International Development and Early Learning Assessment Technical Working Paper

August 2015

Lauren Pisani, Ivelina Borisova and Amy Jo Dowd

Contents

Introduction 3

Why assess children's development and early learning? 3

Why develop IDELA? 3

How was IDELA developed? 4

Where can IDELA be used? 5

What does IDELA measure? 5

Core IDELA Domains and Skills 5

Complementary/Supplementary IDELA items 7

Caregiver survey 7

IDELA Validation 8

Internal consistency 9

Inter-rater reliability 10

Test-retest reliability 10

Construct validity 11

Age 11

Socio-economic status 13

Home learning environment 14

ECCD Program learning environment 16

Intervention effectiveness 17

Next Steps with IDELA 20

References 22

Appendix A. Internal Consistency of IDELA, by country 24

Appendix B. Associations between IDELA domains 25

Approaches to learning 25

Executive function and other IDELA domains 26

Introduction

Why assess children's development and early learning?

Early Childhood Care and Development (ECCD) programs help ensure that young children fulfill their right to healthy development and education, and ultimately assist them in reaching their full potential. There is mounting evidence from around the world demonstrating that the first years of life are critical in the development of children as they shape cognitive, social and language skills, as well as lifelong approaches to learning (Scarborough, 1998; Lonigan, Schatschneider & Westberg, 2008; Lonigan, Burgess & Anthony, 2000; Wagner, Torgesen, Rashotte, Hecht, Barker, Burgess, & Garon, 1997; Young Lives, 2010). Furthermore, children's early learning and development is directly related to their future academic success. Children who begin school with weak prior knowledge and skills in relevant emergent literacy domains, most notably general verbal abilities, basic phonological awareness, familiarity with the basic purposes and mechanisms of reading, and letter knowledge, are particularly likely to have difficulty with learning to read in the primary grades (Snow, Burns, & Griffin, 1998). Reducing the number of children who enter school with inadequate early learning experiences is an important step toward preventing school dropout and later learning difficulties and enabling long-term school success (Engle et al, 2011). Improving children's readiness for school will help them take advantage of their right to education.

Given the importance of early learning experiences, there is increasing interest in knowing whether children have the skills and knowledge necessary for success in the early primary grades. There is consensus that reliable measurement of early learning and development is needed at the local, national, and the global level. First and foremost, assessment of early learning and development ensures that ECCD programs are accountable to families, communities, and donors. In addition, high-quality assessments help to demonstrate that investments in early childhood education are in fact affecting positive change in children's lives. Reliable child development measures can also support program quality and continuous program improvement by shedding light on program strengths and areas that require further support. Further, given the gap in available ECCD information from low income countries, data from rigorous assessments can help governments monitor their progress towards early learning goals, as well as inspire attention to and scale up of effective ECCD initiatives. Finally, valid and reliable data on early learning and development in the global context can help ECCD efforts gain momentum, and put pressure on governments, donors and other stakeholders to invest in the foundational early years of life.

Why develop IDELA?

In the ECCD field, there are few international tools available that can be used to holistically measure children's development and emergent skills. In 2011, Save the Children completed a comprehensive review of the existing child development assessments and documented a number of important limitations with existing tools. Many of the instruments available were limited in their approach, either targeting only one skill area or a specific age group and many were reliant on parent or teacher report rather than directly assessing children's skills. Further, cost associated with using the instruments across countries or projects was a tangible issue as many instruments required special permission and

purchase. Most importantly, the majority of existing tools had been used primarily in high income countries, such as the United States, United Kingdom and Australia, making them difficult to adapt and easily use across countries with diverse populations and resource-poor settings.

Our review concluded that despite the existence of ECCD tools in the global space, none of the instruments available at the time offered a balance between 1) international applicability, especially within low and middle income country contexts, 2) feasibility and ease of administration and adaptation and 3) psychometric rigor. With these criteria in mind, and lessons learned from years of early childhood programming, Save the Children began the process of developing and validating the International Development and Early Learning Assessment (IDELA).

Our goal for IDELA was to develop a holistic, rigorous, open source instrument that is feasible and easily adapted to different national and cultural contexts. IDELA was developed with an aim to support continuous program improvement across Save the Children's and partners' numerous country sites, to increase accountability among ECCD initiatives globally, and to offer cohesive and ongoing data and evidence about children's learning and development across countries that can help governments and global actors to bring successful ECCD programs to scale[1].

How was IDELA developed?

The first phase of the development process began in 2011 with a pilot of more than 65 items covering four developmental domains: physical development, language, math /cognitive development and socio-emotional development. The initial set of items was drawn and adapted from existing assessments such as the Denver, the Ages and Stages Questionnaire, the Bayley Scales of Child Development, and the Early Development Instrument (EDI), among other tools. First, constructs across domains were selected and prioritized based on their international applicability and their relevance to later educational outcomes, and then items across tools were mapped out. The overlaps between tools were identified and questions for direct child assessment were adapted, re-formulated (from parent or teacher report for example), and in some instances developed from scratch. We cast a wide net as a starting point because of the lack of experience globally in administering direct assessments with young children in low income countries. Our goal was to carefully field test questions and to narrow down the initial list of possible items to the most reliable and feasible which could be used across countries and contexts with children between 42 to -78 months[2] .

