 2022, by a group of 12 locally trained data collectors. Data collectors were placed into teams of four, with each team visiting one school per day.

In Kenya, the activity similarly focused on the piloting of instruments for the 6- to 12-year-old age group (in Kenyan primary schools). The data were collected over a five-day period from December 6 to December 10, 2021, by a group of 12 locally trained data collectors. Data collectors were once again placed into teams of four, with each team visiting one school per day.

A total of 15 schools were visited for this phase in each country. In each school, four classrooms were randomly selected for a Classroom Observation, Classroom Inventory, and

Teacher Survey. Two students were also randomly selected to be interviewed in each of the selected classrooms.

Final sample sizes are displayed in **Table 9.**

Table 9. Final sample sizes for Ghana and Kenya primary pilot data

Number of			
	Surveys		
Total schools	Student	Teacher	Classroom observations/ inventories
15	120 (8 per school; 2 per class)	60 (4 per school)	60 (4 per school)

In order to ensure high levels of variability in teaching practices, a range of subjects were observed across the sampled schools (selected randomly), as shown in **Table 10** and **Table 11.**

Table 10. Subjects by grade for classroom observations in Ghana primary pilot

	English	Mathematics	Science	Social studies	Arts	Religion	Total
Grade 1	4	0	1	0	0	1	6
Grade 2	4	4	2	1	0	1	12
Grade 3	5	5	1	3	0	0	14
Grade 4	3	2	1	4	1	1	12
Grade 5	3	3	1	2	0	0	9
Grade 6	2	2	1	1	0	1	7
Total	21	16	7	11	1	4	60

Table 11. Subjects by grade for classroom observations in Kenya primary pilot

	Kiswahili	English	Math	Science	Social studies	Arts	Music	Religion	Total
Grade 1	1	1	1	0	0	1	1	1	6
Grade 2	1	4	4	0	0	0	0	0	9
Grade 3	3	4	3	2	0	1	1	1	15
Grade 4	1	1	2	2	2	3	0	1	12
Grade 5	0	2	6	2	2	0	0	0	12
Grade 6	0	0	1	0	2	0	0	3	6
Total	6	12	17	6	6	5	2	6	60

6.5 Medium- and Full-Scale Data Collection in Primary Schools

The first larger-scale data collection for the primary schools was conducted in Colombia over an eight-week period from March 15 to May 27, 2022, by a group of ten locally trained data collectors. At the primary level, this activity focused only on the piloting of instruments for the 6- to 9-year-old age group (in Colombian primary schools). The initial focus of the Colombian data collection was on ECE, but primary tools were tested in order to align with our research partner's interests. Therefore, this data collection activity at the primary level was smaller than the ensuing data collections in Kenya and Ghana.

A total of 25 public schools were visited for this phase, primarily in Bogotá (consisting of 'ordinary' schools with no ongoing play-based initiatives). In each school, three classrooms were randomly selected for a Classroom Observation, Classroom Inventory, and Teacher Survey. On average, five students were also randomly selected to be interviewed in each of the selected classrooms.

Final sample sizes for Colombia are displayed in **Table 12** and **Table 13.**

Table 12. Final sample sizes for Colombia primary medium-scale data

	Number of		
	Surveys		
Total schools	Student	Teacher	Classroom observations/ inventories
25	380 (~15 per school; ~5 per class)	75 (~3 per school)	75 (~3 per school)

Table 13. Subjects by grade for classroom observations in Colombia primary medium-scale data

	Spanish	Physical education	Mathematics	Science	Social studies	Art	Music	Religion	Total
Grade 1	8	0	11	4	1	0	0	0	24
Grade 2	12	0	7	3	3	0	0	0	25
Grade 3	8	1	9	3	2	1	1	1	26
Total	28	1	27	10	6	1	1	1	75

After revisions were made to the primary-level tools based on findings from the small-scale pilots (in Kenya and Ghana) and the medium-scale pilot in Colombia, the first full-scale data collection was conducted in Kenya. It focused on the 6- to 12-year-old age group in Kenyan public primary schools. The data were collected over a ten-day period from May 9 to May 20, 2022, by a group of 28 locally trained data collectors. Data collectors were placed into teams of four, with each team visiting one school per day.

A total of 70 schools were visited for this phase (40 in Mombasa County and 30 in Kilifi County). The schools in Mombasa were all participating in Aga Khan Foundation's learning through play intervention, while those in Kilifi were public schools without play-based interventions. In each school, four classrooms were randomly selected for a Classroom Observation, Classroom Inventory, and Teacher Survey. Four students were also randomly selected to be interviewed in each of the selected classrooms.

Final sample sizes for Kenya are displayed in **Table 14** and **Table 15**.

Table 14. Final sample sizes for Kenya primary full-scale data

	Number of		
	Surveys		
Total schools	Student	Teacher	Classroom observations/ inventories
70	1,120 (16 per school; 4 per class)	280 (4 per school)	280 (4 per school)

Table 15. Subjects by grade for classroom observations in Kenya primary full-scale data

	Kiswahili	English	Math	Science	Social studies	Arts	Music	Religion	Total
Grade 2	11	10	14	25	0	0	0	10	70
Grade 3	5	9	20	24	2	2	1	7	70
Grade 4	6	14	16	19	8	6	0	1	70
Grade 5	5	11	20	20	7	3	0	4	70
Total	27	44	70	88	17	11	1	22	280

Building off of the findings from the full-scale data collection in Kenya, a full-scale data collection was conducted in Ghana over a ten-day period from June 27 to July 8, 2022, by a group of 28 locally trained data collectors.

The school sample in Ghana consisted of 70 public primary schools in the Greater Accra region. While 48 of these schools were implementing Right to Play's learning through play initiative, 22 schools were 'ordinary' primary schools with no ongoing play-based initiatives. The sampling procedure in Ghana was the same as the one used in Kenya: data collectors were placed into teams of four, with each team visiting one school per day; four classrooms were randomly selected for a Classroom Observation, Classroom Inventory, and Teacher Survey in each school; and four students were randomly selected to be interviewed in each of the selected classrooms.

Final sample sizes for Ghana are displayed in **Table 16** and **Table 17.**

Table 16. Final sample sizes for Ghana primary full-scale data

Number of			
	Surveys		
Total schools	Student	Teacher	Classroom observations/ inventories
70	1,114 (~16 per school; ~4 per class)	277 (~4 per school)	278 (~4 per school)

Table 17. Subjects by grade for classroom observations in Ghana primary full-scale data

	Ghanaian	English	Math	Science	Social studies	Arts	Religion	Total
Grade 2	0	23	19	12	7	0	4	65
Grade 3	1	22	29	8	4	2	1	67
Grade 4	1	17	18	17	8	4	3	68
Grade 5	0	15	18	17	11	2	5	68
Grade 6	1	4	3	0	2	0	0	10
Total	3	81	87	54	32	8	13	278

6.6 Full-Scale Data Collection for ECE

All of the early childhood PLAY tools were administered in Colombia. The data were collected over eight weeks from March 15 to May 27, 2022, by a group of 17 data collectors placed into two teams: caregiver and classroom. Fieldwork was conducted in three geographical areas: Bogotá (the capital city), Caquetá (one of the departments in the

Amazonian region), and Cartagena (one of the main cities in the Caribbean region). In Caquetá, sites were selected in coordination with an ongoing project of Universidad de los Andes. The early childhood settings consisted of community homes, public and private early childhood centres/kindergartens, and public and private schools. The caregiver-child interactions team collected most data in Caquetá, while the classroom observations occurred mainly in Bogotá. Data were collected in pairs initially (i.e. over the first 1–3 weeks) and then individually.

Final sample sizes for Colombia are displayed in **Table 18.**

Table 18. Final sample sizes for Colombia ECE data

	Number of		
	Total dyads/schools	Surveys	Observations
Caregiver 0–2	182	160	165
Caregiver 3–5	146	132	128
Classroom 3–5	104	139	147

The Classroom Observation Tool and Teacher Survey were administered in Jordan. The data were collected for two weeks from April 17 to 28, 2022, by a group of 22 data collectors. Fieldwork was conducted in the Irbid region. Fieldwork was conducted in coordination with an ongoing International Rescue Committee-NYU project. The sample consisted of government preschools where the instructional approach was 'traditional' (i.e. not influenced by a play-based intervention).

Final sample sizes for Jordan are displayed in **Table 19.**

Table 19. Final sample sizes for Jordan classroom ECE data

	Number of		
	Total classrooms	Classroom observations	Teacher surveys
Classroom 3–5	111	147	111

The Classroom Observation Tool was piloted through video coding in in Ghana. A group of 11 data collectors at Innovations for Poverty Action coded videos utilising the Classroom Observation Tool over three weeks from May 24 to June 15, 2022.

Final sample sizes for Ghana are displayed in **Table 20.**

Table 20. Final sample sizes for Ghana ECE video data

	Number of videos	
	Total	Observations
Classroom 3–5	423	423

7. DEVELOPMENT OF THE TOOLKIT

7.1 Introduction and Aim

The development of the PLAY Measurement toolkit and conceptual framework was guided by the aim of the project, which was to measure 'support for learning through play.' For this project, the term 'support' referred mainly to that provided through adult-child interactions, although structural support (e.g. through physical spaces and resources) was also considered.

This section describes the process of development of the primary-age tools, followed by the development of the early childhood tools.

7.2 Development of the Primary School Tools

7.2.1 Observation Tool

The Primary Classroom Observation Tool was used first in the Adapt phase in Kenya. This version of the tool was initially challenging for data collectors because there were many items that were superficially similar. Therefore, we adapted the tool in several ways to simplify and guide the coding process. Our first attempt to simplify the tool was to incorporate several constructs into one section, with all items including a degree of student agency. The following sub-constructs were included:

- problem solving
- iteration
- exploration
- pretend play
- independence
- influence on teacher
- self-expression
- student initiation

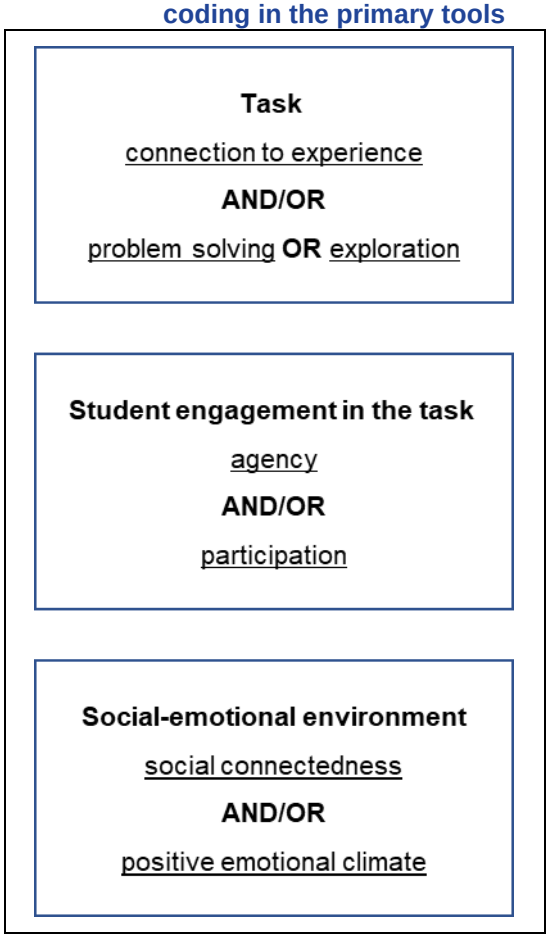
The sub-constructs begin with the most specific first and increase from items with less student agency – such as problem solving (teacher initiated and partly teacher directed) – to items with more student agency, including student-initiated interactions. The reorganisation helped data collectors code items. However, a discussion with the wider NYU/RTI team concluded that there was a value to having agency as a distinct construct rather than an organising principle for several constructs.

Our second approach to simplifying coding evolved through the Ghana Adapt phase. First, we separated the tool into two sections – constructs that related to academic tasks (connection to experience, problem solving, exploration, agency, and participation) and those that related to classroom environment (social connectedness and positive emotional climate). This distinction helped clarify the definition of items – for example, making clear that problem solving referred to academic problems rather than, say, resolving conflict between students. The constructs were further subdivided into (1) the nature of the task (connection to experience, problem solving, and exploration); (2) student engagement (agency and participation); and (3) the classroom environment (social connectedness and positive emotional climate). See **Figure 8**.

These three higher-order categories helped data collectors think through the possible options for coding and cases where one teacher-child interaction may be given more than one code, which we referred to as 'double coding.' We produced a guide to coding that included advice on when to double code. The guidelines noted that double coding is most common when the two codes are in different higher-order categories (e.g. nature of the

task), less likely when they are in the same higher-order category, and least likely when they are in the same construct. The guide also lists common examples of double coding. For example, group work may be commonly coded as both 'teacher allows freedom in approaching an academic task' and 'students work on exercise/project or towards a common goal.' Coding was simplified for one pair of constructs – problem solving and exploration – by defining these constructs as mutually exclusive. Problem solving was defined as being when the teacher has a specific goal towards which they guide child-directed activity, whereas exploration was defined as being open-ended (that is, with no specific goal in mind).

Figure 8. Categories used to guide coding in the primary tools



In addition to assessing whether an item was observed in the classroom, we included a quality metric for each item. In the Kenya Adapt phase, we used a combination of effectiveness and frequency as the quality metric for all items. Data collectors recorded observations in five-minute cycles and recorded whether each item was observed in that five-minute period on a three-point scale (not observed; an ambiguous/low-quality example of the item observed; and a clear/high-quality example of the item observed). This resulted in a score of the total number of cycles in which each item was observed, at each level of effectiveness. We found that this process was quite burdensome for data collectors, and data suggested that getting reliability on this frequency measure would be difficult. We also found that items were observed only once or twice in a lesson, with most items not being observed at all. Thus, we concluded that a detailed frequency measure was not required. In the next iteration – the Ghana Adapt phase – we dropped the approach of tallying the total number of times an item was observed. Instead, data collectors coded each item once for the total lesson. They rated each item on a three-point scale, where items could be unobserved or observed with low or high quality. The quality metric varied by item: frequency (how often the item occurred), effectiveness (whether the item was a clear example of the target behaviour), or participation (how many students were involved). Data showed that many of the items with an effectiveness metric did not capture variation in the observed behaviour, likely because the definitions for high effectiveness and low effectiveness were complicated or nuanced; data collectors tended to code items as either absent or highly effective. As a result, we amended the tool so that most items had a frequency metric that provided a simple and reliable way to score the items. Additionally, we standardised the quality levels for the metric frequency as 'not observed,' 'observed once,' and 'observed more than once' and for the participation metric as 'not observed,' 'less than half the students participate,' and 'more than half the students participate.' We retained an effectiveness metric for a small number of items that are likely to occur only once during a lesson and where we judged that quality was more important than quantity of interactions.

We experimented with different ways of recording data. The procedure favoured by data collectors involved three steps. First, they wrote a narrative description while they were observing the lesson. Second, after the lesson concluded, they converted the narrative into codes and recorded the codes on a paper scoring sheet, with a detailed manual on hand as a reference. In the final step, they entered the codes into a digital form using Tangerine.

Fine-tuning of the tool involved providing several examples of each item – including examples of high and low effectiveness where appropriate – and a brief description of each item for ease of reference.

The choice and wording of items evolved through cognitive interviews and quantitative data collection. We found that genuine problem solving – defined as children solving novel problems without previously being taught a method to solve them – was difficult to observe. For the problem-solving construct, we included items to capture teachers' scaffolding of children's attempts to solve problems. As discussed above, genuine decision-making by children was hard to observe, so items were refined or added to capture more subtle markers of agency.

Learnings

- *The distinction between items should be clear on first reading; subtle shades of meanings tend to get lost.*
- *Agency should be a distinct construct rather than an organising principle because it would otherwise be impossible to understand whether agency operates independently or only in conjunction with other constructs.*
- *Categorising constructs helped clarify the definition of items and helped data collectors deal with 'double coding' incidences. The first distinction was between academic- or classroom environment-related constructs. Academic was then further subdivided into the nature of the task and student engagement in the task.*
- *Data collectors found conducting five-minute observation cycles with tallies of the behaviours to be burdensome. This method also did not produce reliable data. Coding each item once for the entire lesson worked better.*
- *A frequency metric was found to be the most reliable way to score quality. Furthermore, standardising quality scales across all items was helpful for reliability.*
- *Data collectors preferred to record data by (1) writing a narrative description during the lesson, (2) converting this into codes on paper, and (3) entering the codes into a digital form.*
- *It was difficult to observe children solving problems on their own. Capturing teachers' scaffolding of children's attempts to solve problems was easier.*

For our recommendation for the final form of the Primary Classroom Observation Tool, see Section 8.

7.2.2 Teacher Survey

As part of the Adapt phase, we conducted cognitive interviews and explored the best ways to administer the Teacher Survey. Initially, data collectors read the items to teachers; however, we found that teachers preferred to read the items for themselves. The final version of the Teacher Survey included two parts: statements and vignettes. The statements portion involved printing each item on a laminated card with a corresponding printed scale.

Teachers sorted each card on a scale indicating how often they conducted the behaviour described in the item. The process was smooth, and teachers enjoyed being able to direct the activity.

The vignette-based part of the survey presented teachers with brief scenarios depicting common instructional approaches. In the first version of the survey, two types of vignettes were included – those depicting learning through play, and those depicting more traditional instruction. Results from the Adapt phase showed that responses to play-based scenarios were correlated with teacher self-reported agency. Responses to traditional scenarios showed no relation with other measures and poor psychometrics. Consequently, we dropped the traditional scenarios from the Teacher Survey. We also refined the play-based scenarios so that each scenario corresponded to one of the seven constructs. Initially, teachers were asked four questions about each scenario:

- How often do you do something like this?
- How confident are you about using this activity in your classroom?
- How effective is this activity for promoting learner engagement?
- How effective is this activity for supporting learning?

Findings showed that teachers' responses to questions 3 and 4 were very highly correlated. If teachers thought an activity was effective at promoting learner engagement, they also thought the activity was effective at supporting learning. As a result, we dropped the question about learner engagement from the tool.

In general, the Teacher Survey was easier to administer than the Primary Classroom Observation Tool, and only minor revisions were required. The wording of items was revised based on teacher feedback and quantitative analysis. Some items were revised or added to match changes made to the observation tool. The aim was to include equivalent items across the two tools. The findings from the quantitative Adapt phase data from the self-report Teacher Survey showed that agency-supported interactions were less common than other constructs. However, support for agency had the highest correlation with scores from the vignette-based portion of the tool. This suggested that support for agency was potentially an important behaviour, despite being reported less frequently than other constructs.

Learnings

- *Teachers enjoyed the interactive survey experience offered by the vignettes.*
- *Play-based vignettes were more statically reliable and meaningful than traditional scenarios.*
- *Teachers interpreted promoting learner engagement and supporting learning as virtually synonymous.*
- *Though agency-supported interactions were less common, they still emerged as statistically meaningful behaviours.*

For our recommendation for the final form of the Primary Teacher Survey, see Section 8.

7.2.3 Student Survey

We experimented with two versions of the Student Survey. In one version, students indicated their level of agreement with a single statement on a Likert scale. The second

version was a forced-choice response in which students chose to endorse one statement from a pair of contrasting statements. The forced-choice test was rejected for two reasons. First, the literature suggests that forced-choice statements are more cognitively taxing and may not produce reliable results in young children. Second, it was challenging to generate pairs of opposing statements for all items we wanted to measure. After the Kenya Adapt phase, we discarded the forced-choice test and used only the Likert scale. For this scale, we experimented with a four-point one – with response options 'agree strongly,' 'agree,' 'disagree,' and 'disagree strongly' – and a five-point one. We preferred the four-point scale because each item could be split into two questions (Do you agree or disagree? Do you agree/disagree strongly?) to help children respond and because it lent itself to the use of pictures in the response stimuli. We initially used smiley and frowny faces as the response stimuli. After a suggestion from data collectors during Colombia Test phase training, we revised the stimuli to be 'thumbs up' or 'thumbs down' symbols – one for agree/disagree and two for agree strongly/disagree strongly. The rationale for the change was that children were being asked about their agreement with the statement, not about whether the statement made them happy or sad.

The focus of revisions to the survey was to ensure that younger children could understand the procedure and the items. After the Kenya Adapt phase, we revised the instructions and added examples to the start of the survey. We simplified the language of the items where possible and opted for short statements (e.g. 'Your class is fun'). We also added optional descriptions and examples for each item that data collectors could read aloud if students had difficulty understanding an item.

Learnings

- *Forced-choice answer formats are cognitively taxing for children but also more difficult to develop and implement overall.*
- *Particularly for young children, it is important for response stimuli (e.g. pictorial representation) to accurately match the response choices. In this case, the use of 'thumbs up' or 'thumbs down' symbols – one for agree/disagree and two for agree strongly/disagree strongly – rather than smiley/frowny faces was more appropriate, as the children were being asked about their agreement with the statement, not about whether the statement made them happy or sad.*
- *Simple language and optional descriptions/examples to support item understanding are key.*

For our recommendation for the final form of the Primary Student Survey, see Section 8.

7.2.4 Training Approaches

Data collector trainings for the Adapt and Test phases were organised such that participants first built an understanding of the constructs by spending time discussing each construct's items and examples. These discussions included interactive activities such as sorting key words and role-playing classroom scenarios. Then, training activities shifted to focus on tool administration and scoring. Participants visited schools to observe actual classrooms and conduct surveys with teacher or student respondents. At the end of training, we tested participants' reliability and accuracy. Classroom observers watched and coded the same classroom video, which was scored against a gold standard and the group mode. Interviewers demonstrated administration skills as trainers observed and scored them using a performance rubric.

7.3 Development of the ECE Tools

7.3.1 Observation Tools

The early childhood observation tools consist of the classroom 3–5 tool, the caregiver 3–5 tool, and the caregiver 0–2 tool. The classroom 3–5 tool was developed over two Adapt phases in Colombia and Jordan. Findings from the Build phase were also considered during tool development. The caregiver 0–2 and 3–5 tools were administered only in Colombia and thus were fully developed there.

Classroom Tool

The classroom 3–5 tool was initially developed through a two-part Adapt phase in Colombia. In the first part, we engaged with subject-matter experts (SMEs) identified by our in-country partner, Universidad de los Andes. SMEs (n=13) all had a psychology or social science background and experience in the fields of early childhood development and education; they ranged from Ministry of Education staff to university faculty to young professionals and recent university graduates. We administered surveys to SMEs, eliciting feedback about our constructs and items in order to revise the tool for cultural and contextual relevance. For example, we asked SMEs about the likelihood of an item occurring during a 'typical' day in an ECE classroom in Colombia, whether the item was culturally relevant, and whether, as experts in the field, they would endorse the item for the given construct. We also asked open-ended questions, which included having SMEs define constructs in their own words and provide examples of items not represented by the tool that would be important to include. Ultimately, the SME structure was highly beneficial as we were able to have an engaged group of individuals with whom to have regular, iterative discussions during the tool development process. Based on SMEs' feedback, we dropped items identified as irrelevant, confusing, or unlikely to occur. We checked for face validity and moved some items to different constructs. We also added SME-generated items. At this point in tool development, the classroom 3–5 tool had the number of items per construct as shown in **Table 21.**

With the newly revised tools, we engaged a subset of SMEs (n=8) in part two of the Adapt phase. These data collectors would remain throughout the project, becoming 'senior data collectors' in the Test phase. While building capacity and having local expertise is ideal, the primary factor in this type of structure is time. Where there is higher capacity, this type of structure could be implemented more readily. Otherwise, it could still be implemented in contexts of lower capacity but with the understanding that building up knowledge and skills to train would take more time.

In the second part of the Adapt phase, we worked with data collectors via remote training to develop the rubric. We had already decided that the tool would have a four-point scale for usability and clarity, but we wanted local experts' support in developing the scoring rubric. In the first week of training, we worked together to generate exemplars per item and discussed ECE quality in Colombia – for example, what would it mean or 'look like' to score each item as a 0, 1, 2, or 3? In-depth discussions at times revealed items that were irrelevant, confusing, or otherwise not applicable, and we continued to revise the tool iteratively. The training was bilingual in Spanish and English and involved live interpretation and tool translation.

Table 21. Number of constructs in classroom 3–5 tool during Adapt phase

Construct	Total items at start of Adapt I	Number of Items		Total items for Adapt II
		Dropped or moved	Suggested by SMEs	
Agency	17	-7 (drop 5; move 2 to diff construct)	3	13
Exploration	8	-2 (+1 from agency)	1	8
Connection to experience	11	-3	1	9
Problem solving	21	-14	2	9
Positive emotional climate	9	0	4	13
Social connectedness	9	-1 (+1 from agency)	2	11

Note: The far-right column does not represent the final number of items for the Test phase, as the tool continued to be refined after Adapt II, with final changes being made during Test training.

Points of confusion and thus important topics of discussion included distinguishing child exploration from child problem solving; understanding how social connectedness and positive emotional climate are distinct from each other; and understanding what exactly is considered agency in an ECE classroom. We teased apart exploration and problem solving by determining that child exploration refers to the physical process of discovery, while problem solving is aligned to the scientific method and the process of investigation. While social connectedness and positive emotional climate seem to overlap, we came to distinguish the two by affirming that the former refers to peer relationships and perhaps learning social norms, while the latter centres on teachers modelling positive behaviours and responsiveness, among other things. And we focused on how to identify whether behaviour is agentic or the result of an indifferent teacher who allows significant free play time. We determined as a group that agency implies intentionality; there are choices, but the teacher has created a specific context for this. We also noted as a group that the teacher must be using choice authentically and that children must be choosing authentically; teachers using choice as a threat or way to control children's behaviour is not indicative of support for child agency. Enumerators in Colombia then used the classroom 3–5 tool during a video coding exercise in this Adapt phase. We calculated agreement rates and identified potentially problematic items (i.e. >1 average absolute difference on the four-point scale). This applied to ten items across the constructs of agency (three items), connection to experience (one item), problem solving (two items), social connectedness (three items), and positive emotional climate (one item). Based on data and feedback during the training, we further revised the tool by dropping or revising items.

The classroom 3–5 tool was further developed in the Jordan Adapt phase. The training was bilingual in Arabic and English and involved live interpretation and tool translation. NYU staff worked with data collectors to modify cultural references in items and exemplars from Colombia to Jordan. Training revealed the need to remove esoteric jargon. Items were re-worded in the exploration construct to clarify meaning (e.g. distinguishing whether an item refers to physical or sensory exploration, or mental exploration). Some words were vague in the translation and thus revised based on group input. The problem-solving construct was revised to clarify the goal of the item. For example, edits were made to help clarify the goals of engaging children in hypothesis generation (i.e. predicting questions) versus engaging children in generating reasons or explanations for results (i.e. explanation/reasoning questions). After Adapt phase data collection, we followed the process as we did in Colombia. Specifically, we calculated agreement rates and identified potentially problematic

items (i.e. >1 average absolute difference). Only two items had an average absolute difference of 1; all others had a value of <1. No further changes were made based on these results; it was determined that more focused training in the Test phase with attention to these items would be sufficient.

Learnings

- *The inclusion of subject-matter experts, or SMEs, was highly beneficial, not only from the perspective of having an engaged advisory group but also in line with the notion of having more local ownership and expertise on the tools.*
- *Simple language, free of academic jargon, is key. This includes defining key terms, such as:*
  - *Child exploration: physical process of discovery*
  - *Child problem solving: scientific method and process of investigation*
  - *Social connectedness: peer relationships and learning social norms*
  - *Positive emotional climate: teachers modelling positive behaviours and responsiveness, among other things*
  - *Predicting questions: used for the purpose of hypothesis generation*
  - *Explanation questions: used for engaging children in generating reasons or explanations for results*
- *How to identify agency: Child agency should involve authentic choices. Teachers should be the ones purposefully creating the environment for these choices.*
- *Exploration can be either physical/sensory or mental.*

Caregiver Tools

The caregiver 0–2 and 3–5 tools were developed and piloted in Colombia. They were developed in tandem with the classroom 3–5 tool, following the process outlined above with SMEs, followed by part two of the Adapt phase. The tools changed during the Adapt phase, as shown in **Table 22** and **Table 23.**

Table 22. Caregiver 0–2 tool

Construct	Old total	Number of Items		New total
		Dropped or moved	Suggested by SMEs	
Agency	9	-3	2	8
Exploration	7	-1	1	7
Positive emotional climate	8	-1	3	10
Social connectedness	3	0	4	7

Note: The far-right column does not represent the final number of items for the Test phase, as the tool continued to be refined after Adapt II, with final changes being made during Test training.

Table 23. Caregiver 3–5 tool

Construct	Old total	Number of Items		New total
		Dropped or moved	Suggested by SMEs	
Agency	9	-3	3	9
Exploration	10	-4	1	7
Connection to experience	6	0	1	7
Problem solving	6	0	2	8
Positive emotional climate	8	0	3	11
Social connectedness	3	0	3	6

Note: The far-right column does not represent the final number of items for the Test phase, as the tool continued to be refined after Adapt II, with final changes being made during Test training.

The tool continued to evolve during rubric development in the Adapt phase. Discussions centred on the translation of words from English to Spanish, particularly whether the translation accurately captured the intention of the item. Positive emotional climate was discussed, and, as a result, separate items were generated for physical and verbal affection. As with the classroom 3–5 tool, we discussed each item in the agency construct and compared the constructs of agency and exploration to ensure clarity for data collectors. We teased apart the different kinds of support that a caregiver might provide and decided that motivational support would fall under positive emotional climate, while cognitive support would relate to items in problem solving. Although we trained separately in Colombia on the classroom 3–5, caregiver 0–2, and caregiver 3–5 tools, some conversations transferred across tools. For example, we discussed adults as providing either passive or active support, and we spent time as a group clarifying terms relevant to items in problem solving, such as problem, solution, goal, and objective; both of these discussions were applicable across tools. In the future, there could be a combined training – for example, for the two caregiver observation tools, but with breakout groups for developmentally specific aspects. We calculated agreement rates and identified potentially problematic items (i.e. >1 average absolute difference on the four-point scale). For the caregiver 0–2 tool, this applied to seven items: two items in agency, three items in exploration, and two items in positive emotional climate. For the caregiver 3–5 tool, this applied to four items in the following constructs: agency (two items) and exploration (two items). Based on data and feedback during the training, we further revised the tool by dropping or revising items.

The Build phase was the only point in the study where we were able to include unstructured (i.e. naturalistic child point-of-view) observations that were intended to be used to inform the development of structured observation. Analysis results from the Build phase show that caregivers engage their children in playful activities, including household chores and tasks. Therefore, objects and materials related to daily routines and self-care were included in the toy kit for structured observations for the Test phase. The kit also served the purpose of providing a uniform set of materials for data collectors (instead of simply what was available), thus providing a 'level playing field' for all and ensuring some degree of comparability in the data. Additionally, these kits were tailored by age group in order to ensure developmentally appropriate toys and materials. All Test phase observations were conducted in family settings, which included community homes and centres, as well as schools that offered early childhood services.

Learnings

- *Motivational support relates to the positive emotional climate construct.*
- *Cognitive support relates to the problem-solving construct.*
- *Caregivers do engage children in playful activities, which include household chores and tasks. Therefore, objects and materials related to daily routines and self-care are an important component to consider when developing relevant play toolkits and procedures.*

7.3.2 Teacher Survey

The Teacher Survey was developed in Colombia through Build phase results and input from the Universidad de los Andes based on other play-related projects that the university has done in the ECE context. Build phase analysis results facilitated development of the Teacher Survey. Specifically, we used observation data results to generate vignettes of 'typical' scenarios that may occur in an ECE classroom, and items in the traditional format section are related to the classroom 3–5 tool.

Development of the Teacher Survey continued in Jordan. Surveys were administered in the Adapt phase; survey data analysis and cognitive interviews facilitated revisions to the Teacher Survey. The traditional format and vignette portions of the Teacher Survey changed in the following ways: (1) the Likert scale was adjusted; (2) probes were included throughout for clarity; and (3) section titles were removed so teachers were 'blind' to the topic of inquiry for each survey section. This was important for data collection in Jordan but not Colombia. Specifically, as determined by the norms of our partner organisation around data collection and how they typically collect data, data collectors in Jordan would hand teachers the tablet and teachers would self-administer the survey. Therefore, it was important that section headers not be leading. In Colombia, however, data collectors administered the surveys to participants.

Learnings

- *'Blinding' – the topic of inquiry can be a useful method to prevent biases in teacher responses.*
- *It is important to be mindful of differing data collection norms by context (e.g. survey self-administration versus enumerator administration).*

7.3.3 Caregiver Survey

The Caregiver Survey was developed and piloted solely in Colombia. Survey development was based on Build phase results and input from colleagues at the Universidad de los Andes who have extensive experience conducting research with families of children aged 0–5. Analysis results from the Build phase prompted us to think more carefully about the integration of household work and chores in the tools and the role of extended family members. Data show that children learn from adults by watching them do daily household tasks and chores, and sometimes join in and participate in the tasks with caregivers. Routines present an opportunity for children to play, as well as for adults to teach children responsibility, how to care for oneself, and how to perform daily tasks. Build phase data also show that engaging children in play and learning is not limited to the role of primary

caregiver. Extended family members are frequently mentioned in the data as being involved in aspects of caregiving. The activities that caregivers do with children are cited as both playful and an opportunity for teaching and learning. Thus, a section related to chores and work at home was included, and items refer to other adults as supporting caregiving.