The pilot, adaptation and selection of final items followed an iterative process over a period of three years and multiple sites across 11 countries (Bangladesh, Bhutan, Egypt, Ethiopia, Indonesia, Mali, Malawi, Mozambique, Pakistan, Rwanda, and Zambia). Items were selected for inclusion in the final assessment through ongoing qualitative and quantitative evaluation. Qualitatively, we observed and

[1] IDELA was not developed as an individual diagnostic or screening tool and is not meant to be used for high stakes decision making around readiness for school. Rather, IDELA's aim is to use evidence to promote best practice, inclusion and equity in ECCD provision.

[2] In some places IDELA has been used with children as young as 36 months and as old as 84 months but in most countries the most appropriate age group is 42-72 months.

documented how each item performed in terms of:

- Complexity of the adaptation and materials required across settings
- Feasibility of item administration (i.e., Can an assessor with basic education manage the administration of specific tasks, some involving multiple materials?)
- Children's understanding of the tasks at hand
- Ability to standardize training and administration on specific items
- Relevance of items tested to national ECCD standards, where available

In addition, we conducted in depth quantitative analysis of item functioning, including documenting floor and ceiling effects, internal consistency, inter-rater reliability and construct validity. With the above considerations and analysis, the testing and modifying the tool over the past three years with input from multiple country teams has resulted in a 22-item assessment that balances the three key dimensions discussed above: psychometric rigor, feasibility, and international applicability. As a result, IDELA is easily translated and administered in varied cultural contexts, and has strong reliability and validity.

Where can IDELA be used?

IDELA was developed in rural, impoverished communities across eleven low and lower middle income countries largely because these are the communities Save the Children serves and a primary goal of the tool was to support program evaluation and ECCD evidence building in low income countries. The areas where IDELA has been tested and used successfully are some of the most marginalized in the world, which is one of the tool's unique strengths. To date less testing has occurred in wealthier, urban areas, but in 2014 and 2015 (to date) IDELA has been used successfully by international and national partners in a number of urban settings as well in middle and high income countries, including Eastern and Central Europe, Australia, Egypt, and the Philippines, with positive and promising results. We feel confident that IDELA can be successfully adapted and used to inform programs and policies, strengthen ECCD evidence, and improve early learning opportunities for all children[3]. With many interested partners, we hope to continue to expand and document the use of IDELA in diverse global settings, especially marginalized urban communities and additional middle income countries.

What does IDELA measure?

Core IDELA Domains and Skills

There is a consensus among ECCD experts that readiness for school should be understood more broadly than cognitive skills, and instead is best formulated as a holistic concept involving several developmental areas, including motor, language and early literacy, math and problem solving, socio-emotional

[3] Of the 2.2 billion children in the world, 1 billion are living in poverty (UNICEF, 2005), including 76.5 million from the 41 highest income countries (UNICEF, 2014).

development, and approaches to learning[4]. Competence in all these areas will ensure that children are ready to benefit from educational activities offered in the school environment (Janus & Offord, 2000). IDELA is a direct child assessment, (as opposed to relying on parent or teacher report of children's skills), and constitutes a core of 22 items that span the five developmental domains mentioned above. These areas of development represent key early learning and development competencies that most often appear in national ECCD curricula and standards. IDELA is a skill-oriented assessment, tapping into the degree of mastery of specific skills and as such items were developed in way that allows us to see progress over time in all skill areas. Scoring is continuous in most instances (rather than a yes/no response) and a number of items contain integrated stop rules that allow for questions to be answered meaningfully by children of varying abilities and ages.

The core assessment can be used with children ages 42-78 months and takes 30 minutes to administer, on average. Application of IDELA requires a minimal set of materials: a pencil, blank paper, small items for counting (such as beans or buttons), nine picture cards related to eight items on the assessment, and a storybook that contains pictures and text. Table 1 details the skills targeted by the assessment.

Table 1. Core IDELA Domains and Skills

Gross and Fine Motor Development	Emergent Literacy and Language	Emergent Numeracy	Socio-emotional Development
Hopping on one foot	Print awareness	Measurement and comparison	Peer relations
Copying a shape	Expressive vocabulary	Classification/Sorting	Emotional awareness
Drawing a human figure	Letter identification	Number identification	Empathy
Folding Paper	Emergent writing	Shape identification	Conflict resolution
	Initial sound discrimination	One-to-one correspondence	Self-awareness
	Listening comprehension	Simple operations	
		Simple problem solving	
Approaches to Learning: Persistence, motivation and engagement			

The assessment of four of the domains is done through direct child interview, where a trained assessor sits with a child and follows a scripted protocol for each question, and the assessment of children's approaches to learning is done through assessor observation. After six of the most challenging IDELA items (in many instances novel to children), assessors are asked whether the child was persistent,

[4] Approaches to Learning is considered a distinct dimension of school readiness, and includes aspects of development such as curiosity and eagerness to learn, ability to tackle and persist at challenging or frustrating tasks, following directions, taking risks among other skills.

motivated and attentive in her/his effort to complete the task (i.e., stays on task, is not easily distracted and does not want to stop task). In addition, IDELA also includes a series of questions at the end of the child interview about the child's overall attentiveness, motivation and engagement throughout the assessment. Together these multi-dimensional observational measures are used to evaluate children's approaches to learning. The relationship between approaches to learning and the other domains is shown in Appendix B.

Complementary/Supplementary IDELA items

Executive Function

In addition to the 22 core items, there are two direct child assessment items focused on executive function which can be added to the assessment. These items are not part of the core assessment as they may not be applicable in all settings and are considered cross cutting skills that don't fall under the core school readiness domains described above. These additional items have undergone a similarly thorough process of testing and development as the 22 core items and have been found to significantly predict performance on core items. (See Appendix B for more details.)

The two executive function items available to use under IDELA assess inhibitory control and short-term memory. These two items were selected from a larger pool of eight items for the same reasons as the core items: rigor in measuring proposed constructs, feasibility of adaptation and administration across settings, and sensitivity to key child performance differentials. The inhibitory control task is a modified HSTK task in which children are asked to do the opposite of what the assessor instructs (Cameron Ponitz, C., McClelland, C., Matthews, J. S., & Morrison, F. J., 2009)[5]. The short-term memory task is a digit span activity that asks children to remember strings of numbers of differing lengths.

Health and Hygiene Knowledge

Direct child assessment items are also available and can be added to the assessment to respond to specific health and hygiene interventions that may be occurring within early childhood programs. This extended area of assessment is not focused on specific "skills" per se, but instead documents children's knowledge and practices in the following topics: hand washing, teeth brushing, latrine use, healthy food, and use of bed nets. While these items cover a wide range of topics, they do not represent a comprehensive assessment of children's health knowledge and can be supplemented as needed based on contextual factors and program interest. Use of these items is not relevant in all instances, rather program focus or national priorities may guide the decision to include these additional questions into the assessment.

Caregiver survey

A young child's home environment plays a key role in determining his or her chances for positive development. Optimal conditions include a safe and nurturing physical environment, opportunities for children to play, explore and learn, and the presence of developmentally appropriate objects, toys and

[5] HSKT refers to a children's game called head-shoulders-knees-toes. The game has been adapted into an inhibitory control exercise for young children.

books. In addition to the direct child assessment described above, IDELA also offers a caregiver survey. Use of the caregiver survey is not required with the use of the child assessment but it is highly recommended as it provides key information about the quality of children's early learning environment. Collecting information about what is happening in children's homes, along with child level data on early learning and development, provide a much needed, nuanced picture of how the overall quality of care and support affects the developmental outcomes of children in the long and short term. A caregiver survey can also help identify specific targeted interventions needed in early childhood development.

The caregiver survey under IDELA focuses on the topics listed in Table 2 and draws heavily on the fourth round of the Multiple Indicator Cluster Survey (MICS4) developed by UNICEF[6]. Most questions in sections 3, 4, and 5 were adapted or taken directly from the ECCD module in MICS 4. The caregiver survey requires an interview with a primary caregiver and takes between 20-30 minutes to administer. Similar to the direct child assessment, additional items and sections can be added to the survey to meet specific contextual needs.

Table 2. IDELA Caregiver survey

Section	Description
1. General family information	Sex of child, child age, number of children at home, ethnicity, parental literacy, parental education, languages spoken at home
2. ECCD experience and educational expectations	Child participation in ECCD programs, details of participation, parental expectation and aspirations of child's educational attainment
3. Access to early learning materials and resources at home	Types of reading materials at home, types of toys at home
4. Parenting practices and support for learning and development	Adults in the home engaging with children to promote learning and development
5. Inadequate care	Children left alone or in the care of another young child
6. Caregiver self-efficacy	Attitudes about parent's role in child's development
7. Socio-economic status	Roof and wall of home materials, objects/appliances owned, land/animals owned

IDELA Validation

This section details the quantitative analysis and psychometric tests that have been conducted as a part of IDELA's validation, focusing specifically on internal consistency, inter-rater reliability, and construct validity.

[6] http://www.unicef.org/statistics/index_24302.html

Internal consistency

Numerous IDELA datasets have been analyzed to investigate individual item functioning and arrive at the 22 core IDELA items discussed above. To date IDELA has been used in 15 countries and some countries have collected IDELA with different populations of children from different project sites. Here we present analysis of the most recently available datasets, selected to represent a diverse global sample. Thus, a total of 5,304 children are included in this analysis from 11 countries where data was collected in 2013 and 2014 (Table 3).

Table 3. Country samples used to document internal consistency with IDELA

Country	Sample Size
Bangladesh	600
Bhutan	99
Egypt	444
Ethiopia	387
Indonesia	148
Malawi	748
Mali	1260
Mozambique	161
Pakistan	473
Rwanda	722
Zambia	262
Total	5,304

Internal consistency measures the correlation between items that propose to measure the same construct. Thus internal consistency calculations were performed for both the overall IDELA instrument and four of the subscales. The analyses produced standardized Cronbach's alphas and use George and Mallery's (2003) rules of thumb for interpreting the alpha: $\alpha > .9$ is Excellent, $\alpha > .8$ is Good, $\alpha > .7$ is Acceptable, $\alpha > .6$ is Questionable, $\alpha > .5$ is Poor, and $\alpha < .5$ is Unacceptable. As can be seen in Table 4, all domains meet acceptable internal consistency ratings, and the overall instrument has excellent internal consistency. Detailed information regarding the internal consistency for each country sample can be found in Appendix A.

Table 4. Average Internal consistency of IDELA domains and overall instrument

	Minimum	Maximum	Average
Motor Development	0.81	0.88	0.84
Emergent Literacy and Language	0.68	0.89	0.77
Emergent Numeracy	0.71	0.92	0.79
Socio-emotional Development	0.68	0.82	0.75
Total IDELA Instrument	0.84	0.95	0.90

Additional analyses of internal consistency are carried out as IDELA is used in new countries or in diverse national samples, and we expect to continue to document satisfactory levels. For example, information

regarding the internal consistency of the Approaches to Learning domain will be added as more data becomes available in 2015.