Learnings

- *Understanding the role of extended family members in caregiving is critical.*
- *Household routines, work, and chores present not only an opportunity for children to learn from adults but also an opportunity to play.*
- *Items need to look sufficiently different from one another in practice in a classroom setting in order for data collectors to differentiate them.*
- *Agency should be a distinct construct rather than an organising principle to allow for reliable observation.*

7.3.4 Training Approaches

Data collector training for the Adapt phase focused first on perceptions of quality in ECE or caregiving contexts and then examined each construct in detail, at times viewing a short video clip of a classroom or dyadic interaction to help illustrate what each construct means. The training was centred on rubric development in the Adapt phase in order to have exemplars and scoring guidance that made sense for each country. Although remote, the Adapt phase trainings were highly interactive, with in-depth discussions as a whole group and in breakout rooms. As the rubric took shape, we utilised archival videos to practice scoring. We relied on the rich experience of data collectors, as well as the expertise of field-based consultants hired in each country. Test phase trainings were in person and focused solely on understanding the constructs and application of items. To facilitate this, we used archival videos wherever possible, as well as site visits to practice using the tool in person. Training sessions were discussion based; we often met as a whole group before breaking out into smaller groups, and then reconvened to share understandings, points of confusion, and so forth. The tools were revised slightly during the Test training and finalised for the Test phase pilot data collection. The trainings for both phases were very involved and intensive, and most importantly highly collaborative.

8. RECOMMENDATIONS FOR FINAL FORM OF TOOLS AND USE OF TOOLS

In this section, we summarise our recommendations for modifications to each instrument in the PLAY toolkit (refer to **Table 24** for a summary of the components). For each tool, we present an overview of its purpose and original format, a summary of the findings based on analysis, and the accompanying recommendations for modifying it, which are represented in the final form of each tool in the toolkit. The full findings on which these recommendations are based, which include item-level distributions and factor analyses for each tool in each country, can be found in Appendix B of this report. While analysis was done and findings presented for each individual tool by country, single tools for each age group and participant group are presented in the toolkit and expanded upon below. In addition to these tool-specific recommended modifications, we outline in the PLAY 1.0 toolkit overall

recommendations for continuing to review and refine each tool in preparation for the next phase of piloting in PLAY 2.0.

Table 24. Summary of final PLAY toolkit components

Age group	Participant group	PLAY tools
0–2 years	Caregiver-child	• Caregiver-Child Observation Tool • Caregiver Survey
3–5 years	Caregiver-child	• Caregiver-Child Observation Tool • Caregiver Survey
	Classrooms	• ECE Classroom Observation Tool • ECE Teacher Survey • ECE Classroom Inventory*
6–12 years	Classrooms	• Primary Classroom Observation Tool • Primary Teacher Survey • Primary Student Survey • Primary Classroom Inventory*

* This Classroom Inventory is taken within what we refer to as 'supplementary items' in the early childhood observation tools.

8.1 Early Childhood Tools

Factor analyses were performed on each tool by country. Our final recommendations for constructs and items to be included in each of the tools are based on these analyses, as well as ongoing discussions and feedback on how to improve the usability of the tools.

8.1.1 Caregiver-Child Observation Tool

- **Overview:** The Caregiver-Child Observation Tool (refer to **Table 25**) is an observation of caregiver-child interaction that is administered based on the child's age. There is a separate version of the tool for children aged 0–2 and 3–5. Both versions contain observational items across core constructs, as well as supplemental items that collect data on caregiver and child demographics, and structural and process qualities of the family/home setting. All items are rated on a scale of 0–3 (0 = no indication of quality, 1 = low quality, 2 = moderate quality, 3 = high quality).

Table 25. Comparison of original and final caregiver-child observation tools

	Original	Final
0–2	• 4 constructs Agency Exploration Positive emotional climate Social connectedness • 25 items	• 3 constructs Agency Exploration Social connectedness & positive emotional climate • 13 items
3–5	• 6 constructs Agency Connection to experience Exploration Problem solving Social connectedness Positive emotional climate • 37 items	• 3 constructs Agency Problem solving Connection to experience & social connectedness • 15 items

- **Findings:** Table 25 shows the characteristics of our final recommended form of each of the caregiver observation tools, as compared to the original versions.

- *0–2 Caregiver-Child Observation Tool:* Factor analyses initially resulted in a one-factor model. However, due to patterns of fit statistics and the fact that it would allow us to retain more of the original constructs, we ultimately decided on a three-factor structure, with a total of 13 items loading onto the model (six items in factor 1, four in factor 2, and three in factor 3).
- *3–5 Caregiver-Child Observation Tool:* A three-factor structure was found for the 3–-5 Caregiver-Child Observation, with a total of 15 items loading onto the model (five items in each factor).

- **Recommendations:** Despite some differences for both the 0–2 and 3–5 Caregiver-Child Observation Tools between which items loaded onto each factor (as compared to the original groupings), the factors produced from each tool all show strong conceptual underpinnings (i.e. there is conceptual connectivity within each group of items). Therefore, we recommend that the 0–2 Caregiver-Child Observation Tool include the model's 13 items falling into the following three constructs: child agency, exploration, and social connectedness & positive emotional climate (combined). We recommend that the 3–5 Caregiver-Child Observation Tool include 15 items falling into the following constructs: child agency, problem solving, and connection to experience & social connectedness (combined). In each tool, we combine two of the original constructs (social connectedness & positive emotional climate, and connection to experience & social connectedness) into a single one to represent the mix of items that was found in that factor.

8.1.2 Caregiver Survey

- **Overview:** The Caregiver Survey is typically administered to the caregiver by the data collector as a structured interview. The survey is made up of two sections: (1) demographic items and (2) a self-report on child activities and core construct items. The survey is the same for caregivers of children in the 0–2 and 3–5 age group.
- **Findings:** A three-factor structure was found for the Caregiver Survey, with a total of 15 items loading onto the model (five items in each factor).
- **Recommendations:** Based on our findings, we recommend that the Caregiver Survey include these 15 items across the following constructs: positive emotional climate, connection to experience, and social connectedness. In addition, we modified the presentation of each of the constructs so that they are labelled as 'sets' (e.g. the positive emotional climate factor is labelled as set 1) in the tool itself to prevent administration bias and to allow for country- and context-specific naming of constructs.

8.1.3 ECE Classroom Observation Tool

- **Overview:** The ECE Classroom Observation Tool (refer to **Table 26**) contains observational items across core constructs, as well as supplemental observational checklist items that collect data on structural and process qualities of the school and classroom. It is designed to be administered through an observation of a half day, preferably a full morning of observation, which in many systems represents a day's entire session of 2–3 hours).

Table 26. Comparison of original and final ECE classroom observation tools

Original	Final
• 6 constructs Agency Connection to experience	• 3 constructs Exploration & problem solving Connection to experience

Original	Final
Exploration Problem solving Social connectedness Positive emotional climate	Social connectedness
• 46 items	• 17 items
• Four-point scale: 0 = no indication of quality 1 = low quality 2 = moderate quality 3 = high quality	• Three-point scale: 0 = no indication of quality 1 = low to moderate quality 2 = high quality

- **Findings:** Table 26 shows the characteristics of our final recommended form of the ECE Classroom Observation Tool, as compared to the original version. A three-factor model with 17 items was found for Jordan, Colombia, and Ghana. Factors and items were consistent across all countries, with the exception of two items that loaded in Jordan and Ghana but not in Colombia. Factor 1 consisted of seven items, factor 2 had five items, and factor 3 had five items for Jordan and Ghana but three items for Colombia.
- **Recommendations:** Based on these findings, we recommend that the ECE Classroom Observation Tool include all 17 items across the following three constructs: exploration & problem solving, connection to experience, and social connectedness. The first construct, exploration & problem solving, combines these two original constructs into a single one to represent the mix of items that was found in that factor.

In addition to analysis results, we integrated the following into our final recommendations: feedback from training sessions on and use of the tool (primarily from a training experience in South Africa in July 2022 with another LEGO Foundation partner), and group discussions about opportunities for more closely aligning the primary and early childhood tools. For the final recommended items, we made minor wording modifications to simplify the language and make it easier to understand, while still maintaining the original meaning of each item. To better align with primary tools, the quality definitions in the scoring rubric were adjusted for some items so that all were scored along a dimension of either frequency or effectiveness (removing the categories of inclusivity and time). We also inserted additional introductory language to provide more detailed guidance on using and completing the observation for data collectors' reference in the tool itself. Last, we added to the materials for this tool standardised notetaking and scoring forms to provide stronger guidance for completing the observation.

As a result of cross-organisation discussions, the South Africa training, and analysis, we recommend that the scale for the 3–5 Classroom Observation Tool be reduced to a three-point scale: 0 = no indication of quality, 1 = low to moderate quality, and 2 = high quality. This decision was made based on analysis using this collapsed scale, which confirmed that the models still performed sufficiently, as well as the advantage of a simpler tool more closely aligned to that of the Primary Classroom Observation Tool. In addition, in order for the ECE and Primary Classroom Observation Tools to have consistent constructs, we recommend adding agency items back into the tool moving forward in PLAY 2.0. This can be done by drawing from the primary classroom observation agency items for use in the ECE tool or by adding or refining items in collaboration with PLAY 2.0 implementing partners.

ECE Teacher Survey

- **Overview:** The ECE Teacher Survey is a structured interview conducted by the data collector or self-administered by the teacher (depending on what is most appropriate for the context). The survey is made up of three sections: (1) demographic and educational and training background items; (2) traditional-format survey items of PLAY practices; and (3) vignette-based survey items.
- **Findings:** Country-specific factor analyses yielded distinct factor structures for Jordanian and Colombian data, resulting in a total of 32 items across both models. Due to the difference in models from each country's dataset, we presume that the functioning of this tool and its contents may vary more by context than the other early childhood tools. As discussed below, the varying models led us to be conservative in retaining items, to preserve all possible variations of the tool across contexts.
- **Recommendations:** Because of the unique factor structures found for Jordan and Colombia, we recommend that the survey measures of PLAY practices maintain a set of items inclusive of all that loaded for either country (32 total) in order to fully retain the breadth of variation that was found in the Teacher Survey. Retaining all items from both models will allow us to continue to test them in additional contexts and develop a more complete understanding. For the vignette-based items, we found consistency in factor structures across countries and also recommend retaining all 18 original items. Also, consistent with the caregiver survey, we modified the presentation of each of the constructs so that they are labelled as 'sets' in the tool itself to prevent administration bias and to allow for country-specific naming of constructs.

8.2 Primary Tools

Factor analyses were performed on the three primary tools. Our final recommendations for constructs and items to be included in each are based on these analyses, as well as ongoing discussions and feedback on how to improve the usability of the tools.

8.2.1 Primary Classroom Observation Tool

- **Overview:** The Primary Classroom Observation Tool contains items that can be observed in a classroom for 6- to 12-year-olds. It is intended to be used for the duration of a subject (e.g. 45 minutes), such as reading or mathematics. It is conceivable that users may want to observe multiple subjects in the same day. There is also a supplemental checklist to collect information on structural elements such as resources available in the classroom or student work displayed. A comparison of original and final primary school observation tools is shown in **Table 27.**

Table 27. Comparison of original and final primary school observation tools

Original	Final
• 7 constructs Connection to experience Problem solving Exploration Agency Participation Social connectedness Positive emotional climate	• 3 constructs Exploration Agency Social connectedness
• 44 items	• 19 items
• Quality scale metrics Frequency, 32 items Effectiveness, 8 items Participation, 4 items	• Quality scale metrics Frequency, 11 items Effectiveness, 4 items Participation, 4 items
• Quality scale points 0 = no indication of quality 1 = low quality 2 = high quality	• Quality scale points 0 = no indication of quality 1 = low quality 2 = high quality

- **Findings:** The analysis of the Test phase data from Kenya, Ghana, and Colombia yielded four potential models, each with three factors. In these four models, the number of items varied between 12 and 19 items.
- **Recommendations:** Based on these findings, we recommend that the Primary Classroom Observation Tool use the model that includes 19 items across the constructs of exploration, agency, and social connectedness. For the 2.0 version, in order for the Primary and ECE Classroom Observation Tools to have consistent constructs, we recommend adding connection to experience to the primary tool. This can be done by adapting the ECE tool connection to experience items for use in the primary tool or by adding or refining items in collaboration with PLAY 2.0 implementing partners.

8.2.2 Primary Teacher Survey

- **Overview:** The Primary Teacher Survey contains three sections. One section gathers demographics in a self-report oral survey. The second section is directed by the teacher, who sorts short descriptions of classroom practices under headers to indicate the frequency of the practices. The third part consists of scenario-based survey items measuring three elements: frequency, levels of confidence, and effectiveness.
- **Findings:** The analysis of the Test phase data from Kenya, Ghana, and Colombia yielded three potential models, with four or five factors varying between 23 and 27 items. For the scenario-based survey items, the analysis examining frequency, levels of confidence, and effectiveness did not yield anything concerning.
- **Recommendations:** Based on these findings, for the teacher sort, we recommend the four-factor model with 25 items with four constructs: social connectedness, participation, connection to experience, and agency. For the scenario section of the Teacher Survey, we recommend maintaining it as administered.

8.2.3 Primary Student Survey

- **Overview:** The Primary Student Survey is administered individually using a Likert scale with four options. The student responds by indicating agreement or disagreement and is then prompted with a follow-up to indicate the degree of that initial reaction.
- **Findings:** The analysis of the Test phase data from Kenya, Ghana, and Colombia yielded three potential models, with one or two factors varying between 11 and 18 items.
- **Recommendations:** Based on these findings, for the Primary Student Survey, we recommend a two-factor model with 16 items for two constructs: positive emotional climate and problem solving.

8.3 Intended Uses and Users of the PLAY Tools

Each tool in the PLAY Measurement toolkit consists of several constructs. The aim of the tools is to provide valid, reliable quantitative estimates for each construct. The tool can be used wherever a reliable quantitative estimate of these constructs is required. Examples include

- national or regional monitoring efforts that focus on how education and other service systems are supporting child and student engagement and learning in homes, centres, or schools
- impact evaluations of programmes or interventions in homes, centres, or schools that intend to support children's engagement and learning
- assessing the implementation of a programme or intervention to support children's engagement and learning

The toolkit is not designed as a formative assessment or to give feedback to teachers and parents. However, the observation tool could serve this purpose with minimal adaptation – it could be used to record the interactions taking place in a teacher's classroom or in a parent's home, and then the observer and adult being observed could discuss together the interactions that were observed and those that were not observed. This could guide the teacher or parent to identify behaviours they might wish to adopt.

8.4 Which Tools to Use?

For each of the four participant groups, there is an observation tool and an adult survey. These tools are designed to complement each other, and we recommend using them together (the 'comprehensive approach' in **Table 28**). If only one of these tools (the 'parsimonious approach' in Table 29) is to be used, we recommend using the observation tool for two reasons: (1) observational tools have less bias than may be involved in self-reported behaviours like teacher-reported practices in their own classroom; and (2) items on observational tools may be more productively used in professional development and other workforce supports in family support and education systems (consider feedback by a coach based on a teacher's interview response versus feedback based on observation of the classroom). However, the toolkit would need to undergo significant further development before we'd recommend its use in formative assessment of teachers.

Table 28. Approach and tools

Approach	Tools
Comprehensive	Observation Adult survey (self-reported behaviour and vignette-based reasoning) Student survey (for primary age group only)
Parsimonious	Observation only
Focus on learner perspective	Include student survey (for primary age group) in set of tools
Focus on evaluation or monitoring of an intervention	Include observation tool in set of tools

The adult survey tool consists of two parts: a survey of self-reported behaviours and a judgment task based on vignettes. The self-reported behaviour survey measures the same constructs as the observation tool but is quicker to administer and involves less training. There are moderate correlations between self-reported behaviour and the observation instruments, raising the possibility that the adult self-reported behaviour survey may be used alone. However, data from PLAY 2.0 are required to make this a strong recommendation. Moreover, there are some situations in which the use of the adult survey alone is not recommended. When the toolkit is being used to monitor or evaluate an intervention targeted at adults (e.g. a teacher training), the intervention may bias the adults' subject ratings. Thus, improvements in teacher self-rating in an intervention group, compared to a control group, may be the result of bias or changes in perception and not necessarily the result of changes in behaviour. For such purposes, we strongly recommend using the observation tools ('focus on evaluation or monitoring of an intervention' in Table 28).

Currently, we have less strong evidence for the validity of the vignette-based reasoning task and recommend collecting more evidence on this component of the toolkit.

For the primary school tools, there is the additional option of including a Student Survey. The Student Survey does not measure the same constructs as the observation tool and the Teacher Survey. Correlations are weak between student responses and data from the Primary Classroom Observation Tool and Teacher Survey. We recommend using the Student Survey when it is important to assess students' perceptions of classroom support for engagement ('Focus on learner perspective' in Table 28). The Student Survey cannot be used as a proxy for the observation or adult survey.

For the 3–5 age group, there are two sets of tools – one for use in the classroom and one for use in the home. These sets of tools can be used independently or with the same sample of children observed at home and in the classroom.

8.5 How Do the PLAY Tools Relate to Other Measures of Education Quality?

The PLAY Measurement toolkit adds to a number of other measures of education quality currently in use. This is how PLAY relates to other measures of quality:

- In common with other quality measures (e.g. CLASS, TEACH, TIPPS), PLAY has a **domain-general** focus. That is, it is not subject specific, as are tools for the language and literacy environment (CHELLO) and read-alouds (SABR) (Neuman, Koh, & Dwyer, 2008; Pentimonti et al., 2021).
- In common with other quality measures, PLAY measures **adult-child interactions**. However, unlike existing quality measures, PLAY focuses only on specific adult-child interactions, namely those that support (effortless) **self-sustaining engagement** in learning. In the conceptual framework guiding this toolkit, the specific adult-child

interactions hypothesised to promote child engagement are those supporting connection to experience, problem solving, exploration, agency, social connectedness, and a positive emotional climate.

- The PLAY toolkit includes measurement scales for several constructs that are lacking or are measured with only one or two items in other tools. The constructs of **exploration** and **connection to experience** provide domain-general measures of child-centred approaches to conceptual understanding of lesson content. The tool also provides a more detailed measure of child **agency** and extends existing measures of social interaction in the class to develop a measure of **social connectedness** particularly relevant to many cultures in low- and middle-income countries.
- The **relation to learning outcomes has not yet been established** for all constructs measured by PLAY. Thus, PLAY measures some constructs that are hypothesised to relate to learning, whereas existing measures assess constructs with a more established relation to learning. Some constructs may be dropped from future iterations of PLAY if there is no established relation to learning across country contexts.
- PLAY aims to measure interactions that promote a **broad range of outcomes**, including social and emotional outcomes as well as academic outcomes.
- The notion of 'quality' relates only to interactions taking place in classrooms and centres. In homes, we refer to 'caregiver **support** for self-sustaining engagement' rather than 'quality.'

8.6 Next Steps

In this section, we offer recommendations for the final form of the tools based on our findings. In addition, a few steps are required to act on these recommendations to produce a toolkit that can be used at the start of the next phase of this work – PLAY 2.0.

Review construct definitions. Our analyses led to a rearrangement of the items associated with hypothesised constructs. The meaning of each revised construct can be inferred from the new set of items grouped within the construct. This could result in the implied definition of a revised construct differing from what was originally set out in our conceptual framework. In other cases, a revised construct may result from a combination of two original constructs.

In the PLAY 1.0 toolkit, we have conservatively retained the construct definitions proposed in the conceptual framework. For PLAY 2.0, we recommend a collaborative review of constructs to determine where revisions may be needed to their definitions.

Revise or add items. After revising construct definitions, we will review the alignment of items with the revised construct definitions produced in the previous step. We will identify items whose wording can be adjusted for better alignment with the revised definition and potentially suggest new items that could improve the measurement of revised constructs.

Simplify and harmonise quality metrics. The observation tools in PLAY 1.0 use different types of quality metrics based on the frequency and effectiveness of interactions and the level of participation of children in the interactions. In PLAY 1.0, we found that it was easiest to train data collectors on frequency metrics. In PLAY 2.0, we will look to use frequency metrics more widely while retaining effectiveness and participation metrics for items that showed strong psychometric properties or where it was most appropriate based on class structures.

REFERENCES

Adair, J. K. (2014). Agency and expanding capabilities in early grade classrooms: What it could mean for young children. *Harvard Educational Review*, *84*(2), 217-241.

Adair J. K. & Colegrove, K.S. (2014). Communal Agency and Social Development: Examples from First Grade Classrooms Serving Children of Immigrants. *Asia-Pacific Journal of Research in Early Childhood Education* 8(2), 69-91.

Adair, J. K., & Colegrove, K. S. S. (2021). *Segregation by experience: Agency, racism, and learning in the early grades*. University of Chicago Press.

Akyeampong, K. (2014). Reconceptualised life skills in secondary education in the African context: Lessons learnt from reforms in Ghana. *International Review of Education, 60,* 217-234. https://doi.org/10.1007/s11159-014-9408-2

Alcalá, L., Rogoff, B., & Fraire, A. L. (2018). Sophisticated collaboration is common among Mexican-heritage US children. *Proceedings of the National Academy of Sciences*, *115*(45), 11377-11384.

An, D., Eggum-Wilkens, N. D., Chae, S., Hayford, S. R., Yabiku, S. T., Glick, J. E., & Zhang, L. (2018). Adults' conceptualisations of children's social competence in Nepal and Malawi. *Psychology and Developing Societies, 30*(1), 81-104. https://pubmed.ncbi.nlm.nih.gov/30078957/

Apolot, J. M., Ejuu, G., & Lubaale, G. (2020). Pursuing quality education in Karamoja: An analysis of the caregivers' quality indicators from a community perspective for sustainable early childhood education programmes. *American Journal of Education and Practice*, *4*(1), 72-88.

Arnett, J. J. (2016). The neglected 95%: Why American psychology needs to become less American. In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (4th ed.) (pp. 115-132). Washington, DC: American Psychological Association.

Baker-Henningham, H., Walker, S., Powell, C., & Gardner, J. M. (2009). A pilot study of the Incredible Years Teacher Training programme and a curriculum unit on social and emotional skills in community pre-schools in Jamaica. *Child: Care, Health and Development*, *35*(5), 624-631.

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., ... & Walton, M. (2017). From proof of concept to scalable policies: Challenges and solutions, with an application. *Journal of Economic Perspectives*, *31*(4), 73-102. https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.31.4.73

Bang, M. (2020). Cultivating just, culturally thriving, and sustainable worlds: Towards expanding the science of learning, development, and education beyond Western constructions of humans and nature. Lecture, November 2020, New York University.

Baumrind, D. (2013). Authoritative parenting revisited: History and current status. In R. E. Larzelere, A. S. Morris, & A. W. Harrist (Eds.), *Authoritative parenting: Synthesizing nurturance and discipline for optimal child development* (pp. 11–34). American Psychological Association. https://doi.org/10.1037/13948-002

Bazyk, S., Stalnaker, D., Llerena, M., Ekelman, B., & Bazyk, J. (2003). Play in Mayan children. *American Journal of Occupational Therapy*, *57*(3), 273-283.

Bolisetty, S., Castelyn, J., Chowdhury, S., Kim, S., Nyeu, M., Trang, K., & Yoshikawa K. (2022). *Looking at play through the eyes of Rohingya children in Cox's Bazar*. Inter-agency Network for Education in Emergencies.

Bornstein, M. H., Putnick, D. L., & Lansford, J. E. (2011). Parenting attributions and attitudes in cross-cultural perspective. *Parenting, 11*(2-3), 214-237. doi:10.1080/15295192.2011.585568

Bronfenbrenner, U. & Morris, P.A. (2006). The bioecological model of human development. In W. Damon (Series Ed.) & R. M. Lerner (Vol. Ed.), Handbook of child psychology: Theoretical models of human development (pp. 793-828). New York, NY: Wiley.

Brunette, T., Piper, B., Jordan, R., King, S., & Nabacwa, R. (2019). The impact of mother tongue reading instruction in twelve Ugandan languages and the role of language complexity, socioeconomic factors, and program implementation. *Comparative Education Review*, *63*(4), 591-612.

Burchinal, M. (2018). Measuring early care and education quality. *Child Development Perspectives*, *12*(1), 3-9.

Calero, C., Diez, V. G., Soares, Y. S., Kluve, J., & Corseuil, C. H. (2017). Can arts-based interventions enhance labor market outcomes among youth? Evidence from a randomized trial in Rio de Janeiro. *Labour Economics*, *45*, 131-142.

Caspi, A., Elder, G. H., Jr, & Bem, D. J. (1988). Moving away from the world: Life-course patterns of shy children. *Developmental Psychology, 24,* 824–831

Chen, X., Cen, G., Li, D., & He, Y. (2005). Social functioning and adjustment in Chinese children: The imprint of historical time. *Child development*, *76*(1), 182-195.

Chen, X., Chen, X., Zhao, S., Way, N., Yoshikawa, H., Zhang, G., ... & Li, D. (2021). Autonomy-and connectedness-oriented behaviors of toddlers and mothers at different historical times in urban China. *Developmental Psychology*, *57*(8), 1254-1260. https://doi.org/10.1037/dev0001224

Chen, X., Rubin, K. H., & Sun, Y. (1992). Social Reputation and Peer Relationships in Chinese and Canadian Children: A Cross-cultural Study. *Child Development*, *63*(6), 1336–1343. http://doi.org/10.1111/j.1467-8624.1992.tb01698.x

Christie, P. (1985). *The right to learn: The struggle for education in South Africa.* Ravan Press of South Africa.

Christie, P. (2008). Opening the doors of learning: Changing schools in South Africa. Johannesburg: Heinemann.

Coyne, I., & Carter, B. (Eds.). (2018). Being participatory: Researching with children and young people: Co-constructing knowledge using creative techniques. Springer.

Csikszentmihalyi, M. (1997). Flow and education. *NAMTA journal*, *22*(2), 2-35.

Dahlberg, G., Moss, P., & Pence, A. R. (1999). Beyond quality in early childhood education and care: Postmodern perspectives. Psychology Press.

Davis, D., & Raikes, A. (2020). Examining the congruency between the MELE observation tool and LEGO learning through play. University of Nebraska.

Deater-Deckard, K., Lansford, J. E., Malone, P. S., Alampay, L. P., Sorbring, E., Bacchini, D., … & Al-Hassan, S. M. (2011). The association between parental warmth and control in thirteen cultural groups. *Journal of Family Psychology, 25*(5), 790-794. https://doi.org/10.1037/a0025120

De León, L. (2011). Language socialization and multiparty participation frameworks. *The handbook of language socialization*, *72*, 81.

Deci, E. L., Ryan, R. M., Vallerand, R. J., & Pelletier, L. G. (1991). Motivation and Education: The Self-Determination Perspective. *Educational Psychologist*. https://doi.org/10.1080/00461520.1991.9653137

Del Giudice, M. (2014). An evolutionary life history framework for psychopathology. *Psychological Inquiry*, *25*(3-4), 261-300.

Dinarte Diaz, L. I., & Egana-del Sol, P. (2019). *Preventing violence in the most violent contexts: Behavioral and neurophysiological evidence.* The World Bank.

Eccles, J. S. (2005). The present and future of research on activity settings as developmental contexts. In J. L. Mahoney, R. W. Larson, & J. S. Eccles (Eds.), *Organized activities as contexts of development: Extracurricular activities, after-school and community programs* (pp. 353-371). Erlbaum.

Ejuu, G., Apolot, J. M., & Serpell, R. (2019). Early childhood education quality indicators: Exploring the landscape of an African community perspective. *Global Studies of Childhood*, 2043610619832898.

Fong, V. L. (2007). Parent-child communication problems and the perceived inadequacies of Chinese only children. *Ethos*, *35*(1), 85-127. https://doi.org/10.1525/eth.2007.35.1.85

Gallimore, R., Goldenberg, C. N., & Weisner, T. S. (1993). The social construction and subjective reality of activity settings: Implications for community psychology. *American journal of community psychology*, *21*(4), 537-560.

Garcia Coll, C., & Szalacha, L. (2004). The multiple contexts of middle childhood. *The Future of Children,* 81-97.

García-Sánchez, I. M. (2016). Friendship, participation, and multimodality in Moroccan immigrant girls' peer groups. In M. Theobald (Ed.), *Friendship and peer culture in multilingual settings* (pp. 1-31). Bingley, UK: Emerald Group Publishing Limited.

Gaskins, S. (1999). Children's daily lives in a Mayan village: A case study of culturally constructed roles and activities. In A. Goncue (Ed.), *Children's engagement in the world: Sociocultural perspectives* (pp. 25–60). New York: Cambridge University Press.

Gaskins, S. (2013). Pretend play as culturally constructed activity. *The Oxford handbook of the development of imagination*, 224-247.

Gaskins, S. (2015). Childhood practices across cultures: Play and household work. *The Oxford handbook of culture and development*, 185-97.

Gaskins, S., Haight, W., & Lancy, D. F. (2007). The cultural construction of play. *Play and development: Evolutionary, sociocultural, and functional perspectives*, 179-202.

Gazelle, H., & Ladd, G. W. (2003). Anxious solitude and peer exclusion: a diathesis-stress model of internalizing trajectories in childhood. *Child Development*, *74*(1), 257–278. https://doi.org/10.1111/1467-8624.00534

Gjicali, K., Finn, B. M., & Hebert, D. (2020). Effects of belief generation on social exploration, culturally-appropriate actions, and cross-cultural concept learning in a game-based social simulation. *Computers and Education*, *156*, 103959.

Goodfriend, E., Wuermli, A., Hiott, C., Trang, K., Iqbal, Y., & Castelyn, J. (2022). Delivering quality research in culturally dynamic, conflict-affected contexts: Lessons from large-scale pilot research in Cox's Bazar. NYU Global TIES for Children Research Brief.

Gosso, Y., & Carvalho, A. M. A. (2013). Play and cultural context. In R. E. Tremblay, M. Bovin, R. de V. Peters, & P. S. Smith (Eds.), *Encyclopedia on early childhood development*.

Gottlieb, A. (2014). Is it time to detach from attachment theory? Perspectives from the West African rain forest. *Different faces of attachment: Cultural variations on a universal human need*, 187-214.

Gottlieb, G. (1991). Experiential canalization of behavioral development: Theory. *Developmental Psychology*, *27*, 4-13.

Greenfield, P. M. (2009). Linking social change and developmental change: Shifting pathways of human development. *Developmental Psychology*, *45*, 401–418. https://doi.org/10.1037/a0014726

Greenfield, P. M. (2016). Social change, cultural evolution, and human development. *Current Opinion in Psychology, 8*, 84-92. https://doi.org/10.1016/j.copsyc.2015.10.012

Grigorenko, E. L., Geissler, P. W. et al., (2001). The organisation of Luo conceptions of intelligence: A study of implicit theories in a Kenyan village. *International Journal of Behavioral Development, 25*(4), 367-378. https://doi.org/10.1080/01650250042000348

Hamre, B., Hatfield, B., Pianta, R., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher–child interactions: Associations with preschool children's development. *Child Development*, *85*(3), 1257-1274.

Haight, W. L., Wang, X. L., Fung, H. H. T., Williams, K., & Mintz, J. (1999). Universal, developmental, and variable aspects of young children's play: a cross-cultural comparison of pretending at home. *Child development*, *70*(6), 1477-1488.

Harwood, R. L., Schoelmerich, A., Ventura-Cook, E., Schulze, P. A., & Wilson, S. P. (1996). Culture and class influences on Anglo and Puerto Rican mothers' beliefs regarding long-term socialization goals and child behavior. *Child Development*, *67*, 2446-2461.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), 29-29.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2-3), 61-83. https://doi.org/10.1017/S0140525X0999152X

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in “educational” apps: Lessons from the science of learning. *Psychological Science in the Public Interest*, *16*(1), 3-34.

Hirsh-Pasek, K., Hadami, H., Blinkoff, E., & Golinkoff, R.M. (2020). *A new approach to education reform: Playful learning promotes 21st century skills in schools and beyond*. Brookings Institution. https://www.brookings.edu/policy2020/bigideas/a-new-path-to-education-reform-playful-learning-promotes-21st-century-skills-in-schools-and-beyond/

Jensen, H., Pyle, A., Zosh, J. M., Ebrahim, H. B., Scherman, A. Z., Reunamo, J., & Hamre, B. K. (2019). *Play facilitation: the science behind the art of engaging young children: White paper*. The LEGO Foundation.

Jukes, M., Gabrieli, P., Mgonda, N. L., Nsolezi, F., Jeremiah, G., Tibenda, J., & Bub, K. L. (2018). “Respect is an investment”: Community perceptions of social and emotional competencies in early childhood from Mtwara, Tanzania. *Global Education Review*, *5*(2), 160-188.

Jukes, M. C. H., Zuilkowski, S. S., & Grigorenko, E. L. (2018). Do schooling and urban residence develop cognitive skills at the expense of social responsibility? A study of adolescents in the Gambia, West Africa. *Journal of Cross-Cultural Psychology, 49*(1), 82-98. https://doi.org/10.1177/0022022117741989

Kärtner, J., Keller, H., & Chaudhary, N. (2010). Cognitive and social influences on early prosocial behavior in two sociocultural contexts. *Developmental Psychology, 46*(4), 905-914. https://doi.org/10.1037/a0019718

Kawulich, B. B. (2005, May). Participant observation as a data collection method. In *Forum qualitative sozialforschung/Forum: Qualitative social research* (Vol. 6, No. 2).