Inter-rater reliability

IDELA is typically administered by trained enumerators or trained community members. This can include local teachers, government officials, university students, community organizers and others. Priority is given to individuals who have previous experience working with young children but no formal training is required. Training on IDELA typically lasts for four to five days and includes in-office exercises and hands-on field training. That is, enumerators first practice using IDELA with each other in a controlled setting, and then in pilot testing locations with young children in similar communities to those that will be included in the study sample.

To document inter-rater reliability in administering IDELA, 10 percent of children in the overall sample are assessed by two enumerators simultaneously. Long one-way ANOVA techniques were used to calculate the intra-class correlation within pairs of assessors for a measure of reliability. Ratings typically use Fleiss' benchmarks for excellent (ICC > 0.75), good or fair (0.75 >= ICC > 0.4), and poor (0.4 >= ICC). Table 5 displays two examples of inter-rater reliability results from data collected at two sites in September and October 2014. Enumerators in Malawi were a mix of men and women from the local community and all testing locations were rural. Enumerators in Egypt were all women working as community organizers for local NGOs or for the local government, and the testing occurred in rural and semi-urban areas. As seen in Table 5, inter-rater reliability among two different groups of assessors was very high, giving confidence that IDELA is being implemented with fidelity in varied settings.

Table 5. Inter-rater reliability, Malawi and Egypt 2014

	Motor Development	Emergent Literacy	Emergent Numeracy	Socio-emotional Development
Intra-class correlation Malawi	0.89	0.89	0.88	0.89
Rating Malawi	Excellent	Excellent	Excellent	Excellent
Intra-class correlation Egypt	0.93	0.88	0.87	0.9
Rating Egypt	Excellent	Excellent	Excellent	Excellent

Test-retest reliability

A measure of test-retest reliability was undertaken in Ethiopia in June 2015 after a regularly scheduled data collection had been completed in May 2015. Approximately three weeks passed between the original data collection and retest data collection, and the retest sample included 100 children in six villages from the original data collection. The results shown in Table 6 display results of an intra-cluster correlational analysis which found that the test-retest reliability of the IDELA subscales were fair/good

or excellent and the overall IDELA score was excellent (Fleiss, 1986). Also, the differences between original and retest averages was in the expected direction, with children's scores increasing over time.

Table 6. Test-retest reliability

	Original Average (N=100)	Retest Average (N=100)	Retest – Original Average	Intra-cluster correlation (ICC)	Rating
Motor	90%	90%	0%	0.82	Excellent
Literacy	68%	70%	3%	0.79	Excellent
Numeracy	70%	72%	2%	0.66	Fair/Good
Socio-emotional	73%	79%	6%	0.62	Fair/Good
IDELA	73%	76%	3%	0.79	Excellent
Note: ICC > .75 = Excellent, .40 - .75 = Good/Fair, < .40 = Poor. (Fleiss, 1986)					

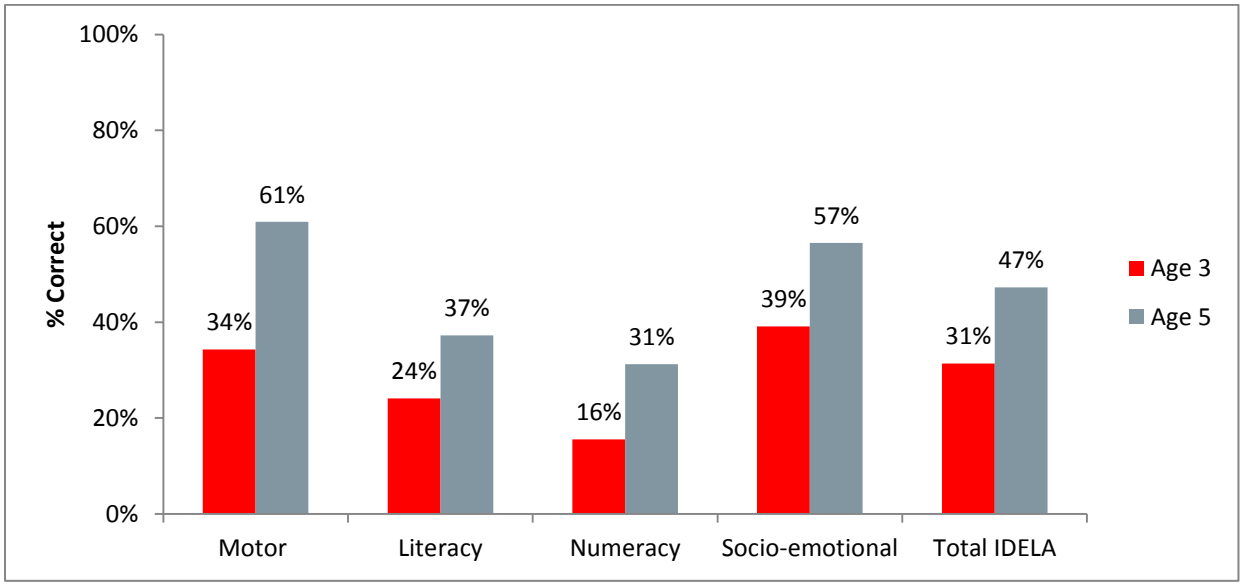
Construct validity

Construct validity is defined as the extent to which an assessment effectively measures the topics it proposes to measure. IDELA is intended to be an instrument that measures children's developmental abilities and skills and one that can be used effectively for programmatic reflection as well as policy decision making. Thus it was critical that IDELA be sensitive to potential differences in children's abilities by background characteristics as well as intervention inputs. The sensitivity of IDELA to differences along hypothesized or known dimensions of importance such as age, socio-economic status, home learning environment, ECCD program learning environment and intervention effectiveness, have been thoroughly tested. Given that the samples in this report were not collected with the sole purpose of validating IDELA, not all datasets contain all the relevant background information about children participating in the assessment. For example, a number of field sites chose not to collect caregiver surveys along with the child assessment so analysis of links between home environments and socio-economic status with IDELA are not possible in every dataset. Therefore, examples of construct validity will be shown in this section along with detailed correlation tables where possible.

Age

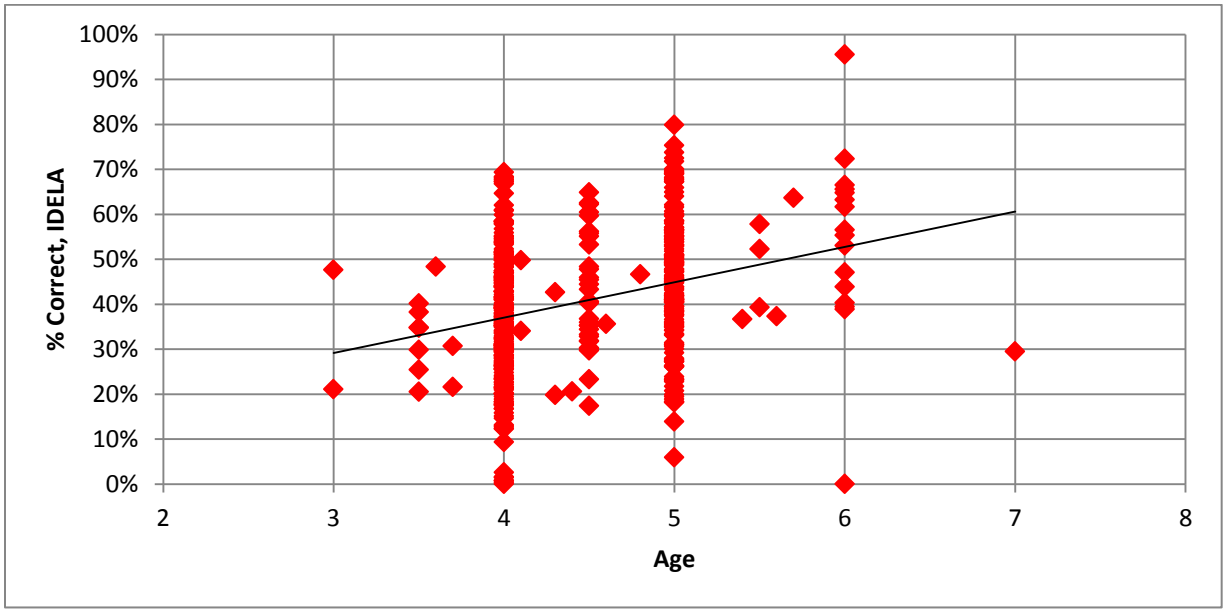
IDELA successfully captures differences in children's learning and development as they mature. Figure 1 displays significant skill variation between 3 and 5-year-old children from a study in Zambia (r = .36), and Figure 2 displays correlation between age and IDELA scores in Egypt (r = .30).

Figure 1. Skill variation by age using IDELA, Zambia 2013



Note: All differences are statistically significant, p <.001 using regression analyses clustered at the school level.

Figure 2. Skill variation by age using IDELA, Egypt 2014



The level of detail within which age information varies between countries and in some places the exact age or birthdate of many children is unknown. Thus the strength of the relationship between age and IDELA varies between countries, as seen in Table 7.

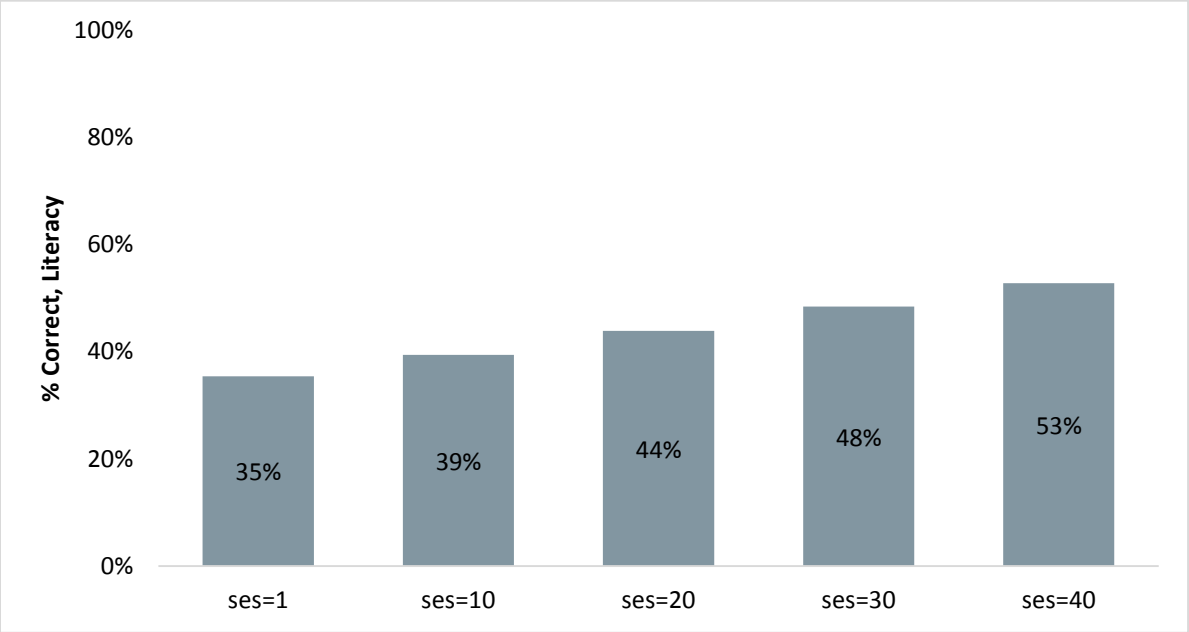
Table 7. Correlations between IDELA domains and age