Keller, H. (2016). Psychological autonomy and hierarchical relatedness as organizers of developmental pathways. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *371*(1686). https://doi.org/10.1098/rstb.2015.0070

Keller, H., Bard, K., Morelli, G., Chaudhary, N., Vicedo, M., Rosabal-Coto, M., ... & Gottlieb, A. (2018). The myth of universal sensitive responsiveness: Comment on Mesman et al. (2017). *Child Development*, *89*(5), 1921-1928.

Keller, H., Lamm, B., Abels, M., Yovsi, R., Borke, J., Jensen, H., ... & Chaudhary, N. (2006). Cultural models, socialization goals, and parenting ethnotheories: A multicultural analysis. *Journal of Cross-Cultural Psychology*, *37*(2), 155-172.

Keller, H., Lohaus, A., Kuensemueller, P., Abels, M., Yovsi, R., Voelker, S., ... & Mohite, P. (2004). The bio-culture of parenting: Evidence from five cultural communities. *Parenting: Science and Practice*, *4*, 25-50.

Kim, S. (2020) Child Point-of-View Interaction Observation Training. [PowerPoint presentation]. Global TIES for Children. Play to Learn training, Bangladesh.

Lancy, D. F. (2007). Accounting for variability in mother–child play. *American Anthropologist*, *109*(2), 273-284.

Lancy, D. F. (2014). *The anthropology of childhood: Cherubs, chattel, changelings* (2nd ed.). Cambridge, UK: Cambridge University Press.

Mack, N. (2005). Qualitative research methods: A data collector's field guide.

Maldonado, C., Yoshikawa, H., Escallon, E., Ponguta, A., Kagan, S.L., Nieto, A.M., Rey-Guerra, C., Cristancho Amaya, J.C., Mateus, A., Rodriguez, A.M., Motta, A.C., Aragon, C.A., & Caro, A. (2021). *Measuring the Quality of Early Childhood Education: Associations with Children's Development from a National Study with the IMCEIC tool in Colombia.* Manuscript under review.

Mehta, J., & Fine, S. (2019). In search of deeper learning: The quest to remake the American high school. Harvard University Press.

Mesman, J., Minter, T., Angnged, A., Cissé, I. A. H., Salali, G. D., & Migliano, A. B. (2017). Universality Without Uniformity: A Culturally Inclusive Approach to Sensitive Responsiveness in Infant Caregiving. *Child Development*, *89*(3), 837–850. http://doi.org/10.1111/cdev.12795

Ministerio de Educación de Perú (2016). *Currículo nacional de la educación básica.* http://www.minedu.gob.pe/curriculo/pdf/curriculo-nacional-de-la-educacion-basica.pdf

Ministerio de Educación Nacional (Colombia). (2017). *Bases curriculares de educación inicial*. https://www.mineducacion.gov.co/1759/articles-341880_recurso_1.pdf

Ministerio de Educación Nacional (Colombia). (2017). *Bases curriculares para la educación inicial y preescolar*.

Morelli, G. A., Rogoff, B., & Angelillo, C. (2003). Cultural variation in young children's access to work or involvement in specialised child-focused activities. *International Journal of Behavioral Development*, *27*(3), 264-274.

Nasser, I., Cheema, J., Saroughi, M., & Alwani, A. (2019). *Advancing education in Muslim societies: Mapping the terrain*.

Nieto, A. M., Leyva, D., & Yoshikawa, H. (2019). Guatemalan Mayan book-sharing styles and their relation to parents' schooling and children's narrative contributions. *Early Childhood Research Quarterly, 47*, 405-417. https://doi.org/10.1016/j.ecresq.2018.08.006

Ng, F. F. Y., Tamis-LeMonda, C. S., Godfrey, E. B., Hunter, C. J., & Yoshikawa, H. (2012). Dynamics of mothers' goals for children in ethnically diverse populations across the first three years of life. *Social Development*, *21*(4), 821-848.

Noyau, C., & Gbeto, K. S. (2004). Les conceptions de l'intelligence dans la culture éwé (Togo): analyse sémantique des expressions du domaine cognitif. In *Congrès de l'ARIC (Association pour la Recherche InterCulturelle)* (pp. actes-en). Université d'Amiens.

Nsamenang, B.A. (2006). Human ontogenesis: An indigenous African view on development and intelligence. *International Journal of Psychology*, *41*(4), 293-297.

Nsamenang, B. (2010). The importance of mixed age groups in Cameroon. In M. Kernan & E. Singer (Eds.), *Peer relationships in early childhood education and care* (pp. 75-87). NY: Routledge.

Nyeu, M. (2020). From indigenous elders' stories to a critical thinking curriculum. Manuscript under review.

O'Donnell, C. R., Tharp, R. G., & Wilson, K. (1993). Activity settings as the unit of analysis: A theoretical basis for community intervention and development. *American Journal of Community Psychology*, *21*, 501-520.

OECD. (2018). Preparing our youth for an inclusive and sustainable world: The OECD PISA global competence framework. OECD.

Office of Head Start (2020). *Program performance standards*. Author. https://eclkc.ohs.acf.hhs.gov/policy/45-cfr-chap-xiii/part-1302-program-operations

Oppong, S. (2020). Towards a Model of Valued Human Cognitive Abilities: An African Perspective. *Frontiers in Psychology*, *11*, 3381.

Piper, B., Zuilkowski, S. S., & Ong'ele, S. (2016). Implementing mother tongue instruction in the real world: Results from a medium-scale randomized controlled trial in Kenya. *Comparative Education Review*, *60*(4), 776-807.

Plata, T., Hein, S., & Ponguta, A.L. (2020). Assessment of the and their alignment with LEGO's learning through play frameworks / constructs. Universidad de los Andes.

Pritchett, L. (2013). *The rebirth of education: Schooling ain't learning*. CGD Books. https://pz.harvard.edu/resources/towards-a-pedagogy-of-play?_ga=2.159768382.229084216.1572796162–1348374209.1541361198

Rad, M. S., Martingano, A. J., & Ginges, J. (2018). Toward a psychology of Homo Sapiens: Making psychological science more representative of the human population. *Proceedings of the National Academy of Sciences of the United States of America, 115*(45), 11401-11405. https://doi.org/10.1073/pnas.1721165115

Raikes, A., Sayre, R., Davis, D., Anderson, K., Hyson, M., Seminario, E., & Burton, A. (2019). The Measuring Early Learning Quality & Outcomes initiative: purpose, process and results. *Early Years*, *39*(4), 360-375.

Ramani, G. B. (2012). Influence of a playful, child-directed context on preschool children's peer cooperation. *Merrill-Palmer Quarterly (1982=).* 159-190.

Reese, E., Haden, C. A., & Fivush, R. (1993). Mother-child conversations about the past: Relationships of style and me Reese, E., Haden, C. A., & Fivush, R. (1993). Mother-child conversations about the past: Relationships of style and memory over time. *Cognitive development*, *8*(4), 403-430.

Richerson, P. J., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago, IL: University of Chicago Press.

Rogoff, B. (2003). *The cultural nature of human development.* Oxford University Press.

Rogoff, B. (2014). Learning by observing and pitching in to family and community endeavors: An orientation. *Human Development*, *57*(2-3), 69-81.

Rogoff, B., Paradise, R., Arauz, R. M., Correa-Chávez, M., & Angelillo, C. (2003). Firsthand learning through intent participation. *Annual Review of Psychology, 54*, 175-203. https://doi.org/10.1146/annurev.psych.54.101601.145118

Roopnarine, J. L. (2012). What is the state of play? *International Journal of Play*, *1*(3), 228–230. http://doi.org/10.1080/21594937.2012.735452

Roopnarine, J.L., & Davidson, K.L. (2015). Parent-Child Play Across Cultures: Advancing Play Research. American Journal of Play 7.2.

Sarason, S. B. (1996). *Revisiting "The culture of the school and the problem of change."* Teachers College Press.

Schlesinger, M. A., & Hirsh-Pasek, K. (2019). Playful Learning Landscapes: Creating skill-building experiences in community spaces. *Childhood Education*, *95*(4), 3-9.

Seidman, E. (2012). An emerging action science of social settings. *American Journal of Community Psychology*, *50*(1-2), 1-16.

Seidman, E., Kim, S., Raza, M., Ishihara, M., & Halpin, P. F. (2018). Assessment of pedagogical practices and processes in low and middle income countries: Findings from secondary school classrooms in Uganda. *Teaching and Teacher Education*, *71*, 283-296.

Serpell, R. (1993). The Significance of Schooling: Life-Journeys in an African Society. Cambridge: Cambridge University Press.

Serpell, R. (2011). Social responsibility as a dimension of intelligence, and as an educational goal: Insights from programmatic research in an African society. *Child Development Perspectives*, *5*(2), 126-133.

Serpell, R. (2021). Culture-sensitive communication in applied developmental research. *Human Development*, 1-16.

Smith, L. B., Yu, C., Yoshida, H., & Fausey, C. M. (2015). Contributions of head-mounted cameras to studying the visual environments of infants and young children. Journal of Cognition and Development, 16(3), 407-419.

Sonderen, E., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of reverse wording of questionnaire items: Let's learn from cows in the rain. *PLOS ONE, 8*(7), e68967. https://doi.org/10.1371/journal.pone.0068967

Sorkhabi, N. (2005). Applicability of Baumrind's parent typology to collective cultures: Analysis of cultural explanations of parent socialization effects. *International Journal of Behavioral Development*, *29*(6), 552-563.

Super, C. M. (1983). Cultural variation in the meaning and uses of children's intelligence. In J. B. Deregowski, S. Dziurawiec, & R. C. Annis (Eds.), *Expectations in cross-cultural psychology*. Lisse, Netherlands: Swets and Zeitlinger.

Super, C. M., & Harkness, S. (1986). The developmental niche: A conceptualization at the interface of child and culture. *International Journal of Behavioral Development*, *9*(4), 545-569. https://doi.org/10.1177/016502548600900409

Tamis-LeMonda, C. S., Way, N., Hughes, D., Yoshikawa, H., Kalman, R. K., & Niwa, E. Y. (2008). Parents' goals for children: The dynamic coexistence of individualism and collectivism in cultures and individuals. *Social Development*, *17*(1), 183-209.

Tobin, J., Hsueh, Y., & Karasawa, M. (2009). *Preschool in three cultures revisited: China, Japan, and the United States*. University of Chicago Press.

Tronick, E. Z., Morelli, G. A., & Ivey, P. K. (1992). The Efe forager infant and toddler's pattern of social relationships: Multiple and simultaneous. *Developmental Psychology*, *28*(4), 568-577. https://doi.org/10.1037/0012-1649.28.4.568

United Nations (2015). *Sustainable Development Goals*. United Nations Educational, Scientific, and Cultural Organization (UNESCO), 2020. SDG 4 global data book. http://tcg.uis.unesco.org/data-resources/

Wadende, P., Oburu, P. O., & Morara, A. (2016). African indigenous care-giving practices: Stimulating early childhood development and education in Kenya. *South African Journal of Childhood Education*, *6*, 1-7.

Wang, M. T., & Eccles, J. S. (2012). Social support matters: Longitudinal effects of social support on three dimensions of school engagement from middle to high school. *Child Development*, *83*(3), 877-895.

Webster-Stratton, C., & Reid, M. J. (2018). The Incredible Years parents, teachers, and children training series: A multifaceted treatment approach for young children with conduct problems. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 122–141). The Guilford Press.

Weisner, T. S. (1996). The 5 to 7 transition as an ecocultural project.

Weisner, T. S. (2014). The socialization of trust: Plural caregiving and diverse pathways in human development across cultures. In H. Otto & H. Keller (Eds.), *Different faces of attachment: Cultural variations on a universal human need* (pp. 263-277). Cambridge University Press.

Wolf, S., Raza, M., Kim, S., Aber, J. L., Behrman, J., & Seidman, E. (2018). Measuring and predicting process quality in Ghanaian pre-primary classrooms using the Teacher Instructional Practices and Processes System (TIPPS). *Early Childhood Research Quarterly*, *45*, 18-30.

World Bank (2019). *TEACH: Manual.* Author. http://wbgfiles.worldbank.org/documents/hdn/ed/saber/Teach/Manual/132204-WP-PUBLIC-Teach-Manual-English.pdf

World Bank (2021). *TEACH ECE: Manual.* Author.

Yoshikawa, H., Way, N., & Chen, X. (2012). Large-scale economic change and youth development: the case of urban China. *New Directions for Youth Development, 135,* 39-55.

Zieman, G. A. (2012). Participant observation. In Action research methods (pp. 49-67). Palgrave Macmillan, New York.

Zosh, J., Hassinger-Das, B., & Laurie, M. (2022). *Learning through play and the development of holistic skills across childhood*. The LEGO Foundation.

Zosh, J. N., Hopkins, E. J., Jensen, H., Liu, C., Neale, D., Hirsh-Pasek, K., Solis, S.L., & Whitebread, D. (2017). *Learning through play: a review of the evidence*. LEGO Foundation.

Zosh, J. M., Hirsh-Pasek, K., Hopkins, E. J., Jensen, H., Liu, C., Neale, D., ... & Whitebread, D. (2018). Accessing the inaccessible: Redefining play as a spectrum. *Frontiers in Psychology*, *9*, 1124.

APPENDIX A: QUALITATIVE SAMPLE SIZES PER COUNTRY

Colombia

Method	Participant / unit of analysis	Age / level	Sample size
Surveys	Caregivers	12–36 months	n=8
	Teachers	ECE	n=8
Focus groups	Caregivers	12–36 months	n=1 (5 adults)
	Teachers	ECE	n=1 (5 adults)
		Primary	n=1 (5 adults)
Illuminative drawings + focus groups	Children	5–6 and 7–8 years	n=2 (8 children; 4 per focus group)
Observations	Classroom, naturalistic	ECE	j=4
		Primary	j=4
	Teacher-child interactions in classroom	ECE	j=4
		Primary	j=4
	Child point of view	12–36 months	n=4

Note: All teachers in the sample taught children aged 36–72 months (ECE) and 6–8 years (primary). The child's age is included throughout the document for caregiver survey data only; it is not possible to discern child age for teacher-level data (i.e. refers to children aged 36–72 months) or caregiver focus group data (i.e. data recorded at the group level, not the individual level, refers to children aged 12–36 months).

Ghana

Method	Participant / unit of analysis	Sample size
Surveys	Teachers	n=11
Focus groups	Teachers	n=1 (5 adults)
		n=1 (6 adults)
		n=1 (4 adults)
Observations	Classroom, naturalistic	j=9
	Teacher-child interactions in classroom	j=8

Jordan

Method	Participant / unit of analysis	Sample size
Focus groups	Teachers	n=2 (5 adults per focus group)
Observations	Classroom, naturalistic	j=2
	Teacher-child interactions in classroom	j=2

Kenya

Method	Participant / unit of analysis	Sample size
Surveys	Caregivers	n=39
	Teachers	n=28
Focus groups	Caregivers	n=3 (5 adults per focus group)
	Teachers	n=3 (5 adults per focus group)
Illuminative drawings + focus groups	Children	n=2 (4 children per focus group)
Observations	Classroom, naturalistic	j=8
	Teacher-child interactions in classroom	j=8

APPENDIX B: FINDINGS FROM DATA ANALYSES

Introduction

An overarching aim of the analytic work was to understand whether there were any common constructs that supported self-sustained engagement and learning through play across contexts. Therefore, the data analyses sought to identify country-specific patterns as well as best-fitting factor models across countries and developmental levels for each of the tools in the study. The factor models showed us how observed variables (i.e. the items) could be grouped into clusters of items based on some unobserved variable (i.e. the latent constructs). The results could then be compared to our initial hypotheses about the constructs constituting support for children's engagement in learning. Results were more successful with certain tools and developmental contexts than others, and though we always considered the statistical properties of our results, in some cases we presented findings that prioritised the conceptual meaning and strength over the statistical rigor.

Approach to Analysis

For each set of tools, we began by considering six dimensions of adult-child interactions for early childhood education (ECE) settings and seven dimensions for primary schools. Each dimension was hypothesised as being associated with self-sustained engagement and learning through play, according to our conceptual framework. The six dimensions shared between ECE and primary schools include support for agency, support for exploration, support for connection to experience, support for problem solving, support for social connectedness, and positive emotional climate. A seventh dimension, which applied only to primary schools, is support for participation.

Throughout this measurement pilot, items were selected for inclusion in the final measures using ongoing quantitative and qualitative evaluation. Quantitatively, each tool analysis undertook the same general steps:

- Assessment of descriptive statistics, internal consistency, and inter-rater reliability (where applicable)
- Exploratory factor analysis (EFA), based on initial properties, dropping selected items and rerunning EFA
  - In some cases, when no suitable EFA solution was found, a confirmatory factor analysis (CFA) was run utilising a good-fitting solution from another country as an initial basis for analysis.
- Initial CFA, based on preferred final EFA (loadings; model fit; conceptual interpretability)
- Final CFA, based on suggested modifications (e.g. adding correlation terms between items)
- Concurrent validity, based on the correlations between the PLAY observation tool, teacher/caregiver survey, and student survey (if applicable)

Note that for the sake of brevity and the overall clarity of results, this analytic summary will focus primarily on final country-level and/or cross-country factor models and validity results.

Qualitatively, we also considered factors such as the face validity of the items (i.e. how easy are they for an observer, teacher, student to understand the meaning?), feasibility of item/tool administration, as well as the consistency and standardisation of the training

protocols, among other things. For more details on this, please refer to Section 7 of the report on tool development.

Outline and Findings

This Appendix is divided into two main sections: primary and ECE. **Section B.I** on the primary data covers three countries: Colombia, Ghana, and Kenya. Each country's primary dataset analysed three tools: classroom observation, teacher survey, and student survey. The analysis is in three sections: first descriptive and reliability statistics are presented, then the results of factor analyses, and finally, concurrent validity results. For each section, analyses are presented for all three sets of tools in each of the three countries. We found that primary classroom data identified three common constructs across contexts – namely exploration, agency, and social connectedness – as evidenced through observational data. Teacher surveys revealed an additional fourth construct of connection to experience. Finally, student data from the primary survey across the three country contexts revealed two common constructs – agency and positive emotional climate. **Section B.II** on the ECE data also covers three countries: Colombia, Jordan, and Ghana. Each country's dataset analysed one common tool: classroom observation. The teacher survey was implemented in Colombia and Jordan only. And additional caregiver surveys and caregiver observations were implemented in Colombia only, divided into two groups: ages 0–2 and 3–5. The analysis covers descriptive and reliability statistics, factor analyses, and concurrent validity results for all tools in each of the countries, as applicable. We found that ECE classroom data identified three common constructs across contexts – namely connection to experience, exploration, and social connectedness – as evidenced through observational data. Teacher surveys revealed more variability between contexts and touched upon some aspects of all 6 originally hypothesised constructs. Similarly, caregiver observation data did not share the same factor structure between the 0-2 and 3-5 tools. However, caregiver surveys did manifest the same three constructs – support for connection to experience, support for social connectedness, and positive emotional climate – across developmental levels, though the results were all from a single country context of Colombia. It should be noted that in Ghana ECE settings, we were only able to test the observation tool on an existing dataset of classroom videos. However, because these videos were tied to an existing intervention and dataset, we had the opportunity to run additional analyses on treatment impacts and student outcomes – data that were not yet available in other settings.

B.I Findings from the Primary School Tools

B.I.1 Inter-rater Reliability

Table B-1 illustrates the inter-rater reliability (IRR) of items in the PLAY observation tool. Inter-rater reliability tells us the level of agreement between raters or observers. Rater reliability statistics typically range from 0 (0%) to 1 (or 100%). There are also several different methods for calculating IRR, some common methods being percentage agreement, intra-class correlations, or Cohen's Kappa.

Colombia had 30 pairs of inter-rater observations, whereas Ghana had 136 pairs and Kenya had 139 pairs of inter-rater observations. For this analysis, Gwet's agreement coefficient (AC) was adopted to examine the inter-rater reliability, as it holds better statistical properties when dealing with skewed data (Kuppens et al., 2011; Wongpakaran et al., 2013). Gwet's AC bigger than 0.6 suggests substantial to almost perfect inter-rater reliability, whereas Gwet's AC smaller than 0.6 and bigger than 0.4 suggests moderate inter-rater reliability. Gwet's AC smaller than 0.4 suggests poor inter-rater reliability. We deleted items with Gwet's AC smaller than 0.4 in the subsequent EFAs. Generally, agreement rates in Colombia were lower than in Ghana or Kenya. We do not have a clear hypothesis for why this was the case.

Table B-1. Inter-rater reliability (IRR) of items in the PLAY primary observation tool

Construct	Item	Colombia N=30	Ghana N=136	Kenya N=139
		IRR	IRR	IRR
Support for connection to experience	Obs_CE1: Teacher connects concepts in the lesson to everyday objects or spaces that are physically present	0.62	0.97	0.80
	Obs_CE2: Teacher connects concepts in the lesson to the students' interests, background, or life outside the classroom	0.74	0.97	0.65
	Obs_CE3: Teacher connects concepts in the lesson to other subjects, topics, or students' prior knowledge about something already learned	0.27	0.95	0.65
	Obs_CE4: Teacher helps students connect to abstract concepts for which they are concrete and familiar	0.76	0.98	0.89
	Obs_CE5 in Ghana and Kenya: Teacher uses language other than the language of instruction		0.93	0.69
Support for problem solving	Obs_PS1: Teacher poses a problem to students	0.78	0.96	0.92
	Obs_PS2: Teacher gives students hints, suggestions, or feedback to help students get to the answer	0.47	0.94	0.80
	Obs_PS3: Teacher supports students to build on other students' (or their own) answers	0.35	0.98	0.74
	Obs_PS4: Students try different solutions (iteration)	0.64	0.99	0.98
	Obs_PS5: Teacher uses or guides students to use a resource to answer a question	0.43	0.98	0.87

Construct	Item	Colombia N=30	Ghana N=136	Kenya N=139
		IRR	IRR	IRR
Support for exploration	Obs_E1: Teacher gives student(s) exposure to something first before being shown how to use/answer it	0.34	0.99	0.80
	Obs_E2: Teacher uses different/various methods to help students learn about a concept	0.66	0.98	0.92
	Obs_E3: Teacher asks a comparison, categorisation, or prediction question or sets up a task designed to promote students thinking by themselves	0.57	0.98	0.85
	Obs_E4: Teacher gives explicit statements to encourage students to continue to explore a concept	0.41	0.97	0.77
	Obs_E5: Students create something connected to a learning goal	0.63	0.99	0.98
	Obs_E6: Students are guided by the teacher to use symbolic play	0.92	0.99	0.98
Support for agency	Obs_A1: Students choose who plays each role in a group activity	0.67	0.99	0.93
	Obs_A2: Students allowed freedom in approaching an academic task	0.56	1.00	0.91
	Obs_A3 in Colombia: Teacher gives individual or a limited number of students responsibility	0.71		
	Obs_A3 in Kenya and Obs_A4 in Colombia: Students' ideas influence teacher's instruction	0.57	0.98	0.92
	Obs_A4 in Kenya and Ghana but Obs_A5 in Colombia: Students express their own ideas or otherwise contribute to class without teacher prompting	0.51	0.98	0.91
	Obs_A5 in Ghana and Kenya but Obs_A6 in Colombia: Students interact with one another without specific direction from the teacher	0.26	0.97	0.69
	Obs_A6 in Ghana and Kenya: Teacher allows students freedom in student movement		0.97	0.72
Support for participation	Obs_P1: Students respond to questions in ways that are not teacher -- student oral response	0.85	0.96	0.76
	Obs_P2: Students practice a new skill introduced by the teacher	0.63	0.95	0.76
	Obs_P3: Students respond to opportunities (from the teacher) to express their own ideas	0.11	0.97	0.69
	Obs_P4: Teacher invites student questions	0.85	0.97	0.95
Support for social connectedness	Obs_SC1: Students work together on an exercise/project or towards a common goal	0.66	0.97	0.87
	Obs_SC2: Students are instructed by the teacher to discuss topics with each other	0.54	0.99	0.87
	Obs_SC3: Student or a group of students assist each other in need	0.63	0.97	0.63
	Obs_SC4: Students demonstrate togetherness or camaraderie	0.38	0.98	0.94
	Obs_SC5: Teacher discusses or otherwise creates a sense of student/class togetherness	0.49	0.99	0.97
	Obs_SC6: Teacher uses physical space to promote interaction	0.81	0.97	0.96

Construct	Item	Colombia N=30	Ghana N=136	Kenya N=139
		IRR	IRR	IRR
	Obs_SC7: Teacher gives individual or a limited number of students responsibility that other students do not have		0.96	0.59
Support for positive emotional climate	Obs_PC1: Teacher uses a mode of instruction that is explicitly joyous throughout the lesson	0.36	0.96	0.74
	Obs_PC2: Teacher provides a special honour to a student(s)	0.86	0.97	0.92
	Obs_PC3: Teacher gives praise or encourages the class to give praise to themselves	0.35	0.93	0.58
	Obs_PC4: Teacher uses student names	0.68	0.97	0.81
	Obs_PC5: Teacher guides students with songs or energisers to start or divide activity	0.49	0.96	0.80
	Obs_PC6: Teacher shows awareness of student emotions	0.85	0.96	0.84
	Obs_PC7: Teacher includes students who did not volunteer to answer (i.e. inclusion)	0.47	0.93	0.64
	Obs_PC8: Teacher provides a risk-free environment for participation	0.22	0.99	0.55
	Obs_PC9: Teacher uses physical proximity to show closeness to students	0.51	0.96	0.61

Note: Pink indicates that the item was not tested in the country.

B.I.2 Internal Consistency

Internal consistency, which is based on item correlations, assesses how much the items within an instrument measure the same construct or characteristic.

Besides inter-rater reliability, Cronbach's alpha was conducted to examine the internal consistency of each construct of the primary observation tool, as shown in in **Table B-2**. Most of the constructs had Cronbach's alpha smaller than 0.6, which suggested poor internal consistency of the originally hypothesised constructs.

Table B-2. Internal consistency of hypothesised constructs in the PLAY primary observation tool

Construct	Cronbach's alpha		
	Colombia (N=76)	Ghana (N=278)	Kenya (N=280)
Support for connection to experience	0.39	0.17	0.3
Support for problem solving	0.58	0.48	0.49
Support for exploration	0.51	0.29	0.28
Support for agency	0.44	0.48	0.41
Support for participation	0.44	0.16	0.18
Support for social connectedness	0.69	0.53	0.23
Support for positive emotional climate	0.61	0.31	0.4

The internal consistency of each construct of the primary teacher-report is shown in **Table B-3**. Some of the constructs had Cronbach's alpha smaller than 0.6, which suggested poor internal consistency of the originally hypothesised constructs.

Table B-3. Internal consistency of hypothesised constructs in PLAY primary teacher-report data

Construct	Cronbach's alpha		
	Colombia (N=76)	Ghana (N=278)	Kenya (N=280)
Support for connection to experience	0.69	0.57	0.57
Support for problem solving	0.58	0.63	0.58
Support for exploration	0.68	0.68	0.66
Support for agency	0.43	0.65	0.68
Support for participation	0.61	0.73	0.5
Support for social connectedness	0.66	0.79	0.59
Support for positive emotional climate	0.55	0.59	0.43

The internal consistency of each construct in the primary student-report is shown in **Table B-4**. The majority of the constructs had Cronbach's alpha smaller than 0.6, which suggested poor internal consistency of the originally hypothesised constructs.

Table B-4. Internal consistency of each construct in PLAY primary student-report data

Construct	Cronbach's alpha		
	Colombia (N=380)	Ghana (N=1,114)	Kenya (N=1,120)
Support for connection to experience	0.41	0.51	0.51
Support for problem solving	0.46	0.59	0.52
Support for exploration	0.38	0.58	0.59
Support for agency	0.32	0.36	0.45
Support for participation	0.38	0.47	0.51
Support for social connectedness	0.45	0.57	0.63
Support for positive emotional climate	0.55	0.5	0.6

B.I.3 Final Factor Analysis Models for Primary School Tools

B.I.3.1 *Summary of Compromised-Fitting PLAY Tools across Countries*

Colombia, Ghana, and Kenya shared certain similarities with regards to the best-fitting PLAY observation tool, teacher survey, and student survey models. To maintain consistency across countries, we proposed a compromised-fit model in each country for each toolkit, after accounting for their conceptual meanings. However, in order to find a common model across the countries, it should be noted that this resulted in a model with some items with quite low factor loadings, as well as a poorer model fit overall. The new CFA loadings and fit statistics can be found in **Tables B-5**, **B-6**, and **B-7**.

Table B-5. Standardised CFA loadings of the common CFA model for PLAY primary observation tool in each country

Kenya		Ghana		Colombia		Factor
E1: Teacher gives student(s) exposure to something first before being shown how to use/answer it	0.47**	E1: Teacher gives student(s) exposure to something first before being shown how to use/answer it	0.08	E1: Teacher gives student(s) exposure to something first before being shown how to use/answer it	0.27*	**Obs1_support for exploration** Kenya alpha=0.515 Ghana alpha=0.29 Colombia alpha=0.66
P3: Students respond to opportunities (from the teacher) to express their own ideas	0.76***	P3: Students respond to opportunities (from the teacher) to express their own ideas	0.48***	P3: Students respond to opportunities (from the teacher) to express their own ideas	0.55***	
E4: Teacher gives explicit statements to encourage students to continue to explore a concept	0.32***	E4: Teacher gives explicit statements to encourage students to continue to explore a concept	0.67***	E4: Teacher gives explicit statements to encourage students to continue to explore a concept	0.79*	
CE2: Teacher connects concepts in the lesson to the students' interests, background or life outside the classroom	0.35***	CE2: Teacher connects concepts in the lesson to the students' interests, background or life outside the classroom	0.22*	CE2: Teacher connects concepts in the lesson to the students' interests, background or life outside the classroom	0.31**	
E3: Teacher asks a comparison, categorisation, or prediction question or sets up a task designed to promote students thinking by themselves	0.07	E3: Teacher asks a comparison, categorisation, or prediction question or sets up a task designed to promote students thinking by themselves	0.11	E3: Teacher asks a comparison, categorisation, or prediction question or sets up a task designed to promote students thinking by themselves	0.74***	
SC1: Students work together on an exercise/project or towards a common goal	0.13	SC1: Students work together on an exercise/project or towards a common goal	0.59***	SC1: Students work together on an exercise/project or towards a common goal	0.83***	**Obs2_support for student-led activities/agency** Kenya alpha=0.31 Ghana alpha=0.54 Colombia alpha=0.72
A1: Students choose who plays each role in a group activity	0.36***	A1: Students choose who plays each role in a group activity	0.7***	A1: Students choose who plays each role in a group activity	0.52***	
A2: students decide what or how to do an academic task	0.33***	A2: Students decide what or how to do an academic task	0.28***	A2: Students decide what or how to do an academic task	0.34**	
SC6: Teacher uses physical space to promote interaction	-0.015	SC6: Teacher uses physical space to promote interaction	0.39***	SC6: Teacher uses physical space to promote interaction	0.57***	
PS4: Students try different solutions (iteration)	0.37***	PS4: Students try different solutions (iteration)	0.22**	PS4: Students try different solutions (iteration)	0.65***	
SC2: Students are instructed by the teacher to discuss topics with each other	-0.06	SC2: Students are instructed by the teacher to discuss topics with each other	0.41***	SC2: Students are instructed by the teacher to discuss topics with each other	0.53***	

Kenya		Ghana		Colombia		Factor
A3 in Kenya and A4 in Colombia: Students' ideas influence teacher's instruction	0.58***	A3 in Kenya and A4 in Colombia: Students' ideas influence teacher's instruction	0.18*	A3 in Kenya and A4 in Colombia: Students' ideas influence teacher's instruction	0.33**	
E5: Students create something connected to a learning goal	-0.01	E5: Students create something connected to a learning goal	0.3***	E5: Students create something connected to a learning goal	0.34**	
P2: Students practice a new skill introduced by the teacher	0.19*	P2: Students practice a new skill introduced by the teacher	0.16*	P2: Students practice a new skill introduced by the teacher	0.28*	
A4 in Kenya and Ghana but A5 in Colombia: Students express their own ideas or otherwise contribute to class without teacher prompting	0.52***	A4 in Kenya and Ghana but A5 in Colombia: Students express their own ideas or otherwise contribute to class without teacher prompting	0.07	A4 in Kenya and Ghana but A5 in Colombia: Students express their own ideas or otherwise contribute to class without teacher prompting	0.31**	
PC9: Teacher uses physical proximity to show closeness to students	0.28**	PC9: Teacher uses physical proximity to show closeness to students	0.2**	PC9: Teacher uses physical proximity to show closeness to students	0.31*	**Obs3_support for togetherness and cooperation** Kenya alpha=0.25 Ghana alpha=0.45 Colombia alpha=0.45
E2: Teacher uses different/various methods to help students learn about a concept	0.42***	E2: Teacher uses different/various methods to help students learn about a concept	0.5***	E2: Teacher uses different/various methods to help students learn about a concept	0.23†	
SC4: Students demonstrate togetherness or camaraderie	0.31**	SC4: Students demonstrate togetherness or camaraderie	0.7***	SC4: Students demonstrate togetherness or camaraderie	0.54***	
SC5: Teacher discusses or otherwise creates a sense of student/class togetherness	0.34**	SC5: Teacher discusses or otherwise creates a sense of student/class togetherness	0.55***	SC5: Teacher discusses or otherwise creates a sense of student/class togetherness	0.61***	
Chi-square (df)	212.6 (148)	**Chi-square (df)**	247.1 (147)	**Chi-square (df)**	217.86 (147)	
p-value	0.0004	**p-value**	<0.0001	**p-value**	0.0001	
CFI	0.75	**CFI**	0.79	**CFI**	0.78	
RMSEA	0.039	**RMSEA**	0.049	**RMSEA**	0.08	
SRMR	0.059	**SRMR**	0.063	**SRMR**	0.1	

Note: † p<0.1, * p<0.05, ** p<0.01, *** p<0.001
Fonts in red suggest insignificant loading values.