	Motor	Emergent Literacy	Emergent Numeracy	Socio-emotional	IDELA
Bangladesh	0.15	0.09	0.19	0.06	0.15
Bhutan	0.39	0.38	0.32	0.40	0.46
Egypt	0.30	0.23	0.25	0.18	0.30
Ethiopia		0.12	0.12	0.05	0.10
Mozambique		0.30	0.26	0.15	0.24
Pakistan		0.27	0.38	0.30	0.39
Rwanda	0.24	0.12	0.18	0.21	0.26
Zambia	0.40	0.32	0.32	0.31	0.36

Socio-economic status

Using IDELA's direct child assessment and caregiver survey together, allow us to capture skill variation by socio-economic status. Figure 3 displays differences in emergent literacy and language related to socioeconomic status found with children in Ethiopia before ECCD programming began. An index of socioeconomic status was created from the caregiver survey using items about rooms in the home, electricity, appliances (TVs, mobiles, radios, and bicycles), land and livestock. The index ranges from 1 (low SES) to 41 (high SES), with a median and mean of 14.

Figure 3. Skill variation by socio-economic status using IDELA, Ethiopia 2013



Note: Socio-economic status significantly predicts Emergent Literacy at $p < .05$, clustered at the school level controlling for mother and father's education, child age and intervention status.

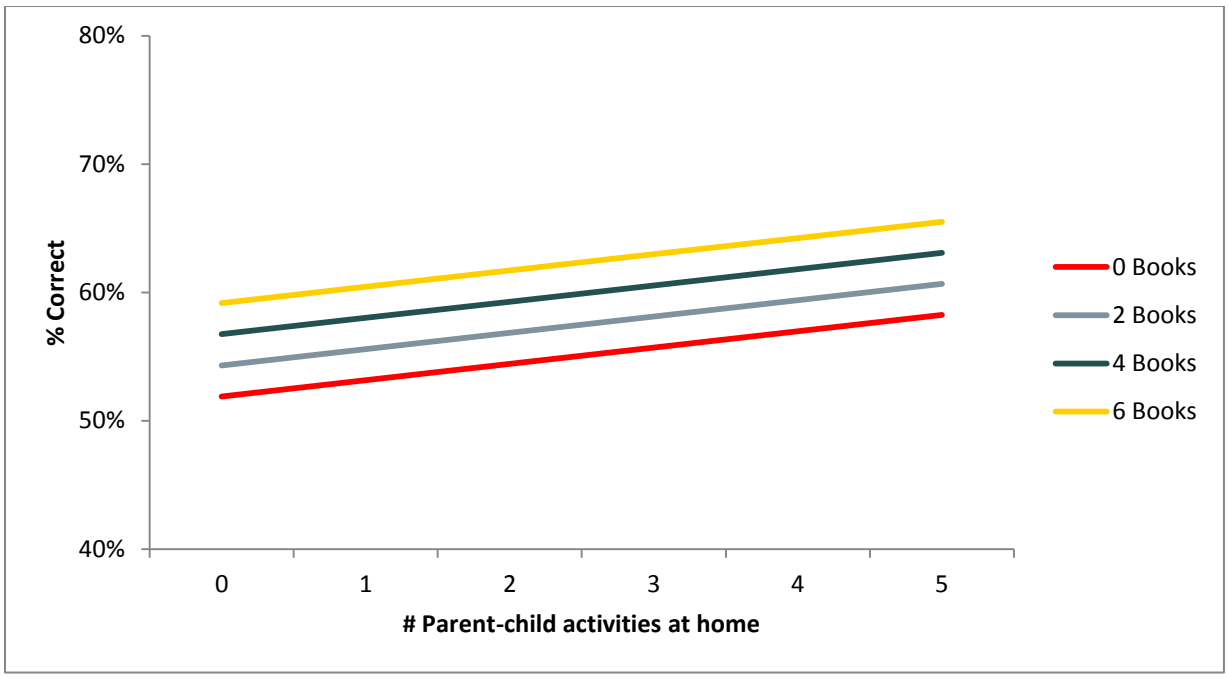
Table 8. Correlations between IDELA subdomains and socio-economic status (SES)

	Motor	Literacy	Numeracy	Socio-emotional	Total IDELA
Ethiopia		0.18	0.13	0.09	0.16
Rwanda	0.22	0.23	0.22	0.13	0.26

Home learning environment

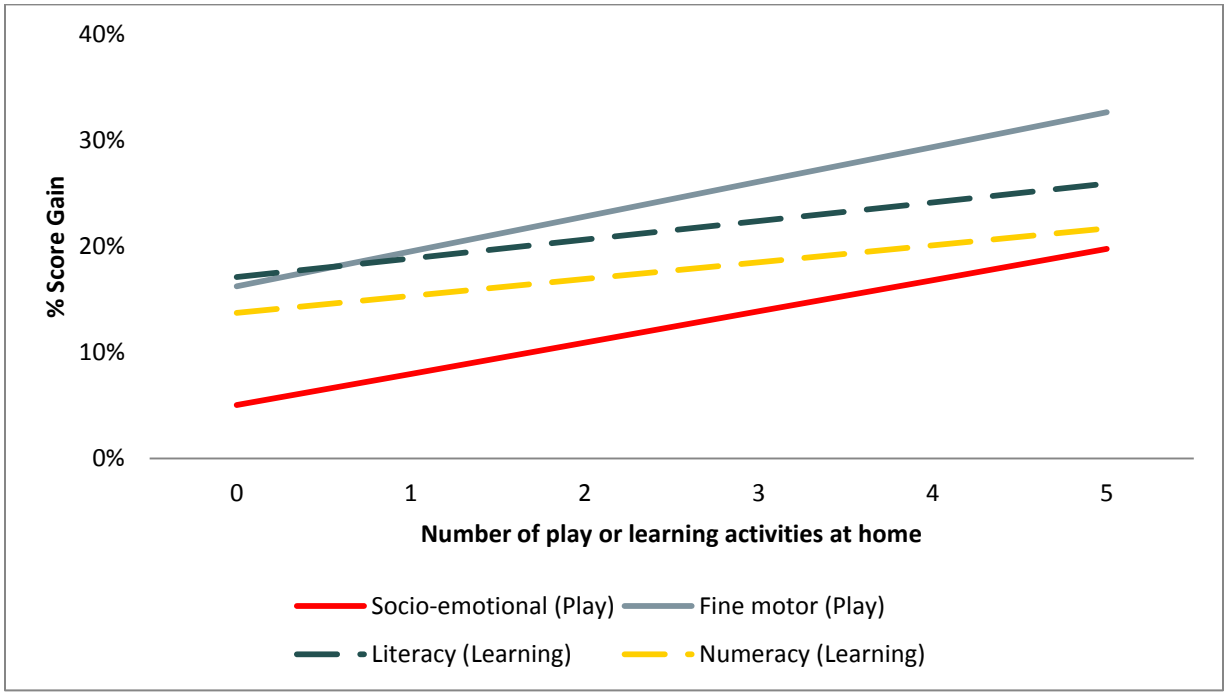
Multiple large scale international studies have documented the importance home environments for children's early learning and development (Mullis, 2007). Thus the IDELA Caregiver Survey asks detailed questions about not only the reading and other learning materials children have at home but also the play and learning activities that caregivers engage in with their children. As research suggests, our analyses indicate that across countries children who have better access to learning materials and are more stimulated by caretakers at home tend to have stronger development and early learning outcomes (Dowd & Pisani, 2013; Friedlander, Dowd, Guajardo, 2012; Wagner, Lockheed, Mullis, Martin, Kanjee, Gove, & Dowd, 2012). Figure 4 provides an example of the additional skills displayed by children who have access to different types of reading materials at home and whose caregivers reported engaging in - activities supporting early development and learning (e.g., telling stories, singing, playing, etc.) before any additional program support is provided. Figure 5 displays the relationship between play activities happening in children's homes and their skill gains across domains in Rwanda.

Figure 4. Average IDELA score, by number of reading material types at home and number of activity types engaged in by caregivers, Bangladesh 2014



Notes: Gains from additional book types and learning activities at home are significant $p < .05$, controlling for child age, sex, and SES and clustering at the center level.

Figure 5. Relationship between play behaviors at home and skill gains, Rwanda 2014



Note: Gains shown control for baseline scores as well as relevant background characteristics.

Table 9. Correlations between IDELA domains and home learning environment

		Motor	Literacy	Numeracy	Socio-emotional	**IDELA**
Bangladesh	# types of reading materials		0.18	0.12	0.14	**0.18**
	# types of toys/learning materials		0.18	0.14	0.13	**0.18**
	Play activities		0.17	0.10	0.18	**0.16**
	Learning activities		0.18	0.11	0.19	**0.18**
Ethiopia	# types of reading materials		0.23	0.22	0.21	**0.24**
	# types of toys/learning materials		0.21	0.17	0.21	**0.21**
	Play activities		0.22	0.21	0.14	**0.21**
	Learning activities		0.26	0.25	0.18	**0.25**
Rwanda	# types of reading materials	0.06	0.09	0.05	0.09	**0.04**
	# types of toys/learning materials	0.10	0.10	0.07	0.14	**0.12**
	Play activities	0.24	0.20	0.23	0.28	**0.35**
	Learning activities	0.19	0.28	0.29	0.25	**0.32**

ECCD Program learning environment

Investigating the relationship between children's outcomes assessed by IDELA and the quality of the learning environments in ECCD centers children attend has been an important connection to establish. We find that IDELA scores do positively correlate with the quality of ECCD learning environments. For example in Rwanda, we find that ECCD center quality, as measured using an adapted version of the Early Childhood Environment Rating Scale (ECERS), is significantly correlated with the average emergent numeracy and emergent literacy of the children in these centers (Figure 6).