Factor 1: Support for exploration
Factor 2: Support for student-led activities/agency
Factor 3: Support for togetherness and cooperation

Table B-6. Standardised CFA loadings of the common CFA model for PLAY primary teacher survey in each country

Kenya		Ghana		Colombia		Factor
C4 in Kenya and Ghana and C5 in Colombia: I encourage learners to understand and empathise with their classmates.	0.557***	C4 in Kenya and Ghana and C5 in Colombia: I encourage learners to understand and empathise with their classmates.	0.564***	C4 in Kenya and Ghana and C5 in Colombia: I encourage learners to understand and empathise with their classmates.	0.524***	**Tr1_support for togetherness and cooperation** Kenya alpha=0.59 Ghana alpha=0.79 Colombia alpha=0.6
SC5: I encourage my learners to work together to achieve common goals.	0.599***	SC5: I encourage my learners to work together to achieve common goals.	0.694***	SC5: I encourage my learners to work together to achieve common goals.	0.575***	
SC4: I create a sense of class togetherness.	0.478***	SC4: I create a sense of class togetherness.	0.65***	SC4: I create a sense of class togetherness.	0.331***	
SC7: I encourage learners to help each other.	0.662***	SC7: I encourage learners to help each other.	0.687***	SC7: I encourage learners to help each other.	0.469***	
SC6: I give learners exercises to work on together in groups.	0.46***	SC6: I give learners exercises to work on together in groups.	0.518***	SC6: I give learners exercises to work on together in groups.	0.483***	
SC3: I encourage learners to listen to each other.	0.366***	SC3: I encourage learners to listen to each other.	0.609***	SC3: I encourage learners to listen to each other.	0.44***	
SC2: I encourage learners to have interest in each other's lives.	0.33***	SC2: I encourage learners to have interest in each other's lives.	0.569***	SC2: I encourage learners to have interest in each other's lives.	0.618***	
C5 in Ghana and Kenya and C6 in Colombia: I encourage learners to answer questions, even if they don't know the correct answer.	0.14*	C5 in Ghana and Kenya and C6 in Colombia: I encourage learners to answer questions, even if they don't know the correct answer.	0.261***	C5 in Ghana and Kenya and C6 in Colombia: I encourage learners to answer questions, even if they don't know the correct answer.	0.304*	**Tr2_support for questioning children and prompting discussion** Kenya alpha=0.25 Ghana alpha=0.45 Colombia alpha=0.51
SC1: I encourage learners to have discussions with each other without my involvement.	0.429***	SC1: I encourage learners to have discussions with each other without my involvement.	0.448***	SC1: I encourage learners to have discussions with each other without my involvement.	0.664***	
P4: I encourage learners to express their own ideas.	0.415***	P4: I encourage learners to express their own ideas.	0.6***	P4: I encourage learners to express their own ideas.	0.463***	
P6: I ask open-ended questions to encourage learners' contributions.	0.221***	P6: I ask open-ended questions to encourage learners' contributions.	0.527***	P6: I ask open-ended questions to encourage learners' contributions.	0.481***	

Kenya		Ghana		Colombia		Factor
CE4: I use learners' experiences outside of school in my lessons.	0.488***	CE4: I use learners' experiences outside of school in my lessons.	0.375***	CE4: I use learners' experiences outside of school in my lessons.	0.618***	**Tr3_support for connection to experience and ideas** Kenya alpha=0.56 Ghana alpha=0.65 Colombia alpha=0.7
CE3: I use learners' background and interests in developing lessons/learning objectives.	0.4***	CE3: I use learners' background and interests in developing lessons/learning objectives.	0.406***	CE3: I use learners' background and interests in developing lessons/learning objectives.	0.57***	
CE2: I use objects or actions to make connections for learners to their prior knowledge.	0.389***	CE2: I use objects or actions to make connections for learners to their prior knowledge.	0.481***	CE2: I use objects or actions to make connections for learners to their prior knowledge.	0.586***	
PS2: I encourage learners to form their own ideas about a problem before I give them the answer.	0.422***	PS2: I encourage learners to form their own ideas about a problem before I give them the answer.	0.567***	PS2: I encourage learners to form their own ideas about a problem before I give them the answer.	0.341***	
PS3: I help learners build on each other's ideas to solve problems.	0.437***	PS3: I help learners build on each other's ideas to solve problems.	0.662***	PS3: I help learners build on each other's ideas to solve problems.	0.499***	
CE1: I relate concepts I teach to everyday objects in the classroom or school grounds.	0.416***	CE1: I relate concepts I teach to everyday objects in the classroom or school grounds.	0.349***	CE1: I relate concepts I teach to everyday objects in the classroom or school grounds.	0.612***	
A1: When a group activity has several roles, I let learners choose who does what.	0.451***	A1: When a group activity has several roles, I let learners choose who does what.	0.506***	A1: When a group activity has several roles, I let learners choose who does what.	0.615***	**Tr4_support for child agency** Kenya alpha=0.63 Ghana alpha=0.7 Colombia alpha=0.74
E1: I have learners use role-play or pretend to be someone or something else.	0.43***	E1: I have learners use role-play or pretend to be someone or something else.	0.504***	E1: I have learners use role-play or pretend to be someone or something else.	0.629***	
E2: I give learners a chance to investigate something first before being shown how to use/answer it.	0.463***	E2: I give learners a chance to investigate something first before being shown how to use/answer it.	0.535***	E2: I give learners a chance to investigate something first before being shown how to use/answer it.	0.823***	
P2: I allow learners to practice new skills on their own.	0.45***	P2: I allow learners to practice new skills on their own.	0.556***	P2: I allow learners to practice new skills on their own.	0.535***	
P5: I use various teaching methods, so learners have different ways to contribute during class.	0.357***	P5: I use various teaching methods, so learners have different ways to contribute during class.	0.549***	P5: I use various teaching methods, so learners have different ways to contribute during class.	0.627***	

Kenya		Ghana		Colombia		Factor
E4: I create or use playful activities for learners to practice what they learn in class.	0.384***	E4: I create or use playful activities for learners to practice what they learn in class.	0.4***	E4: I create or use playful activities for learners to practice what they learn in class.	0.495***	
A2: I let learners decide for themselves how they go about a task.	0.39***	A2: I let learners decide for themselves how they go about a task.	0.464***	A2: I let learners decide for themselves how they go about a task.	0.302**	
A4: I change my teaching based on learners' ideas and suggestions.	0.349***	A4: I change my teaching based on learners' ideas and suggestions.	0.329***	A4: I change my teaching based on learners' ideas and suggestions.	0.537***	
Chi-square (df)	345.023 (264)	**Chi-square (df)**	465.616 (265)	**Chi-square (df)**	409.447 (268)	
p-value	0.0006	**p-value**	<0.0001	**p-value**	<0.0001	
CFI	0.906	**CFI**	0.881	**CFI**	0.7	
RMSEA	0.033	**RMSEA**	0.052	**RMSEA**	0.084	
SRMR	0.054	**SRMR**	0.052	**SRMR**	0.094	

Note: * p<0.05, ** p<0.01, *** p<0.001

- Factor 1: Support for togetherness and cooperation
- Factor 2: Support for questioning children and prompting discussion
- Factor 3: Support for connection to experience and ideas
- Factor 4: Support for child agency

Table B-7. Standardised CFA loadings of the common CFA model for PLAY primary student survey in each country

Kenya		Ghana		Colombia		Factor
C1: Your teacher cares about you.	0.559***	C1: Your teacher cares about you.	0.392***	C1: Your teacher cares about you.	0.374***	**Sr1_support for positive climate** Kenya alpha=0.79 Ghana alpha=0.74 Colombia alpha=0.69
C3: Your teacher gives learners lots of praise.	0.484***	C3: Your teacher gives learners lots of praise.	0.421***	C3: Your teacher gives learners lots of praise.	0.434***	
C4: You like your teacher.	0.473***	C4: You like your teacher.	0.457***	C4: You like your teacher.	0.455***	
C5: You have fun during lessons.	0.51***	C5: You have fun during lessons.	0.48***	C5: You have fun during lessons.	0.522***	
SC2: Your classmates and you often work together towards a common goal.	0.515***	SC2: Your classmates and you often work together towards a common goal.	0.524***	SC2: Your classmates and you often work together towards a common goal.	0.424***	
SC5: Learners in your class help one another.	0.521***	SC5: Learners in your class help one another.	0.476***	SC5: Learners in your class help one another.	0.389***	
P4: Your teacher asks learners if they have questions.	0.433***	P4: Your teacher asks learners if they have questions.	0.431***	P4: Your teacher asks learners if they have questions.	0.395***	
A2: You discuss what you learn with other learners during the lesson.	0.521***	A2: You discuss what you learn with other learners during the lesson.	0.481***	A2: You discuss what you learn with other learners during the lesson.	0.303***	
P5: During the lesson, you get to practice the skills you learn.	0.53***	P5: During the lesson, you get to practice the skills you learn.	0.476***	P5: During the lesson, you get to practice the skills you learn.	0.534***	
SC1: Your class has a sense of togetherness.	0.515***	SC1: Your class has a sense of togetherness.	0.42***	SC1: Your class has a sense of togetherness.	0.389***	
SC4: Learners in your class work well together on exercises.	0.49***	SC:4 Learners in your class work well together on exercises.	0.466***	SC4: Learners in your class work well together on exercises.	0.4***	
PS3: Your teacher encourages you to find answers on your own.	0.44***	PS3: Your teacher encourages you to find answers on your own.	0.453***	PS3: Your teacher encourages you to find answers on your own.	0.395***	**Sr2_support for agency** Kenya alpha=0.55 Ghana alpha=0.34 Colombia alpha=0.26
P2: You share your opinions in class.	0.497***	P2: You share your opinions in class.	0.451***	P2: You share your opinions in class.	0.347***	
P3: During lessons, learners ask a lot of questions to each other.	0.426***	P3: During lessons, learners ask a lot of questions to each other.	0.246***	P3: During lessons, learners ask a lot of questions to each other.	0.297***	
A3: You can leave your desk when you want to get learning materials that are in the classroom.	0.322***	A3: You can leave your desk when you want to get learning materials that are in the classroom.	0.311***	A3: You can leave your desk when you want to get learning materials that are in the classroom.	-0.086	

Kenya		Ghana		Colombia		Factor
A4: You are allowed to do class exercises in your own way.	0.518***	A4: You are allowed to do class exercises in your own way.	0.03	A4: You are allowed to do class exercises in your own way.	0.036	
Chi-square (df)	302.96 (103)	**Chi-square (df)**	283.324 (103)	**Chi-square (df)**	157.756 (102)	
p-value	<0.0001	**p-value**	<0.0001	**p-value**	0.0003	
CFI	0.924	**CFI**	0.907	**CFI**	0.901	
RMSEA	0.042	**RMSEA**	0.04	**RMSEA**	0.038	
SRMR	0.037	**SRMR**	0.035	**SRMR**	0.045	

Note: * p<0.05, ** p<0.01, *** p<0.001

Factor 1: Support for positive climate

Factor 2: Support for agency

B.1.4 Construct Validity

Construct validity tells us how well a tool has identified measurable characteristics that reflect the true meaning of a concept. In other words, construct validity is how well a tool, effectively measures what it claims. We can generally verify construct validity by comparing a measure to other measures of similar constructs and comparing the correlations between the two (or more).

B.1.4.1 Colombia

B.I.1.1.1 Concurrent Validity between Observation Tool, Teacher Survey, and Student Survey

In this section, **Tables B-8** and **B-9** show the concurrent correlations and multilevel mixed-effects linear regressions between the best-fitting PLAY observation tool, teacher survey, and student survey models in Colombia. **Figure B-1** demonstrates the distributions of the three factors extracted from the compromised-fitting PLAY observation tool, four factors extracted from the compromised-fitting PLAY teacher survey, and two factors extracted from the compromised-fitting PLAY student survey in the Colombian context.

Table **B-8** shows the concurrent spearman's rank correlations between the classroom-level PLAY observation tool and teacher survey. There were significant moderate to strong correlations among the three factors within the PLAY observation tool. Support for exploration strongly and positively correlated with support for student-led activities/agency (r=0.678, p<0.001). Support for exploration also strongly correlated with support for togetherness and cooperation (r=0.659, p<0.001). In addition, support for student-led activities/agency also strongly correlated with support for togetherness and cooperation (r=0.848, p<0.001).

Table B-8. Correlations between primary classroom-level observation tool and teacher survey in Colombia

	Obs1_support for exploration	Obs2_support for student-led activities/agency	Obs3_support for togetherness and cooperation	Tr1_support for togetherness and cooperation	Tr2_support for questioning children and prompting discussion	Tr3_support for connection to experience and ideas	Tr4_support for child agency
Obs1_support for exploration	1						
Obs2_support for student-led activities/agency	0.678***	1					
Obs3_support for togetherness and cooperation	0.659***	0.848***	1				
Tr1_support for togetherness and cooperation	-0.056	0.43***	0.257**	1			
Tr2_support for questioning children and prompting discussion	-0.23†	0.189	0.052	0.908***	1		
Tr3_support for connection to experience and ideas	0.063	0.5***	0.563***	0.76***	0.597***	1	
Tr4_support for child agency	-0.079	0.482***	0.257*	0.838***	0.799***	0.718***	1

Note: † p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Among the four factors within the PLAY teacher survey, there were significant strong correlations between the factors. Support for togetherness strongly correlated with questioning children and prompting discussion (r=0.908, p<0.001), support for connection to experience and ideas (r=0.76, p<0.001), and support for child agency (r=0.838, p<0.001). Support for questioning children and prompting discussion also moderately correlated with support for connection to experience and ideas (r=0.597, p<0.001) and support for child agency (r=0.799, p<0.001). Between support for connection to experience and ideas, and support for child agency, there was strong and positive correlation (r=0.718, p<0.001).

Between PLAY observation tool factors and PLAY teacher survey factors, observed support for student-led activities/agency was moderately correlated with teacher-report support for togetherness and cooperation (r=0.43, p<0.001), teacher-report for connection to experience and ideas (r=0.5, p<0.001), and teacher-report support for agency (r=0.482, p<0.001). Observed support for togetherness and cooperation also moderately correlated with teacher-report support for connection to experience and ideas (r=0.563, p<0.001). There were also positive but negligible correlations between observed support for togetherness and cooperation, teacher-report support for togetherness and cooperation, and teacher-report support for agency.

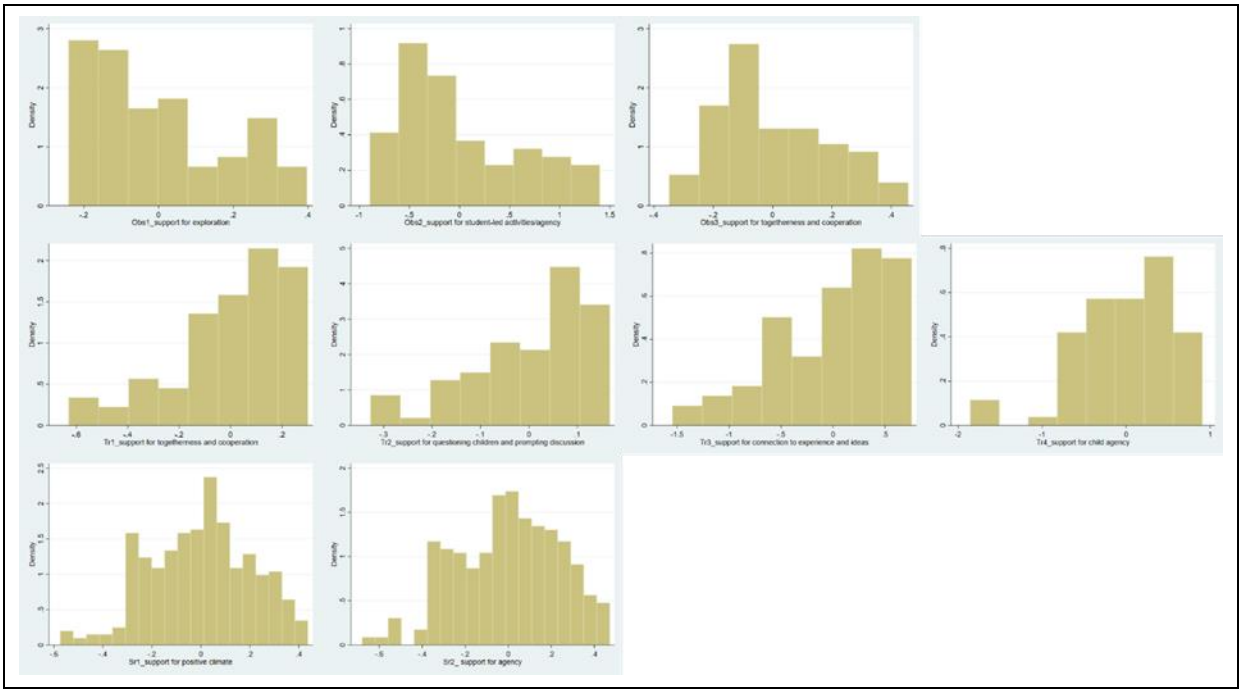
Table B-9 demonstrates the results from multilevel mixed-effects linear regressions, where the group-level variable is the classroom. The majority had no significant correlations between two factors extracted from the PLAY student survey and factors extracted from the PLAY observation tool and teacher survey.

Table B-9. Results of multilevel mixed-effects linear regressions between primary classroom-level observation tool/teacher survey and classroom-level student survey in Colombia

	Sr1_support for positive climate	**Sr2_ support for agency**
Sr2_support for agency		1
Obs1_support for exploration	-0.061	-0.059
Obs2_support for student-led activities/agency	0.004	0.004
Obs3_support for togetherness and cooperation	0.054	0.068
Tr1_support for togetherness and cooperation	-0.063	-0.072
Tr2_support for questioning children and prompting discussion	-0.164†	-0.184†
Tr3_support for connection to experience and ideas	0.008	0.013
Tr4_support for child agency	-0.021	-0.023

Note: † p<0.1

Figure B-1. Distributions of factors extracted from the compromised-fitting PLAY primary observation tool, teacher survey, and student survey in Colombia



B.I.1.1.2 Correlations between Observation Tool/Teacher Survey and Teacher's Sociodemographic

Table B-10 demonstrates the concurrent spearman's rank correlations between three PLAY observation tool factors, four PLAY teacher survey factors, and seven sociodemographic factors. Most of the correlations were insignificant, particularly those associated with school grade, having students from Venezuela, teacher training on primary education, and school type. More years of being a teacher were moderately and positively correlated with teacher-report support for questioning children and prompting discussion (r=0.398, p<0.001). Compared to male teachers, female teachers were more likely to be observed supporting for togetherness and cooperation (r=0.363, p<0.001). There were also gender effects on observed support for exploration and support for student-led activities/agency, but their associations were negligible, at 0.3 level.

Table B-10. Correlations between primary classroom-level observation tool/teacher survey and sociodemographic backgrounds in Colombia

	Obs1_support for exploration	Obs2_support for student-led activities/agency	Obs3_support for togetherness and cooperation	Tr1_support for togetherness and cooperation	Tr2_support for questioning children and prompting discussion	Tr3_support for connection to experience and ideas	Tr4_support for child agency
Years of being a teacher	-0.188	-0.169	-0.248*	0.277*	0.398***	0.12	0.15
Gender (reference: male)	0.254*	0.262*	0.363**	0.069	-0.013	0.261*	0.014
Grade taught by teacher	-0.041	-0.049	-0.069	0.081	0.052	-0.152	0.072
Teacher's highest level of education	-0.227	-0.289*	-0.202	0.0004	-0.005	-0.051	-0.108
Have students from Venezuela (reference: no)	-0.139	-0.086	0.05	0.03	-0.058	0.102	-0.025
Received training on primary education (reference: no)	-0.092	-0.229	-0.189	-0.121	-0.042	-0.132	-0.171
School type (reference: rural)	0.073	0.011	0.036	0.124	0.089	0.191	0.199

Note: * p<0.05, ** p<0.01, *** p<0.001

B.I.4.2 Ghana

B.I.4.2.1 Concurrent Validity between Observation Tool, Teacher Survey, and Student Survey

In this section, **Tables B-11** and **B-12** show the concurrent correlations and multilevel mixed-effects linear regressions between the best-fitting PLAY observation tool, teacher survey, and student survey models in Ghana. **Figure B-2** demonstrates the distributions of the three factors extracted from the compromised-fitting PLAY observation tool, four factors extracted from the compromised-fitting PLAY teacher survey, and two factors extracted from the compromised-fitting PLAY student survey in the Ghanaian context.

Table B-11 shows the concurrent spearman's rank correlations between the classroom-level PLAY observation tool and teacher survey. There were significant moderate to strong correlations among the three factors within the PLAY observation tool. Support for exploration moderately and positively correlated with support for student-led activities/agency (r=0.41, p<0.001). Support for exploration also moderately correlated with support for togetherness and cooperation (r=0.53, p<0.001). In addition, support for student-led activities/agency also strongly correlated with support for togetherness and cooperation (r=0.77, p<0.001). Additionally, it is worth noting there were small but consistently negative correlations between observed support for exploration and all teacher-report factors, which will require further exploration.

Table B-11. Concurrent correlations between primary classroom-level observation tool and teacher survey in Ghana

	Obs1_support for exploration	Obs2_support for student-led activities/agency	Obs3_support for togetherness and cooperation	Tr1_support for togetherness and cooperation	Tr2_support for questioning children and prompting discussion	Tr3_support for connection to experience and ideas	Tr4_support for child agency
Obs1_support for exploration	1						
Obs2_support for student-led activities/agency	0.407***	1					
Obs3_support for togetherness and cooperation	0.529***	0.772***	1				
Tr1_support for togetherness and cooperation	-0.179**	0.139*	0.238***	1			
Tr2_support for questioning children and prompting discussion	-0.167**	0.059	0.204***	0.952***	1		
Tr3_support for connection to experience and ideas	-0.153*	0.131*	0.266***	0.82***	0.91***	1	
Tr4_support for child agency	-0.228***	0.032	0.179**	0.94***	0.95***	0.89***	1

Note: * p<0.05, ** p<0.01, *** p<0.001

Among four factors within the PLAY teacher survey, there were significant strong correlations between the factors. Support for togetherness strongly correlated with questioning children and prompting discussion (r=0.95, p<0.001), support for connection to experience and ideas (r=0.82, p<0.001), and support for child agency (r=0.94, p<0.001). Support for questioning children and prompting discussion also moderately correlated with support for connection to experience and ideas (r=0.91, p<0.001) and support for child agency (r=0.95, p<0.001). Between support for connection to experience and ideas, and support for child agency, there was strong and positive correlation (r=0.89, p<0.001).

Between the PLAY observation tool factors and PLAY teacher survey factors, there were some significant correlations, but they were considered negligible, at 0.3 level.

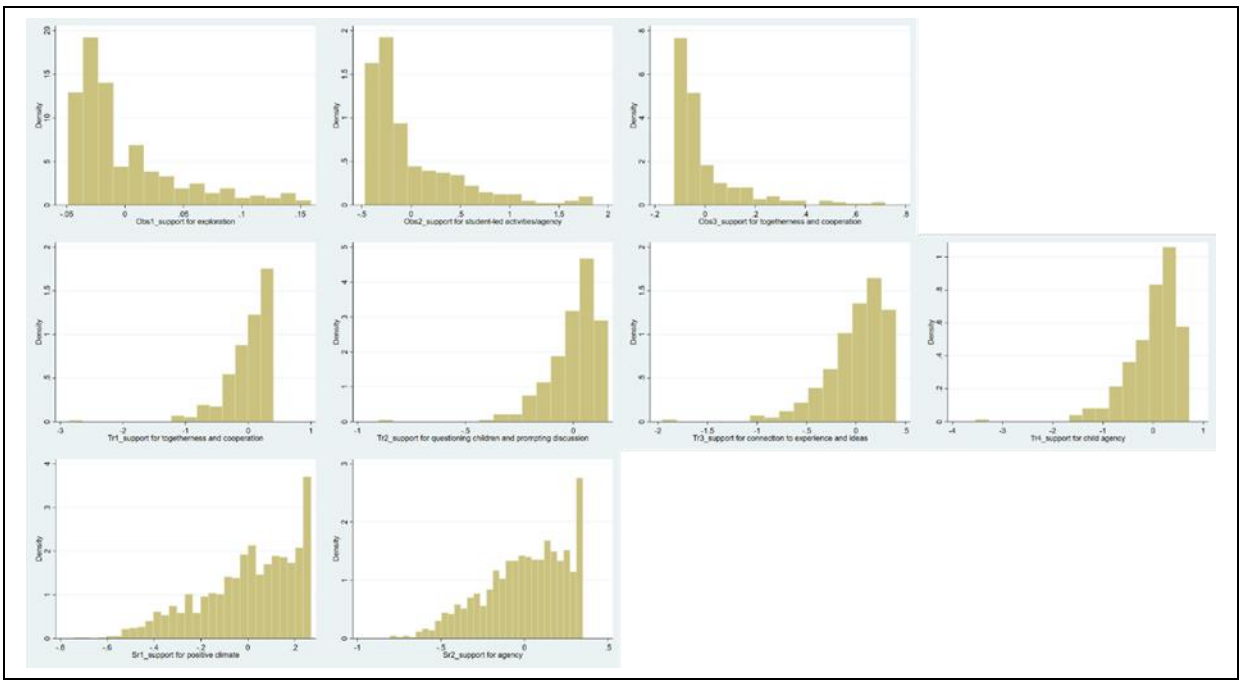
Table B-12 demonstrates the results from multilevel mixed-effects linear regressions, where the group-level variable is the classroom. The majority had no significant correlations between the two factors extracted from the PLAY student survey and the factors extracted from the PLAY observation tool and teacher survey.

Table B-12. Results of multilevel mixed-effects linear regressions between primary classroom-level observation tool/teacher survey and classroom-level student survey in Ghana

	HLM two variables	
	Sr1_support for positive climate	**Sr2_support for agency**
Sr2_support for agency	0.98***	1
Obs1_support for exploration	-0.027	-0.005
Obs2_support for student-led activities/agency	0.0049	0.0073
Obs3_support for togetherness and cooperation	0.0045	0.0055
Tr1_support for togetherness and cooperation	0.043†	0.046†
Tr2_support for questioning children and prompting discussion	0.123†	0.134
Tr3_support for connection to experience and ideas	0.047†	0.052
Tr4_support for child agency	0.025	0.028

Note: † p<0.1, *** p<0.001

Figure B-2. Distributions of factors extracted from the of the common PLAY primary observation tool, teacher survey, and student survey in Ghana



B.I.4.2.2 Correlations between Observation Tool/Teacher Survey and Teacher's Sociodemographic

Table B-13 demonstrates the concurrent spearman's rank correlations between three PLAY observation tool factors, four PLAY teacher survey factors, and five sociodemographic factors. Most of the correlations were insignificant, particularly those associated with years of being a teacher and whether teacher received training on primary education. Although there were a few significant correlations, they were considered negligible, at 0.3 level.

Table B-13. Correlations between primary classroom-level observation tool/ teacher survey and sociodemographic backgrounds in Ghana

	Obs1_support for exploration	Obs2_support for student-led activities/agency	Obs3_support for togetherness and cooperation	Tr1_support for togetherness and cooperation	Tr2_support for questioning children and prompting discussion	Tr3_support for connection to experience and ideas	Tr4_support for child agency
Years of being a teacher	-0.078	0.048	0.003	-0.023	-0.045	-0.067	-0.036
Gender (reference: male)	-0.024	0.112	0.132*	0.126*	0.087	0.101	0.109
Grade taught by teacher	0.024	-0.028	-0.076	-0.049	-0.079	-0.129*	-0.058
Teacher's highest level of education	0.039	0.154*	0.122*	0.086	0.086	0.079	0.065
Received training on primary education (reference: no)	-0.042	0.036	0.087	0.084	0.058	0.044	0.075

Note: * p<0.05

B.I.4.3 *Kenya*

B.I.4.3.1 Concurrent Validity between Observation Tool, Teacher Survey, and Student Survey

In this section, **Tables B-14** and **B-15** show the concurrent correlations and multilevel mixed-effects linear regressions between the best-fitting PLAY observation tool, teacher survey, and student survey models in Kenya. **Figure B-3** demonstrates the distributions of the three factors extracted from the compromised-fitting PLAY observation tool, four factors extracted from the compromised-fitting PLAY teacher survey, and two factors extracted from the compromised-fitting PLAY student survey in the Kenyan context.

Table B-14 shows the concurrent spearman's rank correlations between classroom-level PLAY observation tool and teacher survey. There was significant and moderate correlation only between observed support for student-led activities/agency and support for togetherness and cooperation (r=0.32, p<0.001).

Table B-14. Correlations between primary classroom-level observation tool and teacher survey in Kenya

	Obs1_support for exploration	Obs2_support for student-led activities/agency	Obs3_support for togetherness and cooperation	Tr1_support for togetherness and cooperation	Tr2_support for questioning children and prompting discussion	Tr3_support for connection to experience and ideas	Tr4_support for child agency
Obs1_support for exploration	1						
Obs2_support for student-led activities/agency	0.183**	1					
Obs3_support for togetherness and cooperation	0.101†	0.321***	1				
Tr1_support for togetherness and cooperation	0.034	0.068	0.203***	1			
Tr2_support for questioning children and prompting discussion	0.134*	0.042	0.367***	0.768***	1		
Tr3_support for connection to experience and ideas	0.298***	0.229***	0.121*	0.716***	0.719***	1	
Tr4_support for child agency	0.253***	0.031	0.387***	0.691***	0.925***	0.76***	1

Note: † p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Among the four factors within the PLAY teacher survey, there were significant strong correlations between the factors. Support for togetherness strongly correlated with questioning children and prompting discussion (r=0.77, p<0.001), support for connection to experience and ideas (r=0.72, p<0.001), and support for child agency (r=0.69, p<0.001). Support for questioning children and prompting discussion also moderately correlated with support for connection to experience and ideas (r=0.72, p<0.001) and support for child agency (r=0.93, p<0.001). Between support for connection to experience and ideas, and support for child agency, there was strong and positive correlation (r=0.76, p<0.001).

Between the PLAY observation tool factors and PLAY teacher survey factors, observed support for togetherness and cooperation moderately correlated with teacher-report support for questioning children and prompting discussion (r=0.37, p<0.001) and teacher-report support for child agency (r=0.39, p<0.001).

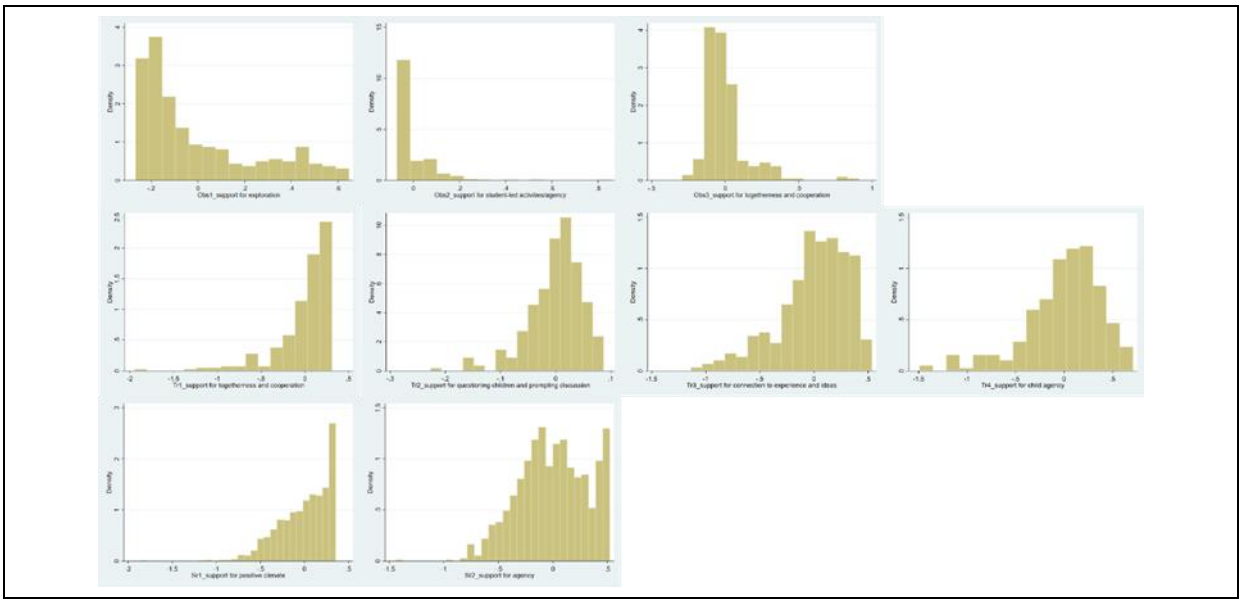
Table B-15 demonstrates the results from multilevel mixed-effects linear regressions, where the group-level variable is the classroom. The majority had no significant correlations between two factors extracted from the PLAY student survey and factors extracted from the PLAY observation tool and teacher survey.

Table B-11. Results of multilevel mixed-effects linear regressions between primary classroom-level observation tool/teacher survey and classroom-level student survey in Kenya

	HLM two variables	
	Sr1_support for positive climate	**Sr2_support for agency**
Sr2_support for agency	0.82***	1
Obs1_support for exploration	-0.049	0.035
Obs2_support for student-led activities/agency	-0.083	-0.059
Obs3_support for togetherness and cooperation	-0.18*	-0.197*
Tr1_support for togetherness and cooperation	-0.009	-0.033
Tr2_support for questioning children and prompting discussion	-0.053	-0.092
Tr3_support for connection to experience and ideas	-0.0004	0.0196
Tr4_support for child agency	0.0021	0.0043

Note: * p<0.05, *** p<0.001

Figure B-3. Distributions of factors extracted from the common model of PLAY observation tool, teacher survey, and student survey in Kenya



B.I.4.3.2 Correlations between Observation Tool/Teacher Survey and Teacher's Sociodemographic

Table B-16 demonstrates the concurrent spearman's rank correlations between three PLAY observation tool factors, four PLAY teacher survey factors, and four sociodemographic factors. Most of the correlations were insignificant, particularly those associated with years of being a teacher. Although there were a few significant correlations, they were considered negligible, at 0.3 level.

Table B-16. Correlations between primary classroom-level observation tool/ teacher survey and sociodemographic backgrounds in Kenya

	Obs1_support for exploration	Obs2_support for student-led activities/agency	Obs3_support for togetherness and cooperation	Tr1_support for togetherness and cooperation	Tr2_support for questioning children and prompting discussion	Tr3_support for connection to experience and ideas	Tr4_support for child agency
Years of being a teacher	-0.035	0.005	0.069	0.03	0.03	0.063	0.016
Gender (reference: male)	0.046	0.053	0.059	0.131*	0.071	0.091	0.038
Grade taught by teacher	-0.049	-0.043	0.014	-0.151*	-0.111	-0.187**	-0.079
Teacher's highest level of education	0.133*	0.074	0.173**	0.031	0.106	0.13*	0.159**

Note: * p<0.05, ** p<0.01

Findings from the ECE Classroom-Based Tools

B.II.1 ECE Rater Reliability

In Colombia, Jordan, and Ghana absolute agreement statistics were used to calculate rater reliability. In Colombia and Jordan, we calculated absolute agreement rates for all classrooms that were observed by multiple raters, and set a threshold of >0.70 and >.65 absolute agreement, respectively to determine which records to use in our analytic sample. In Ghana, because no classrooms were observed by multiple raters, we set a threshold of >0.70 with master coded videos in training in order to establish reliability of raters; all observers met the threshold prior to data collection.

B.II.2 ECE Internal Consistency

Internal consistency, which is based on item correlations, measures how much the items within an instrument measure the same construct or characteristic.

Cronbach's alpha was conducted to examine the internal consistency of each construct of the primary observation tool, as shown in **Table B-17**. Aside from Jordan, where the internal consistency was good to excellent, most of the constructs had Cronbach's alpha smaller than 0.7, which suggested acceptable to poor internal consistency of the originally hypothesised constructs.

Table B-17. Internal consistency of hypothesised constructs in the PLAY ECE observation tool

Construct	Cronbach's alpha		
	Colombia	Jordan	Ghana
Support for agency	0.56	0.86	0.22
Support for connection to experience	0.62	0.83	0.59
Support for exploration	0.75	0.85	0.60
Support for problem solving	0.78	0.86	0.41
Support for social connectedness	0.78	0.93	0.54
Support for positive emotional climate	0.74	0.90	0.48

The internal consistency of each construct of the ECE teacher survey is shown in **Table B-18**. Some of the constructs had Cronbach's alpha smaller than 0.6, which suggested poor internal consistency of the originally hypothesised constructs.

Table B-18. Internal consistency of hypothesised constructs in PLAY ECE teacher survey

Construct	Cronbach's alpha	
	Colombia	Jordan
Agency	0.71	0.78
Connection to experience	0.79	0.84
Exploration	0.72	0.76
Problem solving	0.85	0.82
Social connectedness	0.47	0.78
Positive emotional climate	0.61	0.90

B.II.3 Final Models for ECE Tools

B.II.3.1 *Classroom Observation Factor Analytic Overview from Three Countries*

A common model with a total of 17 items was tested across countries as a reduced confirmatory model for the ECE observation tool using a pooled data set. The results of the analyses are presented in **Tables B-19 and B-20**.

Table B-19. Model fit statistics of ECE classroom observation by country

	χ2 (df)	p-value	CFI	TLI	RMSEA	SRMR
Jordan	234.894 (111)	0.000	0.888	0.863	0.108	0.066
Ghana	226.605 (111)	0.000	0.909	0.889	0.050	0.057
Colombia	115.490 (68)	0.000	0.932	0.909	0.072	0.065

Table B-20. Standardised CFA loadings of the common CFA model for PLAY ECE observation tool in each country

		Mean (SD)			CFA factor loadings								
		Jordan (N=95)	Ghana (N=423)	Colombia (N=133)	Jordan			Ghana			Colombia		
Factor 1: Exploration and problem solving	OAG4	1.37 (1.24)	0.24 (0.68)	1.38 (1.18)	Teacher provides a range of opportunities for children to generate and share ideas and opinions	0.842*	7 items, α=0.883	Teacher provides opportunities for children to generate and share ideas and opinions	0.617*	7 items, α=0.680	Teacher provides opportunities for children to generate and share ideas and opinions	0.712*	6 items, α=0.798
	OEX6	1.37 (1.24)	0.19 (0.55)	0.94 (1.16)	Teacher expresses or show curiosity to lead children to inquiry and information gathering	0.786*		Teacher expresses or shows curiosity to lead children to inquiry and information gathering	0.539*		Teacher expresses or shows curiosity to lead children to inquiry and information gathering	0.538*	
	OCE2	1.28 (1.11)	0.90 (1.19)	1.28 (1.23)	Teacher connects concepts being taught to real-life or children's everyday experiences, showing the relevance of the main topic of the lesson outside the classroom or to 'real life'	0.671*		Teacher connects concepts being taught to real-life or children's everyday experiences, showing the relevance of the main topic of the lesson outside the classroom or to 'real life'	0.405*				
	OCE3	1.24 (1.15)	1.18 (1.23)	1.42 (1.07)	Teacher elicits children's relevant background knowledge	0.760*		Teacher elicits children's relevant background knowledge	0.382*		Teacher elicits children's relevant background knowledge	0.704*	
	OEX5	1.11 (1.25)	0.20 (0.64)	0.56 (0.92)	Teacher rephrases answers and asks questions around multiple uses of materials	0.713*		Teacher rephrases answers and asks questions around multiple uses of materials	0.572*		Teacher rephrases answers and asks questions around multiple uses of materials	0.420*	

		Mean (SD)			CFA factor loadings								
		Jordan (N=95)	Ghana (N=423)	Colombia (N=133)	Jordan			Ghana			Colombia		
	OPS1	0.47 (0.96)	0.12 (0.50)	0.51 (0.56)	Teacher engages children in hypothesis generation [predicting questions]	0.534*		Teacher engages children in hypothesis generation	0.516*		Teacher engages children in hypothesis generation	0.498*	
	OPS3	0.76 (1.11)	0.10 (0.44)	0.53 (0.89)	Teacher asks questions to generate explanations/reasons (e.g. for phenomena/ experiments/results)	0.670*		Teacher engages children towards generating explanations/providing reasons for phenomena/experiments/ results	0.458*		Teacher engages children towards generating explanations/providing reasons for phenomena/experiments/ results	0.633*	
Factor 2: Social connectedness/personal connections	**OCE5**	1.01 (1.16)	0.30 (0.83)	1.11 (1.11)	Teacher elicits student recall of personal events and experiences	0.674*	**5 items, α=0.865**	Teacher elicits student recall of personal events and experiences to learning	0.644*	**5 items, α=0.707**	Teacher elicits children's relevant background knowledge	0.639*	**5 items, α=0.809**
	OSC2	1.00 (1.16)	0.14 (0.58)	1.14 (1.24)	Teacher encourages the children to describe themselves, their families, and communities	0.738*		Teacher encourages the children to describe their personal experience, themselves, their families, and communities	0.609*		Teacher encourages the children to describe their personal experience, themselves, their families, and communities	0.872*	
	OSC3	1.02 (1.24)	0.10 (0.50)	1.35 (1.26)	Teacher promotes children's interest in one another's lives	0.899*		Teacher promotes children's interest in one another's lives	0.538*		Teacher promotes children's interest in one another's lives	0.620*	
	OSC9	1.03 (1.24)	0.11 (0.44)	1.23 (1.25)	Teacher expresses understanding and acceptance of the different personal experiences, stories, and cultures of the students in the class	0.857*		Teacher expresses understanding and acceptance of the different personal experiences, stories, and cultures of the students in the class	0.594*		Teacher expresses understanding and acceptance of different personal experiences	0.614*	
	OPEC 4	1.11 (1.25)	0.04 (0.24)	1.00 (1.09)	Teacher explains children's actions, intentions, and feelings to other children	0.679*		Teacher explains children's actions, intentions, and feelings to other children	0.433*		Teacher explains children's actions, intentions, and feelings to other children	0.457*	

		Mean (SD)			CFA factor loadings								
		Jordan (N=95)	Ghana (N=423)	Colombia (N=133)	Jordan			Ghana			Colombia		
Factor 3: Social and emotional support	OAG6	1.55 (1.23)	0.33 (0.81)	2.44 (0.93)	Teacher observes and listens to children before stepping in	0.813*	5 items, α=0.881	Teacher observes and listens to children before intervening	0.473*	5 items, α=0.545			3 items, α=0.664
	OSC5	1.54 (1.20)	0.27 (0.64)	1.65 (1.27)	Teacher encourages peer active listening	0.795*		Teacher encourages peer active listening	0.680*		Teacher encourages peer active listening	0.809*	
	OSC10	1.35 (1.25)	0.72 (0.89)	1.94 (0.95)	Teacher reinforces/promotes children's understanding of social norms	0.669*		Teacher reinforces/promotes children's understanding of social norms	0.375*		Teacher reinforces/promotes children's understanding of social norms	0.518*	
	OPEC5	1.69 (1.20)	0.28 (0.65)	2.00 (1.05)	Teacher is responsive to all children's emotional needs	0.800*		Teacher is responsive to all children's emotional needs	0.311*				
	OPEC8	1.74 (1.16)	0.38 (0.70)	2.33 (0.99)	Teacher encourages behaviours of friendship and/or social acceptance between children via sharing, cordiality, and affection	0.819*		Teacher encourages behaviours of friendship and/or social acceptance between children via sharing, cordiality, and affection	0.392*		Teacher encourages behaviours of friendship and/or social acceptance between children via sharing, cordiality, and affection	0.524*	

Note: * p<0.05

Factor 1: Support for exploration and problem-solving
Factor 2: Support for social connectedness
Factor 3: Social and emotional support

The following three items did not load in the Colombia CFA analysis:

- OCE2 – Teacher connects concepts being taught to real-life or children's everyday experiences, showing the relevance of the main topic of the lesson outside the classroom or to 'real life'
- OAG6 – Teacher observes and listens to children before intervening
- OPEC5 – Teacher is responsive to all children's emotional needs

Thus, a total of 14 items loaded across all three countries. The internal consistency of the three constructs on the final set of items using the pooled dataset are as follows:

Table B-21. Internal consistency values for pooled ECE data set (Jordan, Ghana, and Colombia)

Factor	Alpha (α)
Support for exploration and problem solving (6 items)	0.84
Support for social connectedness/personal connections (5 items)	0.87
Social and emotional support (3 items)	0.80

We then conducted multigroup measurement invariance analyses using a pooled dataset. However, results of the analysis reveal poor model fit, thus ruling out configural invariance across the three countries, as well as any possibility of testing for metric or scalar invariance.

B.II.3.2 *Teacher Survey Factor Analytic Overview from Two Countries*

For the ECE teacher survey, we were unable to reach a common model. However, the overall summary from the confirmatory models is outlined below in **Tables B-22** and **B-23**. Overall, the fit statistics are not as strong as the observation tool, with statistics that range from adequate to mediocre.

Table B-22. Model fit statistics of ECE teacher survey by country

	χ2 (df)	p-value	CFI	TLI	RMSEA	SRMR
Jordan 4-factor with correlations	728.332 (423)	0.000	0.832	0.815	0.081	0.084
Colombia 3-factor with correlations	228.019 (144)	0.000	0.904	0.886	0.069	0.074

Table B-23. Results of ECE teacher survey confirmatory factor analyses by country

Item	Description	Jordan	Colombia	Jordan	Colombia	Factor
ST_EX3	I ask students questions regarding the properties and specific characteristics of objects and provide opportunities for discovery.	3.58 (0.85)	4.49 (0.84)		0.628*	Support for agency
ST_EX1	I give students a chance to explore something before I show them how to use/answer it.	2.87 (1.14)	4.51 (0.84)	0.803*	0.640*	Support for agency
ST_AG7	I am flexible in my planning to allow room to embed my students' interests.	2.99 (1.22)	4.41 (0.99)	0.716*	0.614*	Support for agency
ST_AG5	I give students the opportunity to make suggestions about what they want to learn.	2.86 (1.16)	4.06 (1.26)	0.684*	0.602*	Support for agency
ST_AG3	I allow students to develop their own way to complete tasks.	3.20(0.95)	4.35 (0.86)	0.646*	0.446*	Support for agency
ST_AG6	I give students the chance to influence my lesson planning.	1.95(1.48)	3.79 (1.67)	0.551*	0.658*	Support for agency
ST_AG1	I give students an environment in which they have access and opportunity to choose what to engage with.	2.93 (1.09)	4.41 (0.89)	0.514*		Support for agency
ST_AG4	I give students opportunities to take on roles/responsibilities/ownership in the classroom.	3.16 (1.08)	4.40 (1.04)	0.414*		Support for agency
ST_AG2	I provide an environment in the classroom where students are free to choose when and where to move around.	2.99 (1.20)	4.18 (1.01)	0.413*		Support for agency
ST_EX4	I model curiosity to lead children to inquiry and information gathering.	3.57 (0.90)	4.56 (0.75)	0.579*		Support for connection to experience and problem solving
ST_EX5	I provide ample time for children to inspect the properties of discovered/available materials.	2.89 (1.18)	3.65 (1.15)	0.588*		Support for connection to experience and problem solving
ST_CE1	I use students' interests in developing lessons/learning objectives.	2.99 (1.00)	4.15 (1.16)	0.606*	0.631*	
ST_CE2	I connect lesson objectives to children's real-life or everyday experiences.	2.98 (1.10)	4.29 (0.94)	0.663*	0.762*	Support for connection to experience
ST_CE3	I connect what children are learning about or doing to what daily experiences and routines they are familiar with.	2.86 (1.09)	4.44 (0.81)	0.711*	0.669*	Support for connection to experience

Item	Description	Jordan	Colombia	Jordan	Colombia	Factor
ST_CE4	I ask children to recall past activities to build upon their interests and to inform the topic being taught.	3.15 (1.03)	4.07 (1.15)	0.688*	0.674*	Support for connection to experience
ST_CE5	I promote children's familiarity with culturally specific practices/objects/geographies/symbols.	3.23 (0.92)	4.30 (0.87)	0.726*	0.664*	
ST_PS1	I engage children in ways that encourage thinking and reasoning.	3.21 (0.85)	4.23 (1.07)	0.731*		
ST_PS2	I promote my students' transfer of knowledge to new applications.	3.00 (1.08)	4.25 (0.88)	0.729*	0.624*	Support for problem solving and social connectedness
ST_PS3	I ask open-ended questions to encourage children's reasoning.	3.25 (1.00)	4.58 (0.77)	0.423*	0.823*	Support for problem solving and social connectedness
ST_PS4	I engage children towards generating explanations/providing reasons for phenomena/experiments/results.	2.90 (1.07)	4.16 (1.11)	0.571*	0.748*	Support for problem solving and social connectedness
ST_PS5	I support children's efforts towards reaching a particular goal.	3.46 (0.98)	4.65 (0.73)	0.619*	0.693*	Support for problem solving and social connectedness
ST_PE1	I interact in a positive and warm manner with children.	3.87 (0.41)		0.866*	0.444*	
ST_PE2	I use positive language and words with children.	3.87 (0.47)	4.93 (0.26)	0.737*		Support for positive emotional climate
ST_PE3	I facilitate the expression of emotions between children.	3.86 (0.44)	4.88 (0.43)	0.755*		Support for positive emotional climate
ST_PE4	I am able to understand my students' emotional needs and respond appropriately.	3.83 (0.50)	4.76 (0.56)	0.758*	0.436*	
ST_PE5	I use positive praise and reinforcement.	3.94 (0.34)	4.77 (0.56)	0.831*		Support for positive emotional climate
ST_PE6	I encourage prosocial classroom behaviour between children via sharing, cordiality, and affection.	3.92 (0.38)	4.94 (0.25)	0.756*		Support for positive emotional climate
ST_PE7	I effectively manage the tensions and difficulties presented in the classroom.	3.74 (0.60)	4.69 (0.65)	0.617*		Support for positive emotional climate
ST_SC4	I encourage peer-listening among children in my classroom.	3.64 (0.78)	4.92 (0.35)	0.887*		Support for social connectedness
ST_SC3	I create opportunities for children in my classroom to interact and relate to one another.	3.66 (0.80)	4.92 (0.30)	0.885*		Support for social connectedness
ST_SC5	I encourage children to describe themselves, their families, and their communities.	3.31 (0.83)	4.27 (0.98)	0.620*		Support for social connectedness

Item	Description	Jordan	Colombia	Jordan	Colombia	Factor
ST_SC2	I educate children on how to approach others with warmth during greetings and interactions.	3.55 (0.93)	4.80 (0.52)	0.543*	0.654*	

Note: * p<0.05; Factors are blocked by color

- Factor: Support for agency
- Factor: Support for connection to experience & problem solving
- Factor: Support for connection to experience
- Factor: Support for problem solving & social connectedness
- Factor: Support for positive emotional climate
- Factor: Support for social connectedness

B.II.4 ECE Construct Validity

To reiterate, construct validity tells us how well a tool has operationalised hypothesised constructs into measurable characteristics that reflect the true meaning of a concept. In other words, construct validity is how well a tool, assessment, etc., effectively measures what it claims. We can generally verify construct validity by comparing a measure to other measures of similar constructs and comparing the correlations between the two (or more).

First, we will review the within country factor analytic models for each tool, which may vary slightly from the pooled data set results presented above. Then we will provide the concurrent validity results for those tools by country. For validity analyses, factor scores were computed in MPlus and exported to Stata. Factor scores provide more precise estimations of the factor values than item averages, as they weight each item based on its factor loading.

B.II.4.1 *Colombia ECE*

B.II.4.1.1 Factor Structures

Classroom Observation

Table B-24 shows the model fit statistics for the final 3-factor confirmatory model in Colombian ECE settings with item-level correlations, confirming adequate to good model fit.

Table B-24. Model fit statistics for final 3-factor confirmatory model in Colombia ECE with item-level correlations

	3-factor
Sample-size adjusted BIC	4897.480
Chi-square (df)	119.251 (68)
P-value	0.000
RMSEA	0.076
CFI	0.926
TLI	0.901
SRMR	0.068

Table B-25. Factor loadings from final 3-factor confirmatory model for Colombia ECE observation

		Standardised factor loading
Factor 1: Support for exploration and problem solving α=0.80		
OAG4	Teacher provides opportunities for children to generate and share ideas and opinions	0.71
OCE3	Teacher elicits children's relevant background knowledge	0.71
OEX5	Teacher probes answers around multiple uses/applications of materials	0.42
OEX6	Teacher models curiosity to capture children’s attention	0.53
OPS1	Teacher engages children in hypothesis generation	0.49
OPS3	Teacher engages children towards generating explanations/providing reasons for phenomena/experiments/results	0.62

		Standardised factor loading
Factor 2: Support for personal connections/social connectedness α=0.81		
OCE5	Teacher elicits student recall of personal events and experiences to learning	0.70
OSC2	Teacher encourages the children to describe themselves, their families, and communities	0.64
OSC3	Teacher promotes children's interest in one another's lives	0.72
OSC9	Teacher validates the different personal experiences, stories, and cultures of the students in the class	0.79
OPEC4	Teacher explains children's actions, intentions, and feelings to other children	0.54
Factor 3: Support for social and emotional support α=0.67		
OSC5	Teacher encourages peer active listening	0.81
OSC10	Teacher reinforces/promotes children's understanding and following of social norms	0.52
OPEC8	Teacher encourages prosocial classroom behaviour between children via sharing, cordiality, and affection	0.52

Descriptive statistics for each of the three factors show that the factors are positively and strongly correlated to each other (f1 and f2: r=0.84, p<0.001; f1 and f3: r=0.99, p<0.001; f2 and f3: r=0.79, p<0.001), and that all factors are somewhat skewed (skewness: f1=-0.03, f2=0.28, f3=-0.18).

Teacher Survey

Table B-26 shows the model fit statistics for the Colombia ECE teacher survey final confirmatory model with item-level correlations, confirming that there is an adequate to mediocre model fit.

Table B-26. Model fit statistics for Colombia ECE teacher survey final confirmatory model with item-level correlations

	3-factor
Sample-size adjusted BIC	5298.874
Chi-square (df)	228.019 (144)
P-value	0.000
RMSEA	0.069
CFI	0.904
TLI	0.886
SRMR	0.074

Table B-27. Factor loadings for final 3-factor confirmatory model Colombia ECE teacher survey

		Standardised factor loading
Factor 1: Support for agency α=0.82		
ST_AG3	I allow students to develop their own way to complete tasks.	0.45
ST_AG5	I give students the opportunity to make suggestions about what they want to learn.	0.60
ST_AG6	I give students the chance to influence my lesson planning.	0.66

		Standardised factor loading
ST_AG7	I am flexible in my planning to allow room to embed my students' interests.	0.61
ST_EX1	I give students a chance to explore something before I show them how to use/answer it.	0.64
ST_EX3	I ask students questions regarding the properties and specific characteristics of objects and provide opportunities for discovery.	0.63
ST_CE1	I use students' interests in developing lessons/learning objectives.	0.63
ST_CE5	I promote children's familiarity with culturally specific practices, objects, geographies, or symbols.	0.66
ST_PE1	I interact in a positive and warm manner with children.	0.44
Factor 2: Support for connection to experience α=0.76		
ST_EX5	I provide ample time for children to inspect the properties of discovered/available materials.	0.55
ST_CE2	I connect lesson objectives to children's experiences at home and in their communities.	0.76
ST_CE3	I connect what children are learning about or doing to what daily experiences and routines they are familiar with.	0.67
ST_CE4	I ask children to recall past activities to build upon their interests and to inform the topic being taught.	0.67
Factor 3: Support for problem solving and social connectedness α=0.81		
ST_PS2	I promote my students' transfer of knowledge to new lessons.	0.62
ST_PS3	I ask open-ended questions to encourage children's reasoning.	0.82
ST_PS4	I engage children towards generating explanations/providing reasons for phenomena/experiments/results.	0.75
ST_PS5	I support children's efforts towards reaching a particular goal.	0.69
ST_SC2	I educate children on how to approach others with warmth during greetings and interactions.	0.65
ST_PE4	I am able to understand my students' emotional needs and respond appropriately.	0.44

B.II.4.1.2 Correlations between Observation Tool, Teacher Survey, and Teacher Demographics

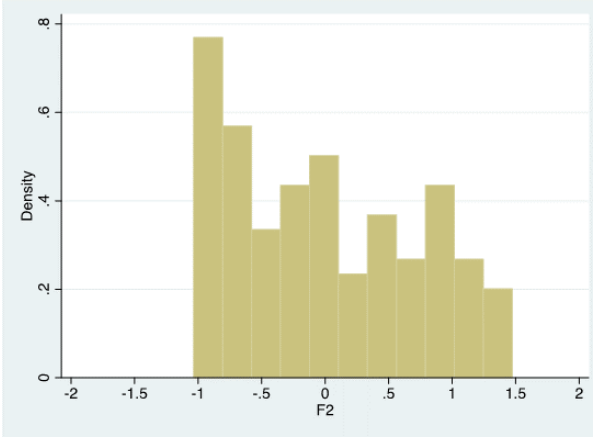
To examine concurrent and discriminatory validity of the tools, pairwise correlations were utilised, including demographic data from the broader PLAY study. Pairwise correlations with PLAY observation and teacher survey measures indicated no significant correlations between each of the three factors from the teacher survey and the three factors from the classroom observation tool (see **Table B-28**).

Table B-28. Concurrent validity with survey measure for Colombia ECE

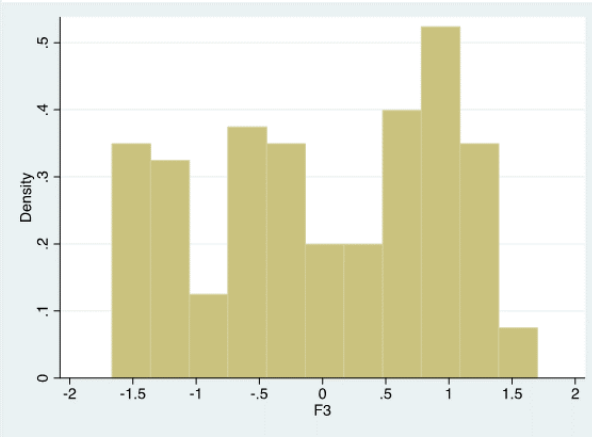
	Support for		
	Agency (survey)	**Connection to experience (survey)**	**Problem solving and social connectedness (survey)**
	r		
Support for exploration and problem solving (observation)	0.005	0.04	-0.002
	121	121	121
Support for connection to experience (observation)	-0.01	0.02	-0.03
	121	121	121
Support for social connectedness (observation)	-0.02	-0.06	-0.04
	121	121	121

Note: † p<0.1

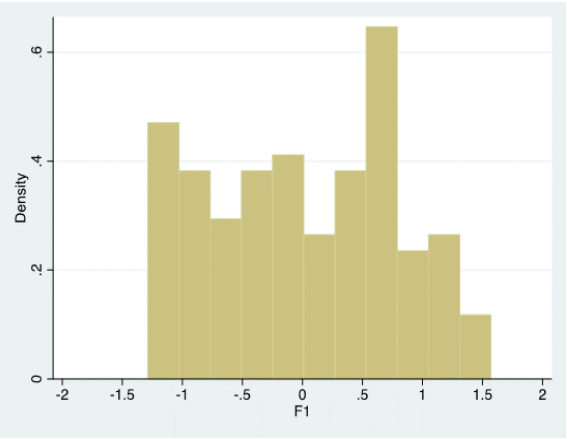
Figure B-4. Distributions of factors from Colombia ECE observation, teacher survey, and teacher demographic factors



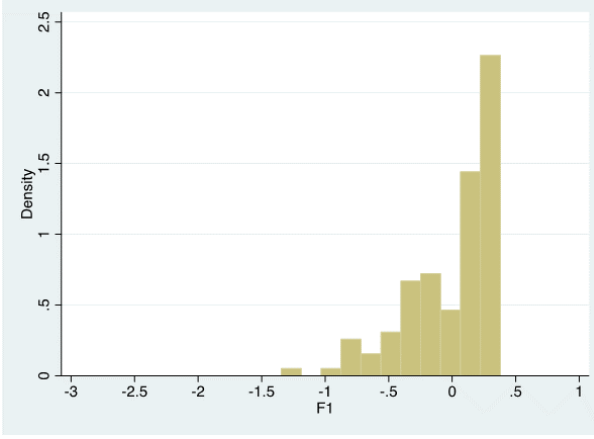
OBS Factor 1: Support for exploration and problem solving



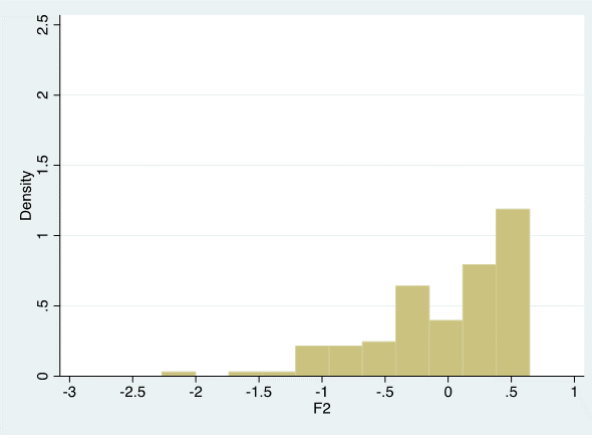
OBS Factor 2: Support for personal connections/social connectedness



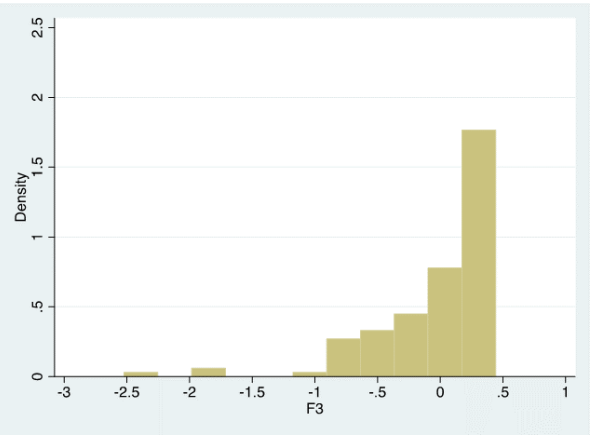
OBS Factor 3: Support for social and emotional support



Teacher Factor 1: Support for Agency

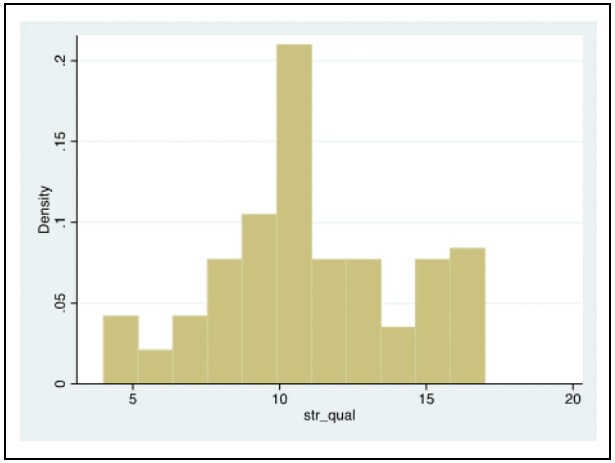


Teacher Factor 2: Support for connection to experience

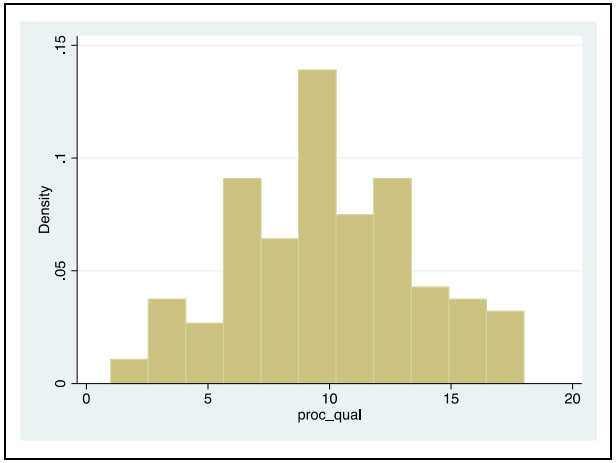


Teacher Factor 3: Support for problem solving social connectedness

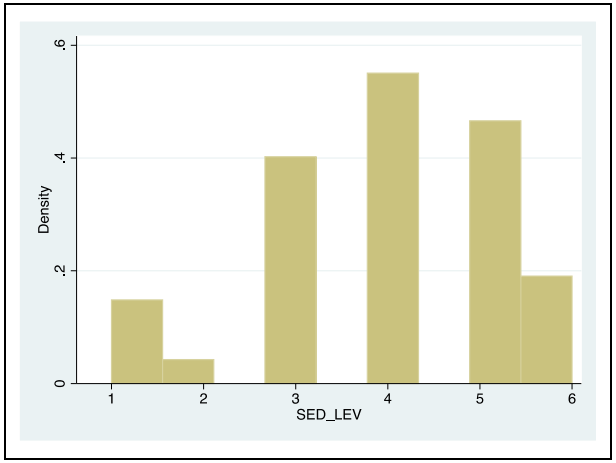
Structural quality measure in PLAY ECE classroom observation tool Colombia



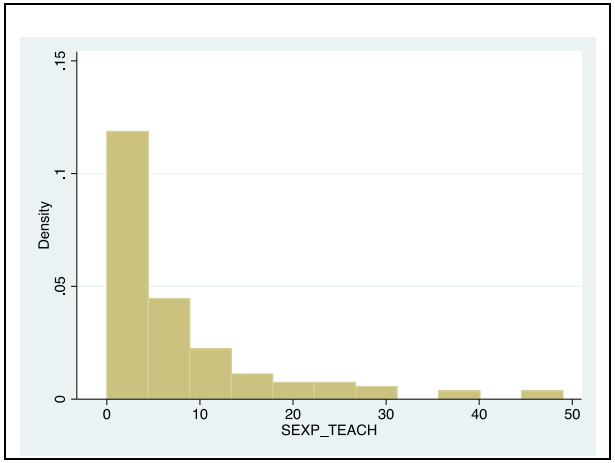
Process quality measure in PLAY ECE classroom observation tool Colombia



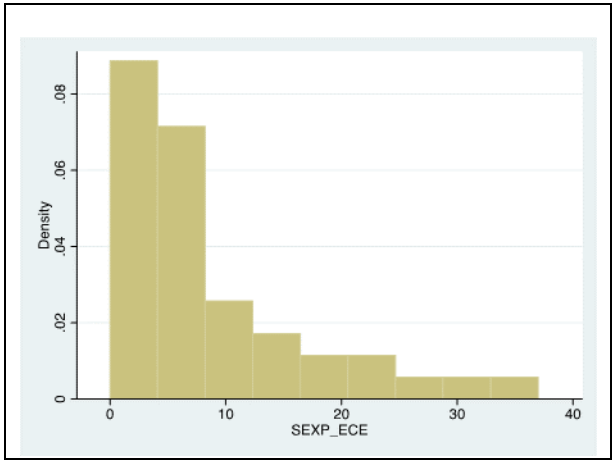
Teacher level of education for Colombia ECE



Teacher years of experience in teaching role for Colombia ECE



Teacher years of experience in ECE teaching role for Colombia ECE



The following classroom and teacher characteristics were used in concurrent validity analyses with the classroom observation and teacher survey factors of the PLAY tool.

Structural quality. Structural quality was observed by the data collectors and included a 19-item checklist of various classroom resources (e.g. Are there fantasy play materials? Dolls, stuffed animals, dress up clothes, masks, pretend food, pots and spoons). A count of yes responses to these items was used for the structural quality scale.

Process quality. Process quality was observed by the data collectors and included a 20-item checklist of children's use of various classroom resources (e.g. Did you observe children using art materials?). A count of yes responses to these items was used for the process quality scale.

Teacher level of education. Teachers reported their highest level of formal education with the following response options: baccalaureate or less, normal higher school, labour/technical, professional/technical, bachelor of arts/science, specialist or master's degree, and doctorate.

Years of experience in teaching role. Teachers reported the number of years they had served in the following teaching roles: teacher in ECE/pre-primary, teacher in primary, teacher in middle, teacher in secondary, and teacher in tertiary.

Years of experience in ECE teaching role. Teachers reported the number of years they had served in an ECE/pre-primary teaching role.

Has training in ECE. Teachers reported whether any of their formal education training focused on ECE (No=0; Yes=1).

Results indicated no significant correlations between each of the three factors of the observation tool and three factors of the teacher survey with any of the teacher characteristics (see **Table B-29**).

Table B-29. Concurrent validity of PLAY ECE factors with teacher characteristics for Colombia

	OBS: Exploration and problem solving	OBS: Personal connections/ social connectedness	OBS: Social and emotional support	Teacher: Agency	Teacher: Connection to experience	Teacher: Problem solving and social connectedness
	r					
Structural quality	0.02	0.01	0.001	0.06	-0.02	0.006
	131	131	131	121	121	121
Process quality	0.03	-0.07	0.02	-0.02	-0.12	-0.15†
	131	131	131	121	121	121
Teacher level of education	-0.02	-0.03	-0.02	0.07	-0.15	-0.18
	87	87	87	85	85	85
Years of experience in teaching role	-0.01	0.003	-0.02	0.03	0.03	0.06
	131	131	131	124	124	124
Years of experience in ECE teaching role	-0.04	-0.02	-0.04	0.007	0.08	0.09
	87	87	87	85	85	85
Has training in ECE	-0.10	-0.13	-0.11	0.04	0.007	0.007
	87	87	87	85	85	85

B.II.4.2 Jordan ECE

B.II.4.2.1 Factor Structures

Classroom Observation

Table B-30 shows the model fit statistics for the Jordan ECE classroom observation tool confirmatory model with item-level correlations, confirming that there is an adequate to mediocre model fit.

Table B-30. Model fit statistics for Jordan ECE observation 3-factor confirmatory model with and without item-level correlations

	χ2 (df)	p-value	CFI	TLI	RMSEA	SRMR
With no inter-item correlations	294.695 (116)	0.000	0.839	0.811	0.127	0.070
With inter-item correlations	234.894 (111)	0.000	0.888	0.863	0.108	0.066

Table B-31. Final 3-factor confirmatory model for Jordan ECE observation

		Standardised factor loading	
Factor 1	**Exploration and problem solving**		**7 items, α=0.883**
OAG4	Teacher provides a range of opportunities for children to generate and share ideas and opinions	0.842	
OCE2	Teacher connects concepts being taught to real-life or children's everyday experiences, showing the relevance of the main topic of the lesson outside the classroom or to 'real life'	0.671	
OCE3	Teacher elicits children's relevant background knowledge	0.760	
OEX5	Teacher rephrases answers and asks questions around multiple uses of materials	0.713	
OEX6	Teacher expresses or show curiosity to lead children to inquiry and information gathering	0.786	
OPS1	Teacher engages children in hypothesis generation [predicting questions]	0.534	
OPS3	Teacher asks questions to generate explanations/reasons (e.g. for phenomena/experiments/results)	0.670	
Factor 2	**Personal connections/social connectedness**		**5 items, α=0.865**
OCE5	Teacher elicits student recall of personal events and experiences	0.674	
OSC2	Teacher encourages the children to describe themselves, their families and communities	0.738	
OSC3	Teacher promotes children's interest in one another's lives	0.899	
OSC9	Teacher expresses understanding and acceptance of the different personal experiences, stories, and cultures of the students in the class	0.857	
PEC4	Teacher explains children's actions, intentions, and feelings to other children	0.679	

		Standardised factor loading	
Factor 3	**Social and emotional support**		**5 items, α=0.881**
OAG6	Teacher observes and listens to children before stepping in	0.813	
OSC5	Teacher encourages peer active listening	0.795	
OSC10	Teacher reinforces/promotes children's understanding of social norms	0.669	
OPEC5	Teacher is responsive to all children's emotional needs	0.800	
OPEC8	Teacher encourages behaviours of friendship and/or social acceptance between children via sharing, cordiality, and affection	0.819	

Descriptive statistics for each of the three factors show that the factors are positively and moderately correlated with each other (f1 and f2: r=0.80, p<0.001; f1 and f3: r=0.62, p<0.001; f2 and f3: r=0.74, p<0.001), and that all factors are positively skewed.

Teacher Survey

Table B-32 shows the model fit statistics for the Jordan ECE teacher survey final four factor confirmatory model with item-level correlations, confirming that there is an adequate to mediocre model fit.

Table B-32. Model fit statistics for final confirmatory model for Jordan ECE teacher survey

	χ2 (df)	p-value	CFI	TLI	RMSEA	SRMR
With correlations	728.332 (423)	0.000	0.832	0.815	0.081	0.084
Without	811.979 (428)	0.000	0.789	0.771	0.090	0.086

Table B-33. Factor loadings from final Jordan ECE teacher survey 4-factor confirmatory model

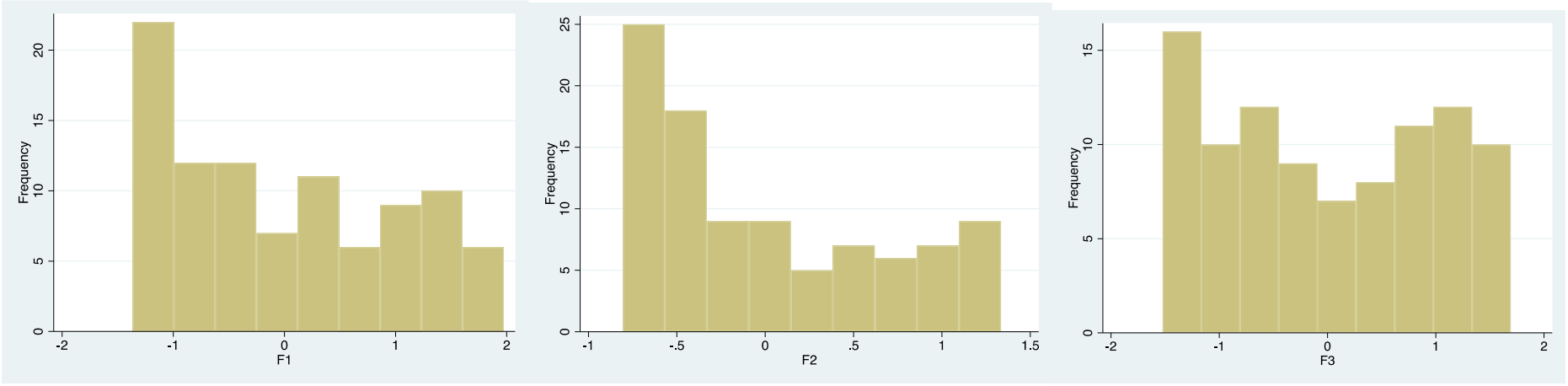
		Jordan CFA final 4-factor model: Factor loadings (31 items)		
Factor 1: Support for agency	ST_EX1	I give students a chance to explore something before I show them how to use/answer it.	0.80	**8 items, α=0.819**
	ST_AG7	I am flexible in my planning to allow room to embed my students' interests.	0.72	
	ST_AG5	I give students the opportunity to make suggestions about what they want to learn.	0.68	
	ST_AG3	I allow students to develop their own way to complete tasks.	0.65	
	ST_AG6	I give students the chance to influence my lesson planning.	0.55	
	ST_AG1	I give students an environment in which they have access and opportunity to choose what to engage with.	0.51	
	ST_AG4	I give students opportunities to take on roles/responsibilities/ownership in the classroom.	0.41	
	ST_AG2	I provide an environment in the classroom where students are free to choose when and where to move around.	0.41	
F2: Support for connection to experience	ST_EX4	I model curiosity to lead children to inquiry and information gathering.	0.58	**12 items, α=0.893**
	ST_EX5	I provide ample time for children to inspect the properties of discovered/available materials.	0.59	
	ST_CE1	I use students' interests in developing lessons/learning objectives.	0.61	

		Jordan CFA final 4-factor model: Factor loadings (31 items)		
	ST_CE2	I connect lesson objectives to children's real-life or everyday experiences.	0.66	
	ST_CE3	I connect what children are learning about or doing to what daily experiences and routines they are familiar with.	0.71	
	ST_CE4	I ask children to recall past activities to build upon their interests and to inform the topic being taught.	0.69	
	ST_CE5	I promote children's familiarity with culturally specific practices/objects/geographies/symbols.	0.73	
	ST_PS1	I engage children in ways that encourage thinking and reasoning.	0.73	
	ST_PS2	I promote my students' transfer of knowledge to new applications.	0.73	
	ST_PS3	I ask open-ended questions to encourage children's reasoning.	0.42	
	ST_PS4	I engage children towards generating explanations/providing reasons for phenomena/experiments/results.	0.57	
	ST_PS5	I support children's efforts towards reaching a particular goal.	0.62	
F3: Positive emotional climate	ST_PE1	I interact in a positive and warm manner with children.	0.87	**7 items, α=0.900**
	ST_PE2	I use positive language and words with children.	0.74	
	ST_PE3	I facilitate the expression of emotions between children.	0.76	
	ST_PE4	I am able to understand my students' emotional needs and respond appropriately.	0.76	
	ST_PE5	I use positive praise and reinforcement.	0.83	
	ST_PE6	I encourage prosocial classroom behaviour between children via sharing, cordiality, and affection.	0.76	
	ST_PE7	I effectively manage the tensions and difficulties presented in the classroom.	0.62	
F4: Social connectedness	ST_SC4	I encourage peer-listening among children in my classroom.	0.89	**4 items, α=0.814**
	ST_SC3	I create opportunities for children in my classroom to interact and relate to one another.	0.89	
	ST_SC5	I encourage children to describe themselves, their families, and their communities.	0.62	
	ST_SC2	I educate children on how to approach others with warmth during greetings and interactions.	0.54	

Factor Descriptive Statistics

Descriptive statistics for each of the three factors show that the factors are positively and moderately correlated with each other (f1 and f2: r=0.80, p<0.001; f1 and f3: r=0.62, p<0.001; f2 and f3: r=0.74, p<0.001; f3 and f4: r=0.538, p<0.001), and that all factors are negatively skewed.

Figure B-5. Distributions of factors from Jordan ECE observation, teacher survey, and teacher demographic factors



OBS Factor 1: Support for exploration and problem solving

OBS Factor 2: Support for personal connections/social connectedness

OBS Factor 3: Support for social and emotional support

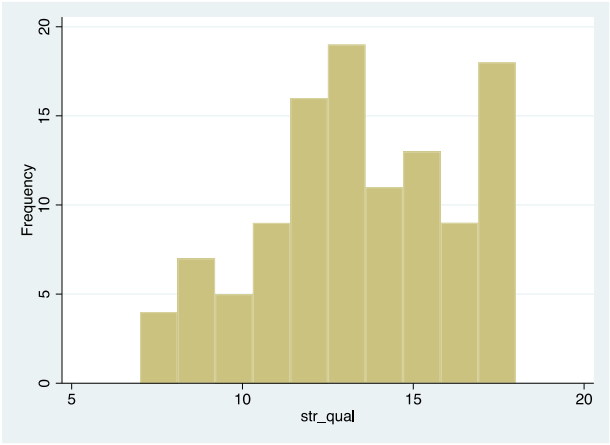
Teacher Factor 1: Support for agency

Teacher Factor 2: Support for agency connection to experience & problem solving

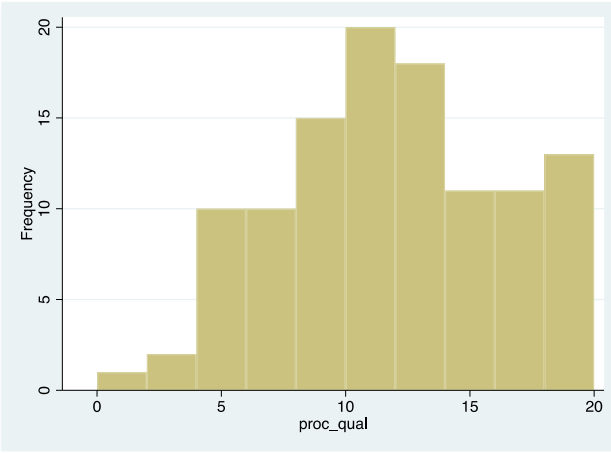
Teacher Factor 3: Positive emotional climate

Teacher Factor 4: Social connectedness

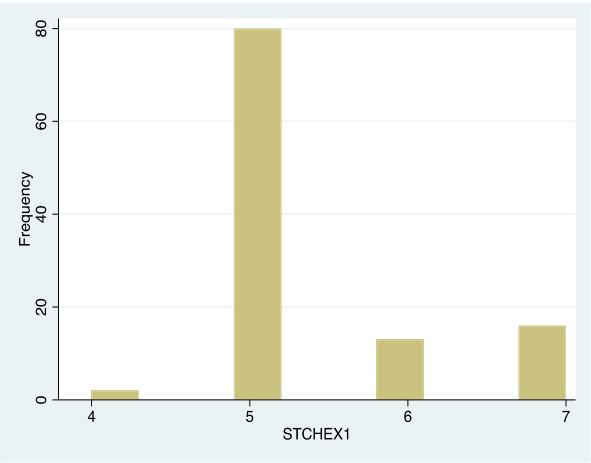
Structural quality measure in PLAY ECE classroom observation tool Jordan



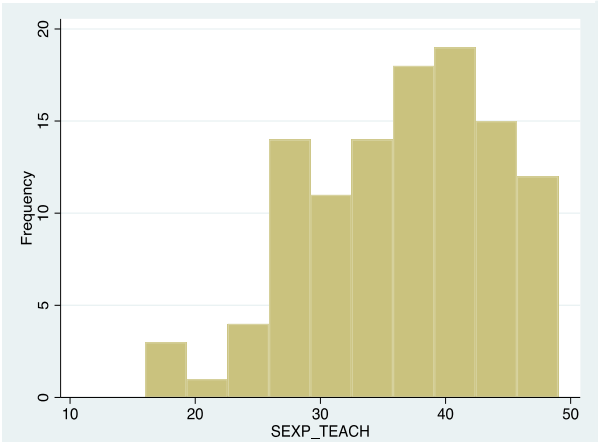
Process quality measure in PLAY ECE classroom observation tool Jordan



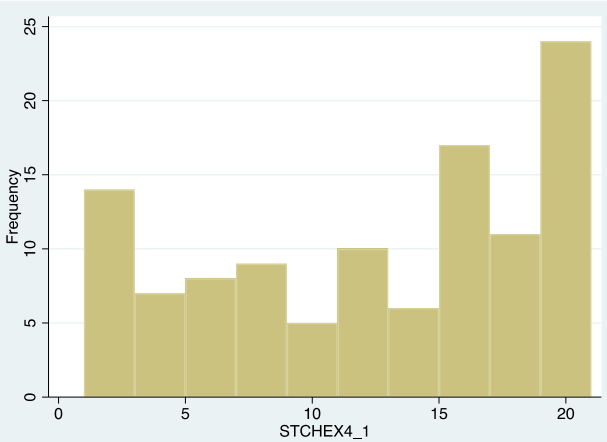
Teacher level of education for Jordan ECE



Teacher years of experience in teaching role for Jordan ECE



Teacher years of experience in ECE teaching role for Jordan ECE



B.II.4.2.2 Correlations between Observation Tool, Teacher Survey, and Teacher Demographics

To examine concurrent and discriminatory validity of the tools, pairwise correlations were utilised, including demographic data from the broader PLAY study. Pairwise correlations with PLAY observation and teacher survey measures indicated no significant correlations between factors 1, 2, and 3 from the teacher survey and the three factors from the classroom observation. However, teacher survey factor 4 (social connectedness) did exhibit trend-level associations with observation factor 1 (exploration and problem solving) and small correlations with factors 2 and 3, personal connections and social and emotional support, of the observation tool (see **Table B-34**).

Table B-34. Correlations between classroom observation with teacher survey for Jordan ECE

		Teacher survey			
		Support for agency	**Connection to experience and problem solving**	**Positive emotional climate**	**Social connectedness**
Observational measure	**Exploration and problem solving (observation)**	0.08	-0.03	-0.02	0.17†
		95	95	95	95
	Personal connections/social connectedness (observation)	0.11	0.05	0.001	0.24*
		95	95	95	95
	Social and emotional support (observation)	0.15	0.02	0.04	0.24*
		95	95	95	95

Note: † p<0.10, * p<0.05

The most surprising results are perhaps the strong associations between structural quality and the three PLAY observational factors *(See section B.I.4.1.2 for description of classroom characteristics)*. Research shows that structural qualities are meaningful towards learning outcomes to the degree they are able to promote process quality (Phillips, Mekos, Scarr, McCartney, & Abbott-Shim, 2000; Pianta et al., 2005; Vandell, 2004; NICHD Study of Early Child Care, 2002), but there is little evidence that they influence child development overall (Pianta, Downer, & Hamre, 2016). However, most observational measures we know of are general measures of classroom quality and don't focus specifically on elements related to play. This could potentially account for some of the differences we are seeing.

Results indicate that both structural and process quality have statistically significant positive correlations to the three observational factors. Furthermore, specific training in early childhood education exhibits small positive, but statistically significant, associations with exploration and problem solving as well as social and emotional climate. Other trend-level associations are exhibited between teacher level of education and the two of the observational factors (see **Table B-35**). Finally, there appear to be no significant correlations between the four factors from the teacher survey and any of the teacher characteristics.

Table B-35. Correlations between of PLAY ECE observation and teacher survey factors with teacher characteristics in Jordan

	OBS: Personal connections/social connectedness	OBS: Exploration and problem solving	OBS: Social and emotional climate	Teacher: Support for agency	Teacher: Connection to experience & problem solving	Teacher: Positive emotional climate	Teacher: Social connectedness
				r			
Structural quality	0.41***	0.31**	0.43**	0.10	0.03	0.15	0.06
	95	95	95	111	111	111	111
Process quality	0.51***	0.53***	0.55***	0.14	-0.005	0.04	0.14
	95	95	95	111	111	111	111
Teacher level of education	0.19†	0.17†	0.13	0.03	-0.01	0.02	0.01
	95	95	95	111	111	111	111
Years of experience in teaching role	0.0002	0.006	0.007	-0.02	-0.04	-0.10	-0.01
	95	95	95	111	111	111	111
Years of experience in ECE teaching role	0.02	-0.02	0.04	0.03	-0.12	-0.079	0.03
	95	95	95	111	111	111	111
Has training in ECE	0.22*	0.19†	0.26**	0.07	0.08	0.009	0.004
	95	95	95	111	111	111	111

Note: † p<0.10, * p<0.05, ** p<0.01, *** p<0.001

B.II.4.3 Ghana ECE

B.II.4.3.1 Factor Structure

Observation Tool

Table B-36 shows the model fit statistics for 3-factor confirmatory model with and without item-level correlations. Without inter-item correlations, the model fit is poor. However, the addition of the inter-item correlations significantly improved the model fit to be quite strong.

Table B-36. Model fit statistics for 3-factor confirmatory model with and without item-level correlations (Ghana ECE observation)

	χ^2 (df)	p-value	CFI	TLI	RMSEA	SRMR
With no inter-item correlations	372.483 (116)	0.000	0.799	0.764	0.072	0.067
With inter-item correlations	226.605 (111)	0.000	0.909	0.889	0.050	0.057

Table B-37 and **Figure B-6** show the final 3-factor confirmatory model for the Ghana ECE observation tool.

Table B-37. Final 3-factor confirmatory model for Ghana ECE observation tool

	Description	Standardised factor loading
Factor 1		
OAG4	Teacher provides opportunities for children to generate and share ideas and opinions	0.62
OCE2	Teacher connects concepts being taught to real-life or children's everyday experiences, showing the relevance of the main topic of lesson outside the classroom or to 'real life'	0.41
OCE3	Teacher elicits children's relevant background knowledge	0.38
OEX5	Teacher rephrases answers and asks questions around multiple uses of materials	0.57
OEX6	Teacher expresses or shows curiosity to lead children to inquiry and information gathering	0.54
OPS1	Teacher engages children in hypothesis generation	0.52
OPS3	Teacher engages children towards generating explanations/providing reasons for phenomena/experiments/results	0.46
Factor 2		
OCE5	Teacher elicits student recall of personal events and experiences to learning	0.64
OSC2	Teacher encourages the children to describe their personal experience, themselves, their families, and communities	0.61
OSC3	Teacher promotes children's interest in one another's lives	0.54
OSC9	Teacher expresses understanding and acceptance of different personal experiences, stories, and cultures of the students in the class	0.59
OPEC4	Teacher explains children's actions, intentions, and feelings to other children	0.43

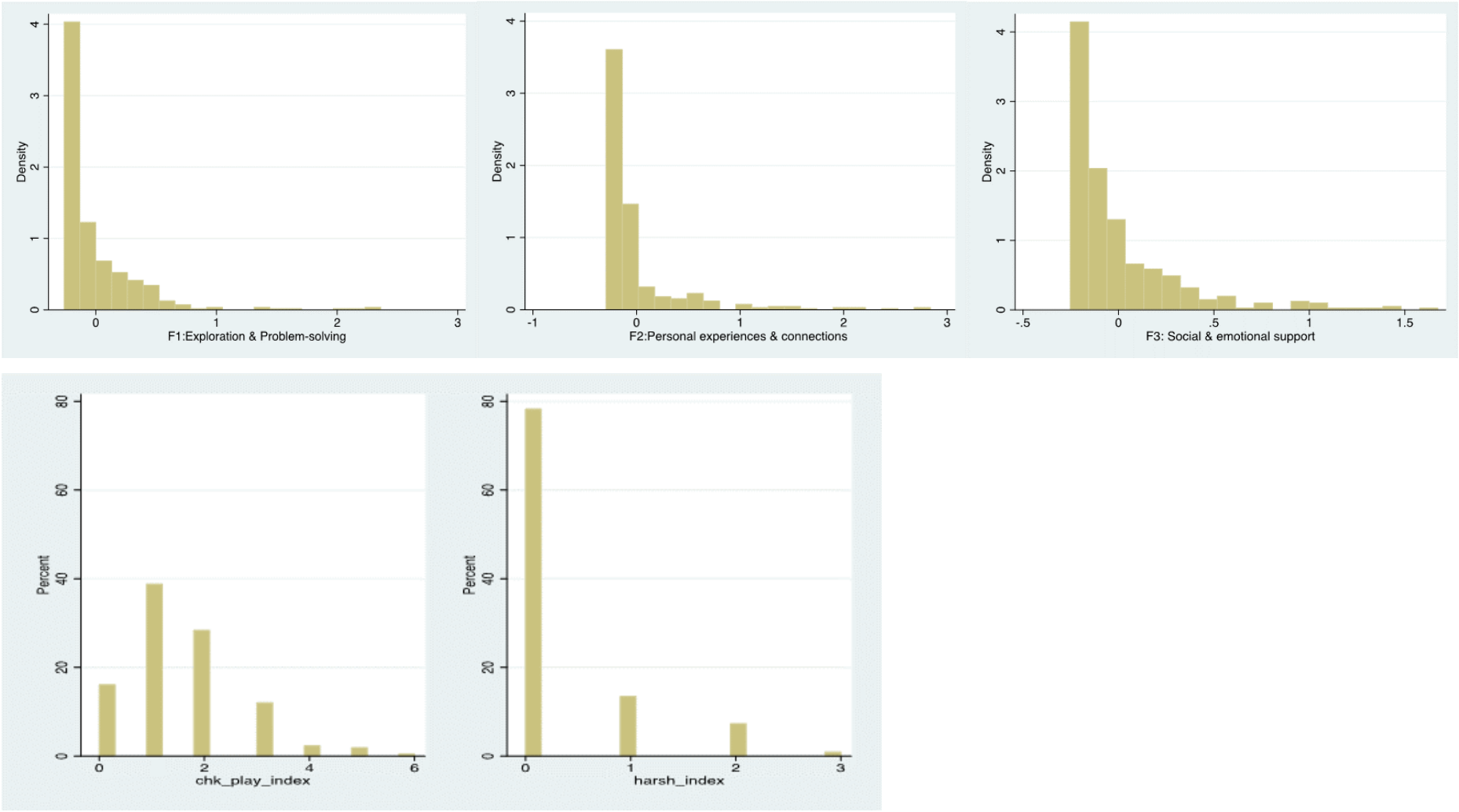
	Description	Standardised factor loading
Factor 3		
OAG6	Teacher observes and listens to children before intervening	0.47
OSC5	Teacher encourages peer active listening	0.68
OSC10	Teacher reinforces/promotes children's understanding and following of social norms	0.38
OPEC5	Teacher is responsive to all children's emotional needs	0.31
OPEC8	Teacher encourages behaviours of friendship and/or social acceptance between children via sharing, cordiality, and affection	0.39

Factor Descriptive Statistics

Descriptive statistics for each of the three factors show that the factors are positively and moderately correlated to each other and that all factors are very negatively skewed (skewness stats: f1=3.4, f2=3.4, f3=4.3). This is not surprising based on the large proportion of zero scores for the items.

Figure B-6 shows a histogram of the three factors for Ghana ECE observation and checklist indices on playful teaching practices and harsh disciplinary practices.

Figure B-6. Distributions of factors from Ghana ECE observation and checklist indices



B.II.4.3.2 Concurrent Validity Analysis

To examine concurrent and discriminatory validity of the three factors, pairwise correlations were utilised, including combining data from the Quality Preschool for Ghana (QP4G) project.

Two features of classroom quality are often considered. *Process quality* is considered the driver of child development and learning (Howes, James, & Ritchie, 2003; Pianta, 2005) and refers to the nature of children's daily interactions and experiences in the classroom, with a broad focus on the social, emotional, physical, and instructional aspects of activities and interactions. *Structural quality* refers to regulable resources, such as class size, student-teacher ratio, and teacher training and education, that are considered important for improving learning outcomes to the extent that they promote process quality. (Notably, though, studies that move beyond associations and aim to examine these associations more rigorously find small or null relations between process quality and child outcomes; Burchinal, 2018).

Step 1: Pairwise correlations with play practices and harsh discipline checklists

The PLAY toolkit included two checklists to cover teaching practices related to play (eight items) and harsh discipline (three items). An index was created for each (histograms shown below). The two indices were correlated, with the correlation coefficient small and negative (r=-0.113).

Results indicate small correlations between two of the three factors and the two checklists (see **Table B-38**). Specifically, exploration and problem solving (factor 1) was not correlated with either checklist. Personal connections had a small and positive correlation with the Playful Teaching Practices checklist. Finally, social and emotional support was correlated to both checklists, with a small positive correlation with playful teaching practices and a small negative correlation with harsh disciplinary practices.

Table B-38. Factor correlations with two PLAY checklists for Ghana ECE observation

	Exploration and problem solving	Personal connections	Social and emotional support
		r	
Checklist: playful teaching practices	0.0576	0.1014*	0.2230*
	423	423	423
Checklist: harsh disciplinary practices	0.0081	-0.0515	-0.1162*
	423	423	423

	Exploration and problem solving	**Personal connections**	**Social and emotional support**
		r	
Checklist: playful teaching practices	0.0576 423	0.1014* 423	0.2230* 423
Checklist harsh disciplinary practices	0.0081 423	-0.0515 423	-0.1162* 423

Note: * p<0.05

Step 2a: Pairwise correlations with QP4G process quality indicators (TIPPS and checklists)

Data were merged with the QP4G classroom observation dataset to explore how the three factors correlated with three factors as coded by the Teacher Instructional Practices and Processes System (TIPPS) (Wolf et al., 2018), as well as a checklist of classroom practices related to positive behaviour management, child-friendly classroom practices, and the use of instructional materials.

There is some evidence of concurrent validity using these measures (see **Table B-39**), though the direction of the associations appears to be representing more general quality of the classroom, given that factors do not always align perfectly with similar constructs (e.g. emotional support).

Exploration and problem solving had small, positive, and significant correlations with all three TIPPS factors, including facilitating deeper learning (a measure of scaffolding and instructional support), supporting student expression (a measure related to helping students develop reasoning and problem-solving skills), and emotional support and behaviour management. Further, this factor had small, positive, and significant correlations with a checklist related to positive behaviour management and child-friendly classroom practices. The factor was not correlated with a checklist measuring the use of instructional materials in teaching.

The personal connections factor had small, significant correlations with supporting student expression and emotional support and behaviour management, but not with facilitating deeper learning. Finally, the factor was positive behaviour management, child-friendly classrooms, and a negative correlation (but very small; r=-0.03, p<0.05) with instructional materials.

Lastly, social and emotional support had small, positive correlations with all TIPPS factors and checklist indices except for instructional materials.

Table B-39. Concurrent validity of PLAY factors with QP4G classroom observations of process quality indicators for Ghana ECE observation

	Exploration and problem solving	Personal connections	Social and emotional support
		r	
TIPPS: Facilitating deeper learning	0.1127*	0.0322	0.0369*
	416	416	416
TIPPS: Supporting student expression	0.0930*	0.1056*	0.1371*
	416	416	
TIPPS: Emotional support & behavior management	0.3439*	0.2898*	0.1312*
	416	416	416
Checklist: positive behavior management	0.0929*	0.0967*	0.0986*
	419	419	419
Checklist: child-friendly classroom	0.2472*	0.2358*	0.1807*
	419	419	419
Checklist: Instructional materials	-0.0195	-0.0349*	0.0223
	419	419	419

	Exploration and problem solving	Personal connections	Social and emotional support
	r		
TIPPS: facilitating deeper learning	0.1127* 416	0.0322 416	0.0369* 416
TIPPS: supporting student expression	0.0930* 416	0.1056* 416	0.1371*
TIPPS: emotional support & behaviour management	0.3439* 416	0.2898* 416	0.1312* 416
Checklist: positive behaviour management	0.0929* 419	0.0967* 419	0.0986* 419
Checklist: child-friendly classroom	0.2472* 419	0.2358* 419	0.1807* 419
Checklist: instructional materials	-0.0195 419	-0.0349* 419	0.0223 419

Note: * p<0.05

Step 2b: Pairwise correlations with structural quality indicators

Correlations between each of the three factors and a set of structural quality indicators were examined. These indicators were collected from the PLAY data set (i.e. number of children in the classroom) and from the QP4G data collectors when they were present in the classroom (a broader set of structural resources). These results are displayed in **Table B-40**.

There are few correlations between the number of children in the class and the three factors, with one exception: the number of boys in the classroom had a small, negative association with social and emotional support scores.

A classroom resource index was created based on a set of ten indicators in the classroom. There was a small, positive association with classroom resources and each of the three factors (r=0.05–0.06), indicating that for each additional classroom resource, teachers scored slightly higher on each PLAY factor.

When broken down to examine each resource individually, it actually appeared that in some cases, having *fewer* resources was correlated with *higher* PLAY scores. Specifically, classroom building structural quality, having student desks, and having an organised seating structure in rows were negatively correlated with personal connections and social and emotional support. Similarly, having books available in the classroom was negatively associated with personal connections. This suggests that with fewer resources, teachers resort to being more creative in how they use other materials and connections to teach and support their students.

Resources that were associated with better PLAY scores included having writing materials, no outside disruptive noise, and having charts in the classroom.

Table B-40. Correlation of PLAY factors with classroom observations of structural quality indicators for Ghana ECE observation

	Exploration and problem-solving	Personal connections	Social and emotional support
No. children in class (total)	0.0024 412	-0.0087 412	-0.0153 412
No. boys	0.0162 411	-0.0277 411	-0.0412* 411
No. girls	-0.0189 412	-0.009 412	-0.0015 412
Classroom resource index (0-10)	0.0648* 416	0.0497* 416	0.0537* 416
resource1: No. of teachers	0.0041 414	-0.0141 414	0.0081 414
resource2: class building structure	0.0159 414	-0.0565* 414	-0.0420* 414
resource3: organized seating	0.0317 414	-0.0497* 414	-0.1407* 414
resource4: small class size (<25)	-0.0019 414	0.0051 414	-0.0237 414
resource5: student desks	-0.1116* 414	-0.0426* 414	-0.0412* 414
resource6: books	-0.0224 414	-0.0446* 414	-0.0227 414
resource7: writing materials	0.0617* 379	0.0601* 379	0.0389* 379
resource8: no outside noise	0.0888* 412	0.0615* 412	0.0642* 412
resource9: charts in class	0.0427* 414	0.0767* 414	0.0799* 414
resource10: adequate lighting	-0.0298 414	0.0024 414	0.0039 414

	Exploration and problem solving	Personal connections	Social and emotional support
No. children in class (total)	0.0024 412	-0.0087 412	-0.0153 412
No. boys	0.0162 411	-0.0277 411	-0.0412* 411
No. girls	-0.0189 412	-0.009 412	-0.0015 412
Classroom resource index (0-10)	0.0648* 416	0.0497* 416	0.0537* 416
Resource 1: no. of teachers	0.0041 414	-0.0141 414	0.0081 414
Resource 2: class building structure	0.0159 414	-0.0565* 414	-0.0420* 414
Resource 3: organised seating	0.0317 414	-0.0497* 414	-0.1407* 414
Resource 4: small class size (<25)	-0.0019 414	0.0051 414	-0.0237 414
Resource 5: student desks	-0.1116* 414	-0.0426* 414	-0.0412* 414
Resource 6: books	-0.0224 414	-0.0446* 414	-0.0227 414
Resource 7: writing materials	0.0617* 379	0.0601* 379	0.0389* 379
Resource 8: no outside noise	0.0888* 412	0.0615* 412	0.0642* 412
Resource 9: charts in class	0.0427* 414	0.0767* 414	0.0799* 414
Resource 10: adequate lighting	-0.0298 414	0.0024 414	0.0039 414

Note: * p<0.05

Resources that were not correlated to PLAY scores included the number of teachers (though the vast majority of classrooms had only one teacher, limiting variation), a small class size (fewer than 25 students), and having adequate lighting in the classroom.

Step 3: Pairwise correlations with teacher characteristics

Exploratory analyses were conducted with a set of teacher characteristics, private- versus public-sector status, and professional well-being indicators (see **Table B-41**).

Table B-41. Correlation of PLAY factors with teacher characteristics for Ghana ECE observation

	Exploration and problem solving	Personal connections	Social and emotional support
		r	
Education level (continuous)	0.0643*	0.0863*	0.1195*
	331	331	331
Private school (versus public)	-0.0137	0.0195	-0.03
	416	416	416
No training in ECD	-0.0715*	-0.0706*	-0.0614*
	410	410	410
Burnout	0.0677*	0.0413*	0.0118
	410	410	410
Personal accomplishment	0.0843*	0.1001*	0.1032*
	410	410	410
Motivation	-0.0105	0.0056	0.0924*
	410	410	410
Job satisfaction	-0.0241	-0.0119	-0.0095
	410	410	410
# of years as an ECE teacher	-0.0626*	-0.0444*	-0.0650*
	409	409	409

	Exploration and problem solving	Personal connections	Social and emotional support
	r		
Education level (continuous)	0.0643* 331	0.0868* 331	0.1195* 331
Private school (versus public)	-0.0137 416	0.0195 416	-0.03 416
No training in ECD	-0.0715* 410	-0.0706* 410	-0.0614* 410
Burnout	0.0677* 410	0.0413* 410	0.0118 410
Personnel accomplishment	0.0843* 410	0.1001* 410	0.1032* 410
Motivation	-0.0105 410	0.0056 410	0.0924* 410
Job satisfaction	-0.0241 410	-0.0119 410	-0.0095 410
# of years as an ECE teacher	-0.0626* 409	-0.0444* 409	-0.0650* 409

Note: * $p<0.05$

The broader literature indicates largely null associations between teachers having at least a bachelor's degree and high-quality education (Early et al., 2007), concluding that policies focused solely on increasing teacher education are not sufficient to improve quality and learning. In Ghana, public-sector kindergarten teachers are required to have a minimum of a diploma in basic education obtained from an approved college of education (Asare & Nti, 2014), while there are no minimum requirements for private sector teachers. The certificate in basic education requires two years of coursework, followed by one year of student teaching. Thus, the level of required credentials is less than in high-income countries. It is possible that this training is critical to preparing teachers. But it is equally possible that it is not sufficient to strongly affect teaching quality.

Regarding teacher professional well-being, there is a growing literature on the stress of teaching and the role that burnout can have in terms of leading to teachers creating harmful learning environments and deteriorating teacher performance (Osheret al., 2007; Tsouloupas et al., 2010), as well as poorer classroom climate (La Paro et al., 2009). There is growing concern that teachers in lower-income countries, including Ghana, are unmotivated and that this may partially explain poor teaching performance and student learning outcomes, high rates of turnover and absenteeism, and misconduct (Moon, 2007; Bennell & Akyeampong, 2007). Given the challenges documented in the teaching profession in Ghana (Bennell & Akyeampong, 2007; Osei, 2006), we explored teachers' psychological and professional well-being as a potential predictor of playful learning practices.

The results show that teacher education is positively related to all three factors, suggesting that more educated teachers are more likely to implement all three of the practices covered in these factors. Similarly, having some training in early childhood development (84% of the sample) was positively correlated with all three factors. Interestingly, the number of years worked as a preschool teacher was negatively correlated with all three factors. It is possible that older teachers are more set in their ways and less likely to implement progressive teaching practices that might be introduced in more current teacher trainings. There were no differences in the scores of the three factors for public versus private school teachers.

Finally, associations with teacher professional well-being were mixed. On the one hand, burnout (emotional exhaustion and depersonalisation as measured by the Maslach Burnout Inventory; Maslach et al., 1996) was related to higher factor scores on exploration and problem solving as well as personal connections, but not social and emotional support. Motivation was related positively correlated to social and emotional support, and a sense of personal accomplishment (as measured by the Maslach Burnout Inventory) was positively correlated with all three factors. Teachers' self-reported job satisfaction was not correlated with any of the three factors.

Notably, all correlations were quite small (most <r=0.10).

B.II.4.3.3 Ghana – Impacts on Intervention Treatment Conditions

The Ghana data were unique in that they were derived by observing video recordings of classrooms from a previously concluded evaluation. The QP4G project was a large-scale, school-randomised controlled trial designed to improve preschool quality and child development outcomes. The intervention consisted of two treatment arms: a one-year teacher training and coaching programme with parental-awareness meetings, and the same training and coaching programme without parental-awareness meetings (Wolf, Aber, Behrman & Peele, 2019). The context and intervention are described in further detail below. This allowed us to analyse PLAY data in relation to existing datasets as follows:

- Impacts of QP4G interventions on PLAY factors
- Mean differences by public and private schools
- Associations of PLAY factors with child outcomes
- Variation by treatment status and kindergarten (KG) level

Impacts of QP4G Interventions on PLAY Factors

In the first step, we analysed the impact of two randomised interventions on the PLAY scores. Before presenting findings, we describe the context and nature of the intervention.

In 2007, the government of Ghana expanded two years of pre-primary (kindergarten) education to be included in the country's universal education. The quality of pre-primary education in Ghana is low, particularly in urban and semi-urban settings, where low-cost private and public preschools account for over 90% of pre-primary options. The majority of kindergarten teachers are untrained, and many have only a primary school education. At the same time, research suggests that parents of kindergarteners may be largely uninformed about what high-quality kindergarten teaching looks like, and they may pressure teachers to use methods that are developmentally inappropriate (for example, rote repetition of letters and numbers).

This research took place in six districts of the Greater Accra Region, across public and private schools serving children enrolled in kindergarten, aged four to six.

Researchers evaluated the impact of a short, in-service kindergarten teacher training programme delivered by the National Nursery Teacher Training Centre, with and without a parental-awareness programme, on teacher well-being, classroom quality, child learning, and parental knowledge of and attitudes towards early childhood education. Researchers randomly assigned 240 schools to either receive the teacher training, teacher training with parental awareness component, or neither (comparison group). Each group comprised ~80 schools (~40 public and ~40 private).

In-service teacher training: The National Nursery Teacher Training Centre training for kindergarten teachers and head teachers began with a five-day course, followed by a two-day refresher training four months later, and a one-day refresher four months after that. The programme offered experiential training for teachers and included ongoing monitoring and support. The training focused on helping teachers provide age-appropriate and play-based instructional techniques, as well as an encouraging, positive classroom environment.

Parental-awareness meetings: This programme consisted of three educational sessions (one per term) held during PTA meetings. District education coordinators screened videos, followed by discussion, which focused on (1) play-based learning, (2) parents' role in child learning, and (3) encouraging parent-teacher and parent-school communication.

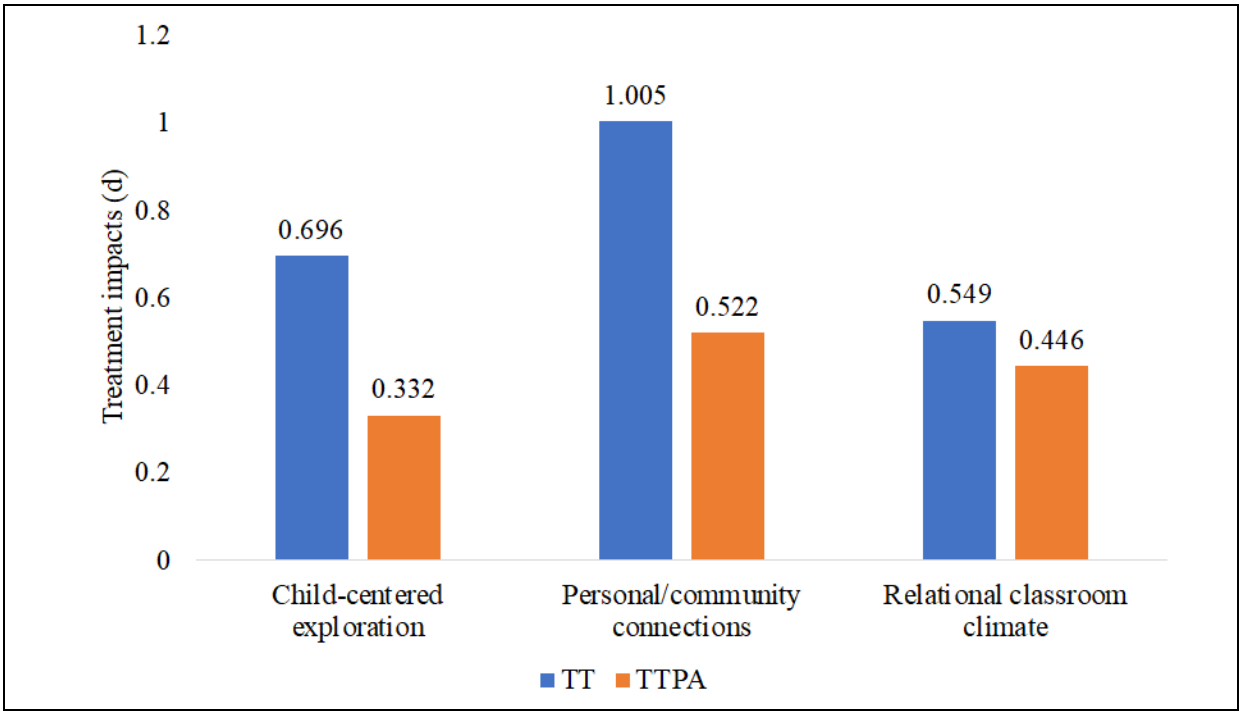
At each of the 240 schools, researchers randomly chose 15 kindergarten students from KG1 and KG2 classrooms and measured their school readiness skills, including early academic skills, social-emotional development, and behavioural outcomes. Researchers followed the students as they entered their second year of kindergarten and primary school, and one year later when children were in their first and second years of primary, in order to understand the lasting impacts of the programme of children, as well as the teachers to understand sustained impacts on teaching quality.

In the findings below, TT refers to classrooms that received the teacher training intervention alone. TTPA refers to classrooms that received both the teacher training and parental-awareness interventions. Impacts are assessed using two-level hierarchical linear models (classrooms nested in schools). Coefficients represent effect sizes, with treatment means computed relative to the control group. **Table B-42** shows the impact estimates of QP4G treatments on PLAY observation factors for Ghana ECE.

Table B-42. Impacts estimates of QP4G treatments on PLAY observation factors for Ghana ECE

	Unconditional			Conditional (district, private)		
	d	SE	p-value	d	SE	p-value
Child-centred exploration						
TT	0.69	0.23	0.002	0.70	0.23	0.002
TTPA	0.32	0.23	0.16	0.33	0.23	0.16
Personal/community connections						
TT	0.99	0.26	0.00	1.005	0.26	0.00
TTPA	0.50	0.26	0.05	0.52	0.26	0.05
Relational classroom climate						
TT	0.49	0.17	0.003	0.55	0.16	0.001
TTPA	0.37	0.17	0.03	0.45	0.17	0.008

Figure B-10. Impacts estimates of QP4G treatments on PLAY observation factors for Ghana ECE



Note: Treatment effects control for district and private school status and are derived from multilevel models with classrooms nested in schools.

Mean Differences by Public and Private Schools

There were no statistically significant differences in mean PLAY factor scores for public and private schools.

Table B-43. Mean factor scores by public and private school status for Ghana ECE observation

	Public (N=193)	Private (N=224)
Child-centred exploration	0.007	-0.004
Personal/community connections	-0.01	0.006
Relational classroom climate	0.006	-0.01

Note: No differences are statistically significant.

B.II.4.3.4 Associations of PLAY Factors with Child Outcomes

In the following section, all child outcomes are standardised; coefficients can be interpreted as effect sizes. Directly assess child outcomes are first presented as a school readiness composite alongside the data collector-reported approaches to learning, and then broken down by the five domains.

Key results:

- In general, there are small, statistically significant correlations between PLAY factors and child outcomes (r=0.01–0.08).
- In cross-sectional, unconditional models:
  - When examining school readiness as a composite variable, PLAY factors do not statistically significantly predict school readiness. Child-centred exploration does marginally statistically significantly predict approach to learning as reported by data collector.
  - When broken down by domain scores, there are some associations between PLAY factors and child outcomes. Personal/community connections positively predicts executive functions (d=0.18, p<0.05), and marginally statistically significantly predicts social-emotional skills (d=0.15, p<0.10). Surprisingly, relational classroom climate negatively predicts executive function skills (d=-0.26, p<0.05).
- When controlling for TIPPS factors, in cross-sectional models:
  - Personal/community connections significantly predicts the school readiness composite (d=0.17, p<0.05) and child-centred exploration predicts approaches to learning (d=0.19, p<0.05).
  - When examining the five domains separately, personal/community connections most consistently predicts child outcomes, including literacy (d=0.22, p<0.10), numeracy (d=0.18, p<0.10), social-emotional (d=0.18, p<0.05), and executive function (d=0.22, p<0.05). Relational classroom climate negatively predicts executive function (d=-0.29, p<0.05).
- In longitudinal lagged models (controlling for baseline/fall score):
  - Personal/community connections marginally statistically significantly predicts school readiness (d=0.07, p<0.10), and relational classroom climate negatively and statistically significantly predicts school readiness (d=-0.14, p<0.05).
  - When examining the five domains separately, there are quite a few associations between PLAY factors and child outcomes. Specifically, personal/community connections predict literacy (d=0.12, p<0.05) and social-emotional skills (d=0.15, p<0.05). Relational classroom climate negatively predicts three domains: literacy (d=-0.15, p<0.10), numeracy (d=-0.12, p<0.05), and executive function (d=-0.22, p<0.05).
  - Controlling for these TIPPS factors, all associations held and were fairly consistent.

Table B-44. Cross-sectional bivariate correlations with standardised child outcomes for Ghana ECE observation

	Child-centred exploration	Personal/community connections	Relational classroom climate
School readiness composite	0.07*	0.07*	0.004
Literacy	0.08*	0.08*	0.01
Numeracy	0.07*	0.07*	0.01
Social-emotional	0.05*	0.07*	0.02*
Executive function	0.05*	0.04*	-0.02*
Motor skills	0.03*	0.01	-0.01
Approaches to learning	0.03*	0.04*	0.06*

Note: N=3,035 children
* p<0.05

Table B-45. Cross-sectional three-level hierarchical models with classroom quality factors predicting school readiness composite and approaches to learning for Ghana ECE observation

	School readiness composite	Approaches to learning	School readiness composite	Approaches to learning	School readiness composite	Approaches to learning
PLAY: Child-centred exploration	0.12	0.14†			0.13	0.19*
	(0.08)	(0.09)			(0.085)	(0.09)
PLAY: Personal/community connections	0.14	-0.08			0.17*	-0.05
	(0.09)	(0.09)			(0.09)	(0.09)
PLAY: Relational classroom climate	-0.13	0.16			-0.17	0.14
	(0.12)	(0.13)			(0.12)	(0.13)
TIPPS: Facilitating deeper learning			0.10*	0.02	0.11*	0.02
			(0.05)	(0.05)	(0.05)	(0.05)
TIPPS: Supporting student expression			0.21**	0.06	0.21**	0.06
			(0.07)	(0.08)	(0.07)	(0.08)
TIPPS: Emotional support/behaviour management			-0.005	-0.07	-0.06	-0.10†
			(0.05)	(0.05)	(0.05)	(0.05)
TIPPS: Checklist, dev appropriate activities			-0.04	0.002	-0.03	-0.0005
			(0.02)	(0.02)	(0.02)	(0.02)
Constant	0.02	0.03	-0.75**	-0.10	-0.70**	-0.02
	(0.03)	(0.04)	(0.21)	(0.23)	(0.21)	(0.23)
Observations	3,006	3,006	3,006	3,006	3,006	3,006
Number of groups	213	213	213	213	213	213

Note: Standard errors in parentheses
** p<0.01, * p<0.05, †p<0.1
Models do not include any controls; sample size=213 schools, 390 classrooms, 3,006 children

Table B-46. Three-level hierarchical models with classroom quality factors predicting school readiness composite and approach to learning with lagged outcomes for Ghana ECE observation

	School readiness composite	Approaches to learning	School readiness composite	Approaches to learning	School readiness composite	Approaches to learning
PLAY: Child-centred exploration	-0.004	0.07			-0.01	0.11
	(0.04)	(0.07)			(0.04)	(0.07)
PLAY: Personal/community connections	0.07	-0.09			0.07†	-0.08
	(0.04)	(0.07)			(0.04)	(0.07)
PLAY: Relational classroom climate	-0.12*	0.18†			-0.14*	0.17
	(0.06)	(0.10)			(0.06)	(0.10)
TIPPS: Facilitating deeper learning			0.05*	-0.02	0.05*	-0.02
			(0.02)	(0.04)	(0.02)	(0.04)
TIPPS: Supporting student expression			0.09*	0.04	0.09*	0.04
			(0.04)	(0.06)	(0.04)	(0.06)
TIPPS: Emotional support/behaviour management			-0.004	-0.06	-0.008	-0.07†
			(0.02)	(0.04)	(0.02)	(0.04)
TIPPS: Checklist, dev appropriate activities			0.003	0.02	0.004	0.02
			(0.008)	(0.02)	(0.008)	(0.02)
Lagged outcome (FU)	0.72**	0.36**	0.72**	0.36**	0.72**	0.36**
	(0.01)	(0.02)	(0.014)	(0.02)	(0.01)	(0.02)
Constant	0.01	0.03	-0.37**	-0.04	-0.39**	0.007
	(0.02)	(0.03)	(0.11)	(0.19)	(0.11)	(0.19)
Observations	2,647	2,646	2,647	2,646	2,647	2,646
Number of groups	212	212	212	212	212	212

Note: Standard errors in parentheses
† p<0.1, * p<0.05, ** p<0.01
3-level models; does not include any controls; sample size=212 schools, 388 classrooms, 2,647 children

Table B-47. Cross-sectional three-level hierarchical models with classroom quality factors predicting child outcomes across five domains for Ghana ECE observation

	Literacy	Numeracy	Social-emotional	Executive function	Motor	Literacy	Numeracy	Social-emotional	Executive function	Motor	Literacy	Numeracy	Social-emotional	Executive function	Motor
PLAY: Child-centered exploration	0.177	0.152	0.0882	0.112	0.0971						0.177	0.147	0.0841	0.143	0.106
	(0.111)	(0.104)	(0.0864)	(0.0884)	(0.0868)						(0.113)	(0.107)	(0.0885)	(0.0896)	(0.0891)
PLAY: Personal/community connections	0.176	0.154	0.154+	0.179*	0.0114						0.215+	0.183+	0.183*	0.217*	0.0395
	(0.115)	(0.107)	(0.0890)	(0.0911)	(0.0895)						(0.113)	(0.107)	(0.0888)	(0.0900)	(0.0895)
PLAY: Relational classroom climate	-0.153	-0.134	-0.0860	-0.260*	-0.0508						-0.191	-0.174	-0.134	-0.289*	-0.0744
	(0.159)	(0.148)	(0.123)	(0.126)	(0.123)						(0.156)	(0.146)	(0.122)	(0.123)	(0.123)
TIPPS: Facilitating deeper learning						0.141*	0.0975+	0.0973*	0.0686	0.0905+	0.160**	0.115*	0.113*	0.0864+	0.0960*
						(0.0607)	(0.0573)	(0.0476)	(0.0485)	(0.0476)	(0.0605)	(0.0572)	(0.0474)	(0.0481)	(0.0478)
TIPPS: Supporting student expression						0.280**	0.252**	0.118	0.239**	0.145+	0.275**	0.246**	0.113	0.238**	0.144+
						(0.0979)	(0.0918)	(0.0765)	(0.0777)	(0.0762)	(0.0970)	(0.0912)	(0.0759)	(0.0767)	(0.0762)
TIPPS: Emotional support / behavior mgmt						0.0276	0.0170	-0.0241	-0.0238	-0.0215	-0.0395	-0.0391	-0.0693	-0.0808	-0.0465
						(0.0618)	(0.0579)	(0.0483)	(0.0491)	(0.0481)	(0.0650)	(0.0611)	(0.0509)	(0.0514)	(0.0511)
TIPPS: Checklist, dev appropriate activities						-0.0454*	-0.0239	0.0135	-0.0379*	-0.0241	-0.0486*	-0.0264	0.0113	-0.0385*	-0.0247
						(0.0219)	(0.0203)	(0.0171)	(0.0172)	(0.0168)	(0.0218)	(0.0203)	(0.0170)	(0.0171)	(0.0169)
Constant	0.0235	0.0241	0.0125	0.0117	0.0233	-1.023**	-0.905**	-0.589**	-0.679**	-0.496*	-0.927**	-0.822**	-0.523*	-0.622**	-0.461*
	(0.0378)	(0.0336)	(0.0287)	(0.0290)	(0.0279)	(0.286)	(0.267)	(0.223)	(0.226)	(0.221)	(0.287)	(0.269)	(0.224)	(0.226)	(0.224)
Observations	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006
Number of groups	213	213	213	213	213	213	213	213	213	213	213	213	213	213	213

Standard errors in parentheses

** p<0.01, * p<0.05, + p<0.1

Note - models do not include any controls; Sample size = 213 schools, 390 classrooms, 3,006 children.

	Literacy	Numeracy	Social-emotional	Executive function	Motor	Literacy	Numeracy	Social-emotional	Executive function	Motor	Literacy	Numeracy	Social-emotional	Executive function	Motor
PLAY: Child-centred exploration	0.177 (0.111)	0.152 (0.104)	0.0882 (0.0864)	0.112 (0.0884)	0.0971 (0.0868)						0.177 (0.113)	0.147 (0.107)	0.0841 (0.0885)	0.143 (0.0896)	0.106 (0.0891)
PLAY: Personal/ community connections	0.176 (0.115)	0.154 (0.107)	0.154+ (0.0890)	0.179* (0.0911)	0.0114 (0.0895)						0.215+ (0.113)	0.183+ (0.107)	0.183* (0.0888)	0.217* (0.0900)	0.0395 (0.0895)
PLAY: Relational classroom climate	-0.153 (0.159)	-0.134 (0.148)	-0.0860 (0.123)	-0.260* (0.126)	-0.0508 (0.123)						-0.191 (0.156)	-0.174 (0.146)	-0.134 (0.122)	-0.289* (0.123)	-0.0744 (0.123)
TIPPS: Facilitating deeper learning						0.141* (0.0607)	0.0975+ (0.0573)	0.0973* (0.0476)	0.0686 (0.0485)	0.0905+ (0.0476)	0.160** (0.0605)	0.115* (0.0572)	0.113* (0.0474)	0.0864+ (0.0481)	0.0960* (0.0478)

	Literacy	Numeracy	Social-emotional	Executive function	Motor	Literacy	Numeracy	Social-emotional	Executive function	Motor	Literacy	Numeracy	Social-emotional	Executive function	Motor
TIPPS: Supporting student expression						0.280** (0.0979)	0.252** (0.0918)	0.118 (0.0765)	0.239** (0.0777)	0.145+ (0.0762)	0.275** (0.0970)	0.246** (0.0912)	0.113 (0.0759)	0.238** (0.0767)	0.144+ (0.0762)
TIPPS: Emotional support/behaviour management						0.0276 (0.0618)	0.0170 (0.0579)	-0.0241 (0.0483)	-0.0238 (0.0491)	-0.0215 (0.0481)	-0.0395 (0.0650)	-0.0391 (0.0611)	-0.0693 (0.0509)	-0.0808 (0.0514)	-0.0465 (0.0511)
TIPPS: Checklist, develop appropriate activities						-0.0454* (0.0219)	-0.0239 (0.0203)	0.0135 (0.0171)	-0.0379* (0.0172)	-0.0241 (0.0168)	-0.0486* (0.0218)	-0.0264 (0.0203)	0.0113 (0.0170)	-0.0385* (0.0171)	-0.0247 (0.0169)
Constant	0.0235 (0.0378)	0.0241 (0.0336)	0.0125 (0.0287)	0.0117 (0.0290)	0.0233 (0.0279)	-1.023** (0.286)	-0.905** (0.267)	-0.589** (0.223)	-0.679** (0.226)	-0.496* (0.221)	-0.927** (0.287)	-0.822** (0.269)	-0.523* (0.224)	-0.622** (0.226)	-0.461* (0.224)
Observations	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006	3,006
Number of groups	213	213	213	213	213	213	213	213	213	213	213	213	213	213	213

Standard errors in parentheses
** p<0.01, * p<0.05, + p<0.1
Note – models do not include any controls; Sample size = 213 schools, 390 classrooms, 3,006 children.

Table B-48. Three-level hierarchical models with classroom quality factors predicting child outcomes across five domains with lagged outcomes for Ghana ECE observation

	Literacy	Numeracy	Social-emotional	Executive function	Motor	Literacy	Numeracy	Social-emotional	Executive function	Motor	Literacy	Numeracy	Social-emotional	Executive function	Motor
PLAY: Child-centered exploration	0.0311	0.0204	0.0326	0.0711	0.00288						0.0259	0.00433	0.0278	0.0833	-0.00849
	(0.0603)	(0.0501)	(0.0681)	(0.0732)	(0.0691)						(0.0615)	(0.0515)	(0.0698)	(0.0754)	(0.0715)
PLAY: Personal/community connections	0.124*	0.0750	0.148*	0.0988	-0.0186						0.140*	0.0765	0.170*	0.117	-0.0114
	(0.0627)	(0.0529)	(0.0716)	(0.0767)	(0.0727)						(0.0620)	(0.0529)	(0.0713)	(0.0768)	(0.0732)
PLAY: Relational classroom climate	-0.153+	-0.123+	-0.0882	-0.215*	-0.0333						-0.185*	-0.141+	-0.125	-0.241*	-0.0448
	(0.0890)	(0.0730)	(0.100)	(0.107)	(0.101)						(0.0878)	(0.0727)	(0.0997)	(0.107)	(0.101)
TIPPS: Facilitating deeper learning						0.0838*	0.0138	0.0867*	0.0471	0.0504	0.0919**	0.0179	0.0990**	0.0549	0.0483
						(0.0329)	(0.0279)	(0.0378)	(0.0407)	(0.0386)	(0.0329)	(0.0280)	(0.0377)	(0.0407)	(0.0388)
TIPPS: Supporting student expression						0.115*	0.126**	0.0675	0.153*	0.0947	0.119*	0.129**	0.0648	0.160*	0.0982
						(0.0542)	(0.0452)	(0.0618)	(0.0664)	(0.0624)	(0.0540)	(0.0451)	(0.0614)	(0.0660)	(0.0625)
TIPPS: Emotional support / behavior management						-0.00301	0.0214	-0.0167	-0.0111	-0.00170	-0.0264	0.0121	-0.0491	-0.0406	0.00394
						(0.0341)	(0.0285)	(0.0390)	(0.0419)	(0.0394)	(0.0362)	(0.0303)	(0.0412)	(0.0443)	(0.0420)
TIPPS: Checklist, dev appropriate activities						-0.00285	-0.00429	0.0126	-0.0156	-0.0135	-0.00272	-0.00302	0.0112	-0.0142	-0.0122
						(0.0126)	(0.0102)	(0.0142)	(0.0151)	(0.0141)	(0.0126)	(0.0102)	(0.0141)	(0.0151)	(0.0142)
Lagged outcome (FU)	0.628**	0.700**	0.434**	0.405**	0.405**	0.627**	0.699**	0.434**	0.403**	0.402**	0.626**	0.698**	0.433**	0.402**	0.403**
	(0.0149)	(0.0140)	(0.0177)	(0.0175)	(0.0190)	(0.0149)	(0.0140)	(0.0177)	(0.0175)	(0.0190)	(0.0149)	(0.0140)	(0.0176)	(0.0175)	(0.0190)
Constant	0.0118	0.0153	0.00690	0.0132	0.0197	-0.518**	-0.420**	-0.425*	-0.479*	-0.329+	-0.510**	-0.429**	-0.385*	-0.475*	-0.350+
	(0.0239)	(0.0174)	(0.0257)	(0.0272)	(0.0241)	(0.160)	(0.132)	(0.182)	(0.195)	(0.183)	(0.161)	(0.133)	(0.182)	(0.196)	(0.185)
Observations	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647
Number of groups	212	212	212	212	212	212	212	212	212	212	212	212	212	212	212

	Literacy	Numeracy	Social-emotional	Executive function	Motor	Literacy	Numeracy	Social-emotional	Executive function	Motor	Literacy	Numeracy	Social-emotional	Executive function	Motor
PLAY: Child-centred exploration	0.0311 (0.0603)	0.0204 (0.0501)	0.0326 (0.0681)	0.0711 (0.0732)	0.00288 (0.0691)						0.0259 (0.0615)	0.00433 (0.0515)	0.0278 (0.0698)	0.0833 (0.0754)	-0.00849 (0.0715)
PLAY: Personal/ community connections	0.124* (0.0627)	0.0750 (0.0529)	0.148* (0.0716)	0.0988 (0.0767)	-0.0186 (0.0727)						0.140* (0.0620)	0.0765 (0.0529)	0.170* (0.0713)	0.117 (0.0768)	-0.0114 (0.0732)
PLAY: Relational classroom climate	-0.153+ (0.0890)	-0.123+ (0.0730)	-0.0882 (0.100)	-0.215* (0.107)	-0.0333 (0.101)						-0.185* (0.0878)	-0.141+ (0.0727)	-0.125 (0.0997)	-0.241* (0.107)	-0.0448 (0.101)
TIPPS: Facilitating deeper learning						0.0838* (0.0329)	0.0138 (0.0279)	0.0867* (0.0378)	0.0471 (0.0407)	0.0504 (0.0386)	0.0919** (0.0329)	0.0179 (0.0280)	0.0990** (0.0377)	0.0549 (0.0407)	0.0483 (0.0388)
TIPPS: Supporting student expression						0.115* (0.0542)	0.126** (0.0452)	0.0675 (0.0618)	0.153* (0.0664)	0.0947 (0.0624)	0.119* (0.0540)	0.129** (0.0451)	0.0648 (0.0614)	0.160* (0.0660)	0.0982 (0.0625)

	Literacy	Numeracy	Social-emotional	Executive function	Motor	Literacy	Numeracy	Social-emotional	Executive function	Motor	Literacy	Numeracy	Social-emotional	Executive function	Motor
TIPPS: Emotional support/behaviour management						-0.00301 (0.0341)	0.0214 (0.0285)	-0.0167 (0.0390)	-0.0111 (0.0419)	-0.00170 (0.0394)	-0.0264 (0.0362)	0.0121 (0.0303)	-0.0491 (0.0412)	-0.0406 (0.0443)	0.00394 (0.0420)
TIPPS: Checklist, develop appropriate activities						-0.00285 (0.0126)	-0.00429 (0.0102)	0.0126 (0.0142)	-0.0156 (0.0151)	-0.0135 (0.0141)	-0.00272 (0.0126)	-0.00302 (0.0102)	0.0112 (0.0141)	-0.0142 (0.0151)	-0.0122 (0.0142)
Lagged outcome (FU)	0.628** (0.0149)	0.700** (0.0140)	0.434** (0.0177)	0.405** (0.0175)	0.405** (0.0190)	0.627** (0.0149)	0.699** (0.0140)	0.434** (0.0177)	0.403** (0.0175)	0.402** (0.0190)	0.626** (0.0149)	0.698** 0.0149)	0.433** (0.0176)	0.402** (0.0175)	0.403** (0.0190)
Constant	0.0118 (0.0239)	0.0153 (0.0174)	0.00690 (0.0257)	0.0132 (0.0272)	0.0197 (0.0241)	-0.518** (0.160)	-0.420** (0.132)	-0.425* (0.182)	-0.479* (0.195)	-0.329+ (0.183)	-0.510** (0.161)	-0.429** (0.133)	-0.385* (0.182)	0.475* (0.196)	-0.350+ (0.185)
Observations	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647	2,647
Number of groups	212	212	212	212	212	212	212	212	212	212	212	212	212	212	212

Note: Standard errors in parentheses
** p<0.01, * p<0.05, †p<0.1
Models do not include control variables.
Sample size=213 schools, 390 classrooms, 2,673 children

B.II.4.3.5 Variation by Treatment Status and KG Level

Summary of Findings

To test for subgroup differences, first models were run separately for each subgroup. Second, interactions between each PLAY factor and each grouping variable (i.e. treatment status and KG level) were added to the model to examine if any coefficients between subgroups were statistically significantly different.

Overall, there was limited evidence of systematic differences in the associations of PLAY factors with child outcomes by treatment status and KG level. In the models by treatment status, only one of the 21 interaction terms was statistically significant ($p<0.05$). This was the interaction between child-centred exploration and numeracy skills, indicating that the association was moderate sized and statistically significant for control group classrooms and zero and not statistically significant in treatment classrooms.

Regarding KG level, children in mixed KG classrooms (all KG levels combined; less than 5% of the sample) were excluded, and differences for children in KG1 versus KG2 classrooms were explored. Of the 21 interaction terms tested, none reached statistical significance. Two interaction terms reached marginal statistical significance ($p<0.06$). Specifically, the interaction between child-centred exploration and KG level significantly predicted numeracy skills, with children in KG1 showing a small, positive association (b=0.09, SE=0.09, p=0.33) and children in KG2 showing a small, negative association (b=-0.07, SE=0.06, p=0.26) (see **Table B-49**).

Second, the interaction between personal/community connections and KG level significantly predicted approaches to learning, with children in KG1 showing a small, negative association (b=-0.18, SE=0.15, p=0.31) and children in KG2 showing no association (KG2: b=0.26, SE=0.10, p=0.94) (see **Table B-50**).

Table B-49. Regression results, by treatment status for Ghana ECE observation

	School Readiness Composite	Literacy	Numeracy	Social-emotional	Executive function	Motor	Approaches to Learning
				Control			
Child-centered exploration	0.0704	0.0794	0.476**	0.0855	0.145	-0.0162	0.168
	(0.151)	(0.213)	(0.175)	(0.249)	(0.244)	(0.258)	(0.224)
Personal/community connections	-0.0364	0.0575	-0.0202	0.0582	-0.259	0.250	0.0468
	(0.156)	(0.215)	(0.178)	(0.261)	(0.247)	(0.261)	(0.248)
Relational classroom climate	-0.220	-0.235	-0.193	-0.320	-0.290	-0.208	0.357
	(0.140)	(0.196)	(0.161)	(0.232)	(0.224)	(0.237)	(0.221)
Lagged outcome (FU)	0.684**	0.615**	0.712**	0.393**	0.431**	0.279**	0.368**
	(0.0266)	(0.0274)	(0.0260)	(0.0328)	(0.0332)	(0.0298)	(0.0310)
Constant	-0.0640*	-0.0282	-0.00561	-0.152**	-0.121*	0.0390	0.0163
	(0.0317)	(0.0420)	(0.0358)	(0.0558)	(0.0494)	(0.0518)	(0.0665)
Observations	797	797	797	797	797	797	797
Number of groups	65	65	65	65	65	65	65
				Treatment (pooled)			
Child-centered exploration	-0.0149	0.0302	-0.0252	0.0134	0.0638	-0.00190	0.0566
	(0.0406)	(0.0614)	(0.0522)	(0.0684)	(0.0770)	(0.0690)	(0.0713)
Personal/community connections	0.0570	0.115+	0.0797	0.132+	0.106	-0.0435	-0.0785
	(0.0431)	(0.0648)	(0.0560)	(0.0735)	(0.0819)	(0.0740)	(0.0751)
Relational classroom climate	-0.109+	-0.148	-0.126	-0.0601	-0.216+	-0.00404	0.109
	(0.0655)	(0.0988)	(0.0817)	(0.108)	(0.122)	(0.109)	(0.117)
Lagged outcome (FU)	0.743**	0.637**	0.697**	0.444**	0.397**	0.497**	0.355**
	(0.0164)	(0.0178)	(0.0168)	(0.0211)	(0.0207)	(0.0245)	(0.0201)
Constant	0.0340+						
	(0.0200)						
Observations	1,827	1,827	1,827	1,827	1,827	1,827	1,826
Number of groups	145	145	145	145	145	145	145

Standard errors in parentheses
** p<0.01, * p<0.05, + p<0.1

	School Readiness Composite	Literacy	Numeracy	Social-emotional	Executive function	Motor	Approaches to Learning
	Control						
Child-centred exploration	0.0704 (0.151)	0.0794 (0.213)	0.476** (0.175)	0.0855 (0.249)	0.145 (0.244)	-0.0162 (0.258)	0.168 (0.224)
Personal/ community connections	-0.0364 (0.156)	0.0575 (0.215)	-0.0202 (0.178)	0.0582 (0.261)	-0.259 (0.247)	0.250 (0.261)	0.0468 (0.248)
Relational classroom climate	-0.220 (0.140)	-0.235 (0.196)	-0.193 (0.161)	-0.320 (0.232)	-0.290 (0.224)	-0.208 (0.237)	0.357 (0.221)
Lagged outcome (FU)	0.684** (0.0266)	0.615** (0.0274)	0.712** (0.0260)	0.393** (0.0328)	0.431** (0.0332)	0.279** (0.0298)	0.368** (0.0310)
Constant	-0.0640* (0.0317)	-0.0282 (0.0420)	-0.00561 (0.0358)	-0.152** (0.0558)	-0.121* (0.0494)	0.0390 (0.0518)	0.0163 (0.0665)
Observations	797	797	797	797	797	797	797
Number of groups	65	65	65	65	65	65	65
	Treatment (pooled)						
Child-centred exploration	-0.0149 (0.0406)	0.0302 (0.0614)	-0.0252 (0.0522)	0.0134 (0.0684)	0.0638 (0.0770)	-0.00190 (0.0690)	0.0566 (0.0713)
Personal/ community connections	0.0570 (0.0431)	0.115+ (0.0648)	0.0797 (0.0560)	0.132+ (0.0735)	0.106 (0.0819)	-0.0435 (0.0740)	-0.0785 (0.0751)
Relational classroom climate	-0.109+ (0.0655)	-0.148 (0.0988)	-0.126 (0.0817)	-0.0601 (0.108)	-0.216+ (0.122)	-0.00404 (0.109)	0.109 (0.117)
Lagged outcome (FU)	0.743** (0.0164)	0.637** (0.0178)	0.697** (0.0168)	0.444** (0.0211)	0.397** (0.0207)	0.497** (0.0245)	0.355** (0.0201)
Constant	0.0340+ (0.0200)						
Observations	1,827	1,827	1,827	1,827	1,827	1,827	1,826
Number of groups	145	145	145	145	145	145	145

Standard errors in parentheses
** p<0.01, * p<0.05, †p<0.1
Note: Of all coefficient differences tests, only one was statistically significant (marked in red in the table), suggesting that the association between child-centred exploration and numeracy outcomes was moderate sized and statistically significant for control group classrooms.

Table B-50. Regression results, by KG level for Ghana ECE observation

	School Readiness Composite	Literacy	Numeracy	Social-emotional	Executive function	Motor	Approaches to Learning
				KG1			
Child-centered exploration	0.0493	0.0456	0.0879	0.159	0.129	-0.0746	0.103
	(0.0802)	(0.118)	(0.0893)	(0.115)	(0.140)	(0.134)	(0.150)
Personal/community connections	0.139+	0.251*	0.115	0.164	0.0763	-0.0154	-0.176
	(0.0791)	(0.116)	(0.0883)	(0.114)	(0.138)	(0.132)	(0.148)
Relational classroom climate	-0.251*	-0.350*	-0.201	-0.183	-0.191	-0.0312	0.319
	(0.113)	(0.165)	(0.126)	(0.161)	(0.196)	(0.188)	(0.210)
Lagged outcome (FU)	0.704**	0.613**	0.716**	0.452**	0.387**	0.328**	0.285**
	(0.0234)	(0.0248)	(0.0242)	(0.0274)	(0.0267)	(0.0247)	(0.0253)
Constant	-0.0449+	-0.0761*	-0.0590*	-0.0788*	-0.129**	-0.184**	-0.119**
	(0.0234)	(0.0341)	(0.0267)	(0.0324)	(0.0391)	(0.0372)	(0.0417)
Observations	1,275	1,275	1,275	1,275	1,275	1,275	1,274
Number of groups	179	179	179	179	179	179	179
				KG2			
Child-centered exploration	-0.00624	-0.0218	-0.0702	-0.0251	0.0678	0.000984	0.0177
	(0.0502)	(0.0742)	(0.0621)	(0.0897)	(0.0813)	(0.0685)	(0.102)
Personal/community connections	0.0753	0.128+	0.111+	0.182*	0.101	-0.0277	0.0259
	(0.0505)	(0.0746)	(0.0629)	(0.0903)	(0.0823)	(0.0692)	(0.102)
Relational classroom climate	-0.0994	-0.123	-0.130	-0.0576	-0.227*	-0.0215	0.122
	(0.0696)	(0.103)	(0.0861)	(0.124)	(0.113)	(0.0951)	(0.141)
Lagged outcome (FU)	0.621**	0.570**	0.603**	0.328**	0.337**	0.497**	0.327**
	(0.0229)	(0.0230)	(0.0211)	(0.0259)	(0.0260)	(0.0478)	(0.0266)
Constant	0.102**	0.130**	0.154**	0.166**	0.170**	0.162**	0.132**
	(0.0204)	(0.0293)	(0.0250)	(0.0336)	(0.0308)	(0.0301)	(0.0380)
Observations	1,165	1,165	1,165	1,165	1,165	1,165	1,165
Number of groups	181	181	181	181	181	181	181

	School Readiness Composite	Literacy	Numeracy	Social-emotional	Executive function	Motor	Approaches to Learning
	KG1						
Child-centred exploration	0.0493 (0.0802)	0.0456 (0.118)	0.0879 (0.0893)	0.159 (0.115)	0.129 (0.140)	-0.0746 (0.134)	0.103 (0.150)
Personal/ community connections	0.139+ (0.0791)	0.251* (0.116)	0.115 (0.0883)	0.164 (0.114)	0.0763 (0.138)	-0.0154 (0.132)	-0.176 (0.148)
Relational classroom climate	-0.251* (0.113)	-0.350* (0.165)	-0.201 (0.126)	-0.183 (0.161)	-0.191 (0.196)	-0.0312 (0.188)	0.319 (0.210)
Lagged outcome (FU)	0.704** (0.0234)	0.613** (0.0248)	0.716** (0.0242)	0.452** (0.0274)	0.387** (0.0267)	0.328** (0.0247)	0.285** (0.0253)
Constant	-0.0449+ (0.0234)	-0.0761* (0.0341)	-0.0590* (0.0267)	-0.0788* (0.0324)	-0.129** (0.0391)	-0.184** (0.0372)	-0.119** (0.0417)
Observations	1,275	1,275	1,275	1,275	1,275	1,275	1,274
Number of groups	179	179	179	179	179	179	179
	KG2						
Child-centred exploration	-0.00624 (0.0502)	-0.0218 (0.0742)	-0.0702 (0.0621)	-0.0251 (0.0897)	0.0678 (0.0813)	0.000984 (0.0685)	0.0177 (0.102)
Personal/ community connections	0.0753 (0.0505)	0.128+ (0.0746)	0.111+ (0.0629)	0.182* (0.0903)	0.101 (0.0823)	-0.0277 (0.0692)	0.0259 (0.102)
Relational classroom climate	-0.0994 (0.0696)	-0.123 (0.103)	-0.130 (0.0861)	-0.0576 (0.124)	-0.227* (0.113)	-0.0215 (0.0951)	0.122 (0.141)
Lagged outcome (FU)	0.621** (0.0229)	0.570** (0.0230)	0.603** (0.0211)	0.328** (0.0259)	0.337** (0.0260)	0.497** (0.0478)	0.327** (0.0266)
Constant	0.102** (0.0204)	0.130** (0.0293)	0.154** (0.0250)	0.166** (0.0336)	0.170** (0.0308)	0.162** (0.0301)	0.132** (0.0380)
Observations	1,165	1,165	1,165	1,165	1,165	1,165	1,165
Number of groups	181	181	181	181	181	181	181

Note: Standard errors in parentheses
** p<0.01, * p<0.05, †p<0.1

B.III FINDINGS FROM ECE CAREGIVER TOOLS

B.III.1 Rater Reliability

In Colombia, observers achieved absolute agreement in the field using live/paired observations. Any data that had an absolute agreement rate of <0.70 was dropped. Three cases were dropped, for a total sample of 166 for the 0-2 tool, and four cases were dropped for a total sample of 114 for the 3-5 tool.

B.III.2 Internal Consistency

Internal consistency, which is based on item correlations, measures how much the items within an instrument measure the same construct or characteristic.

Besides inter-rater reliability, Cronbach's alpha was conducted to examine the internal consistency of each construct of the primary observation tool, as shown in in **Table B-51**. Most of the constructs had Cronbach's alpha smaller than 0.6, which suggested poor internal consistency of the originally hypothesised constructs.

Table B-51. Internal consistency of constructs in the PLAY caregiver observation tool (Colombia only)

Construct	Cronbach's alpha	
	0-2	3-5
Agency	0.71	0.57
Exploration	0.52	0.28
Social Connectedness	0.39	0.37
Positive Emotional Climate	0.52	0.67

B.III.3 Final Models for ECE Caregiver Tools

Caregiver 0–2 Tools

Table B-52 shows the model fit statistics for 1-factor confirmatory model with item-level correlations. The fit statistics demonstrate excellent model fit.

Table B-52. Model fit statistics for final Colombia caregiver (0–2) confirmatory model with item-level correlations

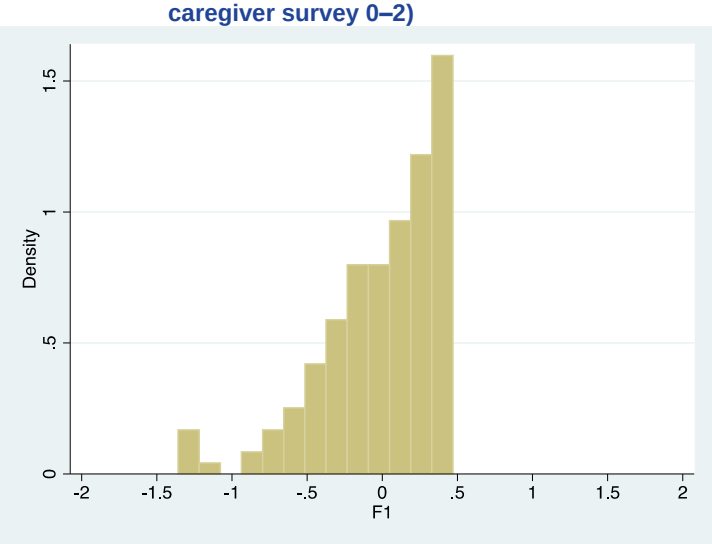
	1-factor
Sample-size adjusted BIC	3074.297
Chi-square (df)	21.236 (16)
P-value	0.1696
RMSEA	0.044
CFI	0.989
TLI	0.980
SRMR	0.033

Table B-53. Factor loadings from final Colombia caregiver (0–2) 1-factor confirmatory model

		Standardised factor loading
Factor 1: Support for child-directed exploration, α=0.83		
OAG1	Caregiver permits child to choose which material(s) to engage with	0.52
OAG2	Caregiver permits child to choose how to engage with material(s)	0.70
OAG3	Caregiver follows child's lead/initiative or expression of interest	0.79
OAG4	Caregiver provides positive facial/gesture/tone feedback that shows approval for child initiative	0.66
OAG5	Caregiver observes child before intervening	0.58
OEX2	Caregiver supports a child's motor initiative (e.g. turning an object over)	0.48
OEX3	Caregiver allows child to mouth objects, conduct simple manipulations such as rotating objects	0.61
OEX4	Caregiver allows ample time for child's inspection of discovered/available materials through reaching out, grasping, mouthing, manipulating	0.46

Descriptive statistics for the factor show that it is negatively skewed (skewness stats: f1=-1.22).

Figure B-11. Distribution Factor 1: support for child-centred exploration of materials (Colombia caregiver survey 0–2)



Caregiver 3-5 Tools

Table B-54 shows the model fit statistics for 3-factor confirmatory model with item-level. The fit statistics demonstrate excellent model fit.

Table B-54. Model fit statistics for final Colombia caregiver observation (3–5) confirmatory model with item-level correlations

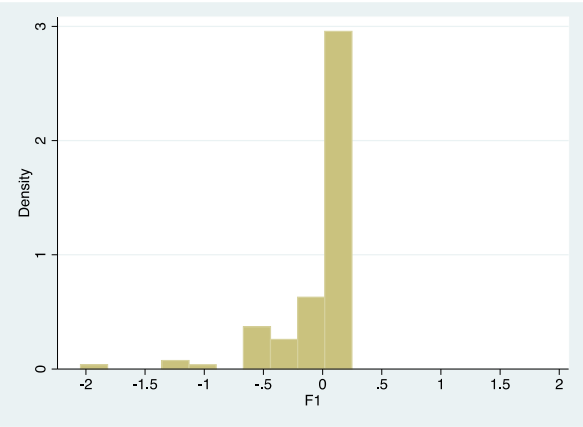
	3-factor
Sample-size adjusted BIC	2642.871
Chi-square (df)	68.691 (59)
P-value	0.1819
RMSEA	0.037
CFI	0.979
TLI	0.972
SRMR	0.065

Table B-55. Factor loadings from final Colombia caregiver observation (3–5) 3-factor confirmatory model

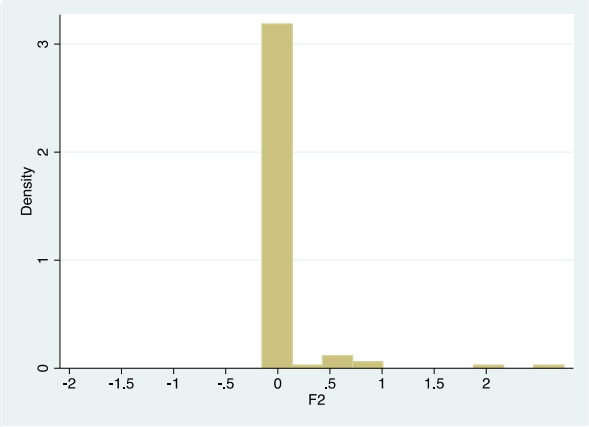
		Standardised factor loading
Factor 1: Support for agency, α=0.72		
OAG1	Caregiver permits child to choose which material(s) to engage with	0.84
OAG2	Caregiver permits child to choose how to engage with material(s)	0.75
OAG3	Caregiver follows child's lead/initiative or expression of interest	0.47
OAG5	Caregiver observes child before intervening	0.58
OEX5	Caregiver asks open-ended questions about physical objects	0.42
Factor 2: Support for problem solving, α=0.79		
OAG7	Caregiver asks questions/notices child's preferences	0.53
OPS2	Caregiver allows ample time for child to solve problems/figure something out (e.g. providing a break)	0.67
OPS3	Caregiver allows child to figure out how to do something by themselves when stuck by a challenge	0.91
OPS4	Caregiver provides ongoing feedback that facilitates problem solving	0.65
OPS5	Caregiver promotes problem-solution reflection by asking questions	0.68
Factor 3: Support for connection to experience and social connectedness, α=0.60		
OCE3	Caregiver engages the child in an activity that they may do at home together (singing a song, dancing, acting like a favourite cartoon character)	0.62
OSC2	In the presence of a new person (including the data collector), caregiver facilitates greetings/introductions	0.59
OSC5	There are multiple displays of physical affection between caregiver and child	0.53

Descriptive statistics for each of the three factors show that the factors are positively and strongly correlated to each other (f1 and f2: r=-0.0103, p=0.91; f1 and f3: r=-0.19, p<0.05; f2 and f3: r=0.45, p<0.001) and that factors 1 and 2 are skewed (skewness: f1=-2.77, f2=5.46, f3=0.84)

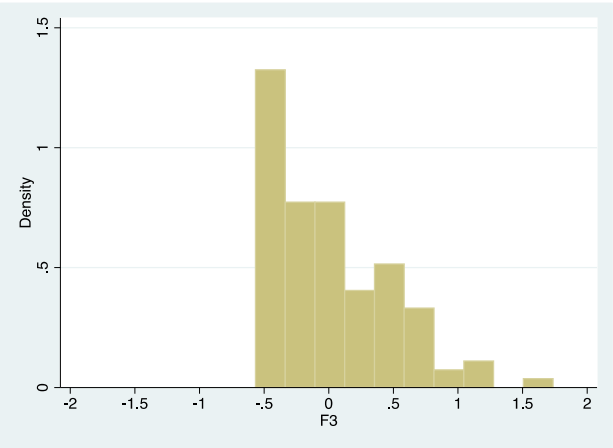
Figure B-12. Distribution of factors, Colombia caregiver observation (3–5)



CareOBS Factor 1: Support for agency



CareOBS Factor 2: Support for problem solving



CareOBS Factor 3: Support for connection to experience & social connectedness

B.III.4 Construct Validity

B.III.4.1 Correlations between Observation Tool, Caregiver Survey, and Caregiver Characteristics

<u>0-2 Measures</u>

The PLAY toolkit included a 0-2 caregiver survey. This measure exhibited three factors. Results indicated small positive correlations between each of the three factors from the caregiver survey and the one factor from the caregiver observation tool (see **Table B-56**).

Table B-56. Concurrent validity with survey measure (Colombia caregiver 0–2)

	Support for child-centred exploration of materials (observation)
	r
Support for connection to experience (survey)	0.28*
	166
Support for social connectedness (survey)	0.18*
	166
Positive emotional climate (survey)	0.26*
	166

Note: * p<0.05

The following caregiver characteristics were used in convergent validity analyses with the one factor that emerged from the PLAY caregiver observation tool.

Frequency of activities. Caregivers reported the frequency in which they engaged in the following activities with their child in the last week: book-reading, singing, playing games, and playing with a toy (response options: never, once in the last week, a few times in the last week (2–3) times, daily, or more than once a day). An average of these frequencies was used as the frequency of activities scale.

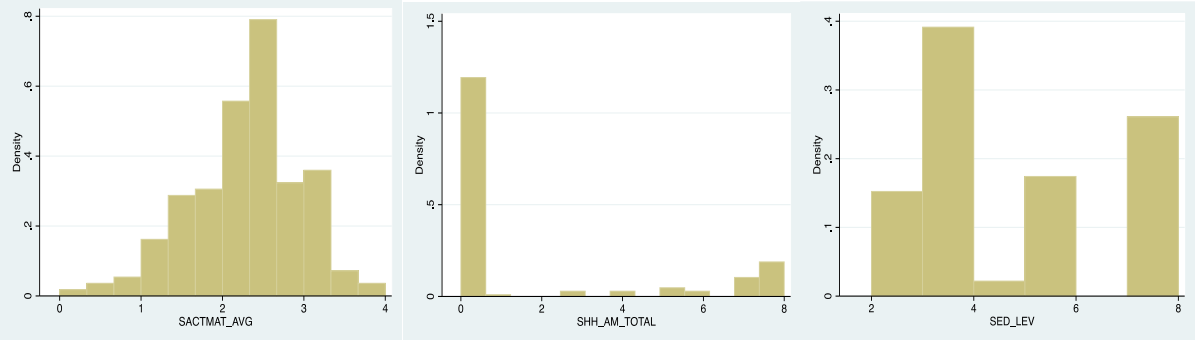
Household resources. Caregivers also reported whether they had access to household resources, including electricity, radio, television, computer, internet, refrigerator, laundry facilities, and drinking/potable water. A count of yes responses to these items was used for the household resources scale.

Caregiver level of education. Caregivers reported their highest level of formal education with the following response options: none, primary, high school, normal higher school, labour/technical, professional/technical, bachelor of arts/science, or graduate studies.

Caregiver employed. Caregivers reported whether they were currently employed in a job that provides income.

Rurality. Caregivers reported whether they were in an urban or rural area.

Figure B-13. Distribution of caregiver characteristics (Colombia 0–2)



Frequency Activities (0-2) Household Resources (0-2) Caregiver Education (0-2)

Results indicated a small positive correlation between the factors and the frequency of activities scale (see **Table B-57**). There was also a small negative correlation between the factor and rurality, suggesting that caregivers in rural households are less likely to implement these strategies. There were no significant correlations between household resources, caregiver level of education, or caregiver employment status with the factor.

Table B-57. Pairwise correlations with caregiver characteristics (Colombia 0–2)

	Support for child-centred exploration of materials
	r
Frequency of activities	0.24*
	165
Household resources	0.09
	169
Caregiver level of education	-0.003
	46
Caregiver employed	0.06
	46
Rurality	-0.18*
	166

Note: * p<0.05

3-5 Measures

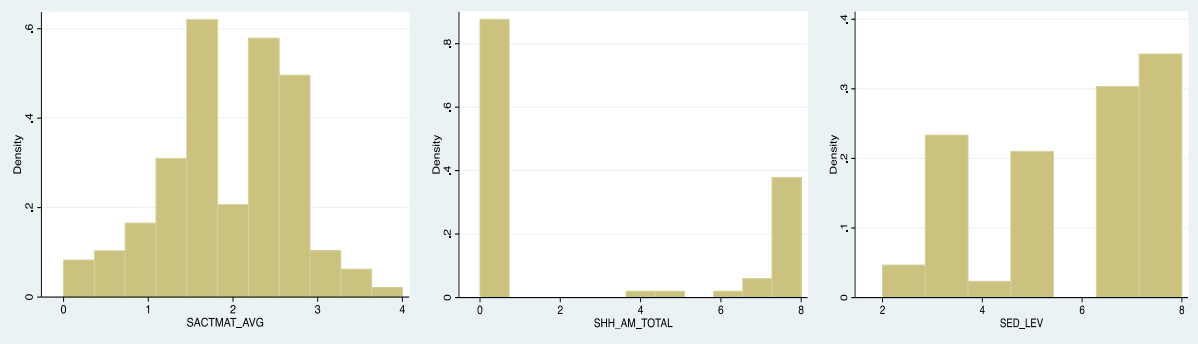
The PLAY toolkit included a 3-5 caregiver survey. This measure exhibited three factors. Results indicated no significant correlations between each of the three factors from the caregiver observation (see **Table B-58**).

Table B-58. Concurrent validity Colombia caregiver observation (3–5) with survey measure

	Support for agency (observation)	**Support for problem solving (observation)**	**Support for connection to experience and social connectedness (observation)**
	r		
Support for connection to experience (survey)	0.09	-0.05	-0.02
	114	114	114
Support for social connectedness (survey)	0.12	0.03	-0.01
	114	114	114
Positive emotional climate (survey)	0.12	0.03	0.02
	114	114	114

Note: * p<0.05

Figure B-14. Distribution of caregiver characteristics (Colombia 3-5)



Frequency Activities (3-5) Household Resources (3-5) Caregiver Education (3-5)

Pairwise correlations between the 3-5 caregiver survey and caregiver characteristics *(See above section for definition of caregiver characteristics)* indicated small to moderate positive correlations between the support for agency factor and the frequency of activities scale, the household resources scale, and the caregiver level of education (see **Table B-59**). There were no significant correlations between the support for problem solving factor and any of the caregiver characteristics. There were also no significant correlations between the support for connection to experience and social connectedness factor and any of the caregiver characteristics.

Table B-59. Pairwise correlations with caregiver characteristics (Colombia 3–5)

	Support for agency	Support for problem solving	Support for connection to experience and social connectedness
	r		
Frequency of activities	0.21*	0.08	0.09
	114	114	114
Household resources	0.18*	0.06	0.07
	118	118	118
Caregiver level of education	0.41*	0.17	-0.16
	34	34	34
Caregiver employed	0.28	0.12	-0.06
	33	33	33
Rurality	-0.04	-0.04	-0.15
	114	114	114

Note: * p<0.05

REFERENCES

Asare, K., & Nti, S. K. (2014). Teacher education in Ghana: A contemporary synopsis and matters arising. *SAGE Open*, 1–8.

Bennell, P., & Akyeampong, K. (2007). Teacher motivation in Sub-Saharan Africa and South Asia. London: DFID.

Burchinal, M. (2018). Measuring early care and education quality. *Child Development Perspectives, 12*(1), 3-9. https://doi-org.proxy.library.upenn.edu/10.1111/cdep.12260

Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., . . . & Henry, G. T. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child Development, 78*(2), 558–580.

La Paro, K. M. L., Hamre, B. K., Locasale-Crouch, J., Pianta, R. C., Bryant, D., Early, D., et al. (2009). Quality in kindergarten classrooms: Observational evidence forth need to increase children's learning opportunities in early education classrooms. *Early Education and Development, 20,* 657–692.

Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). *Maslach burnout inventory manual* (3rd ed.). Palo Alto, CA: Consulting Psychologists Press.

Moon, B. (2007). School-based teacher development in sub-Saharan Africa: Building a new research agenda. *Curriculum Journal, 18*(3), 355–371.

NICHD Early Child Care Research Network. (2002). Child-care structure→ process→ outcome: Direct and indirect effects of child-care quality on young children's development. *Psychological Science*, *13*(3), 199-206.

Osei, G. M. (2006). Teachers in Ghana: Issues of training, remuneration and effectiveness. *International Journal of Educational Development, 26*, 38–51.

Osher, D., Sprague, J., Weissberg, R. P., Axelrod, J., Keenan, S., Kendziora, K., et al. (2007). A comprehensive approach to promoting social, emotional, and academic growth in contemporary schools. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology* (5th ed., Vol. 5, pp. 1263–1278). Bethesda, MD: NASP.

Phillips, D., Mekos, D., Scarr, S., McCartney, K., & Abbott–Shim, M. (2000). Within and beyond the classroom door: Assessing quality in child care centers. *Early Childhood Research Quarterly*, *15*(4), 475-496.

Pianta, R., Downer, J., & Hamre, B. (2016). Quality in early education classrooms: Definitions, gaps, and systems. *The Future of Children*, 119-137.

Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, *9*(3), 144-159.

Tsouloupas, C. N., Carson, R. L., Matthews, R., Grawitch, M. J., & Barber, L. K. (2010). Exploring the association between teachers' perceived student misbehavior and emotional exhaustion: The importance of teacher efficacy beliefs and emotion regulation. *Educational Psychology, 30*(2), 173–189.

Vandell, D. L. (2004). Early child care: The known and the unknown. *Merrill-Palmer Quarterly (1982-)*, 387-414

Wolf, S., Aber, J. L., Behrman, J. R., & Peele, M. (2019). Longitudinal causal impacts of preschool teacher training on Ghanaian children's school readiness: Evidence for persistence and fade-out. *Developmental Science*, *22*(5), e12878.

Wolf, S., Raza, M., Kim, S., Aber, J.L, Behrman, J., & Seidman, E. (2018). Measuring process quality in pre-primary classrooms in Ghana using the Teacher Instructional Practices and Processes System (TIPPS). *Early Childhood Research Quarterly, 45*, 18-30. https://doi.org/10.1016/j.ecresq.2018.05.003