Figure 6. Correlation between quality learning environment (ECERS) and IDELA emergent numeracy and literacy scores, Rwanda 2014

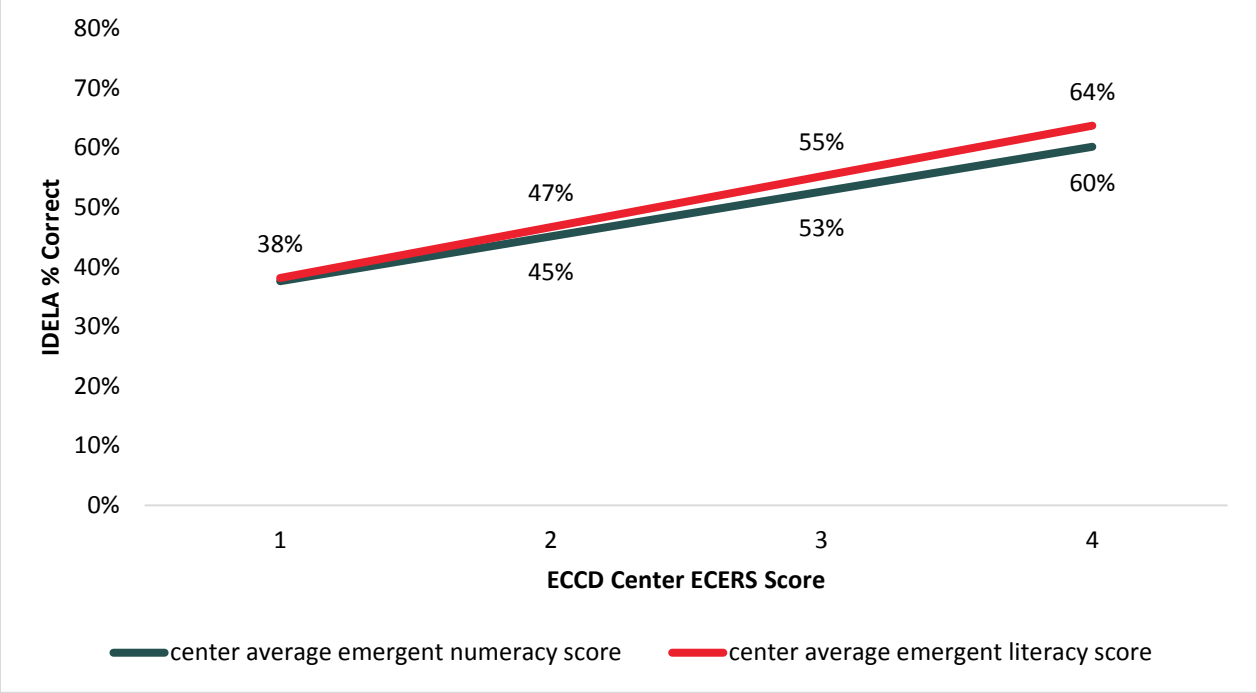


Table 10. Correlations between ECERS and Average IDELA scores

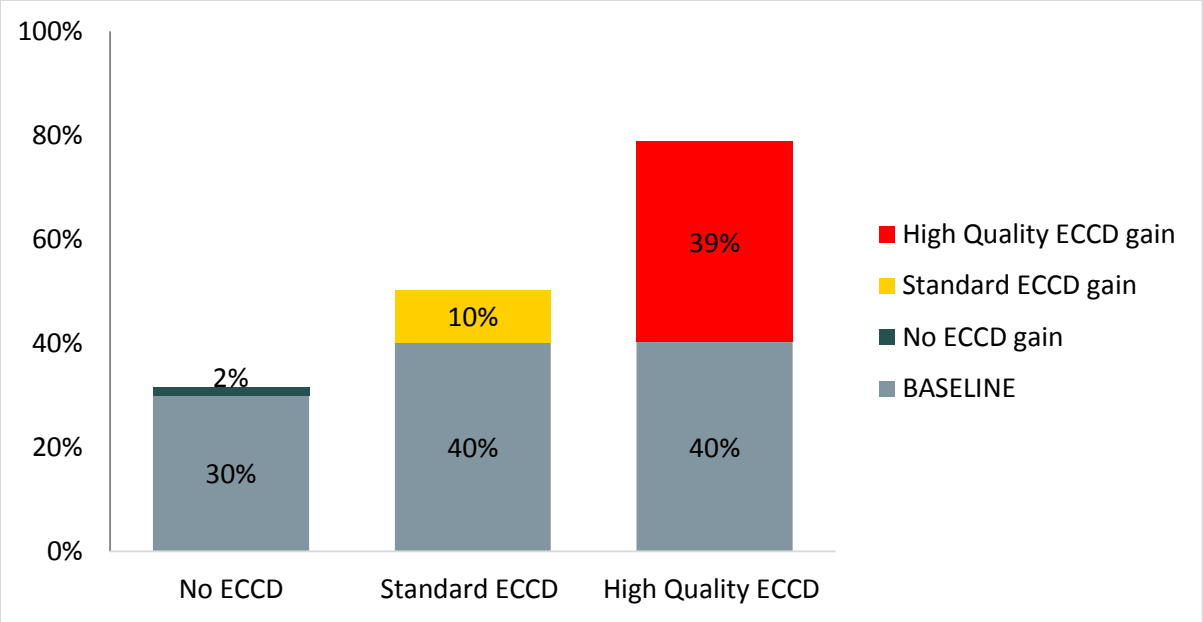
ECERS Subscale	Emergent Numeracy	Emergent Literacy	IDELA
Space and Furnishings	.43	.20	.21
Activities	.45	.61	.36
Interactions	.77	.52	.47
Programme Structure	.53	.54	.31
Language and Literacy	.50	.52	.41
Mathematics	.67	.53	.63
ECERS Total	.68	.55	.45

Intervention effectiveness

IDELA's ability to document children's early development as it relates to different programmatic inputs is also of great interest as it allows reflection and comparison of different ECCD program approaches and ultimately evidence based decision making around scaling up of successful programs. Across multiple

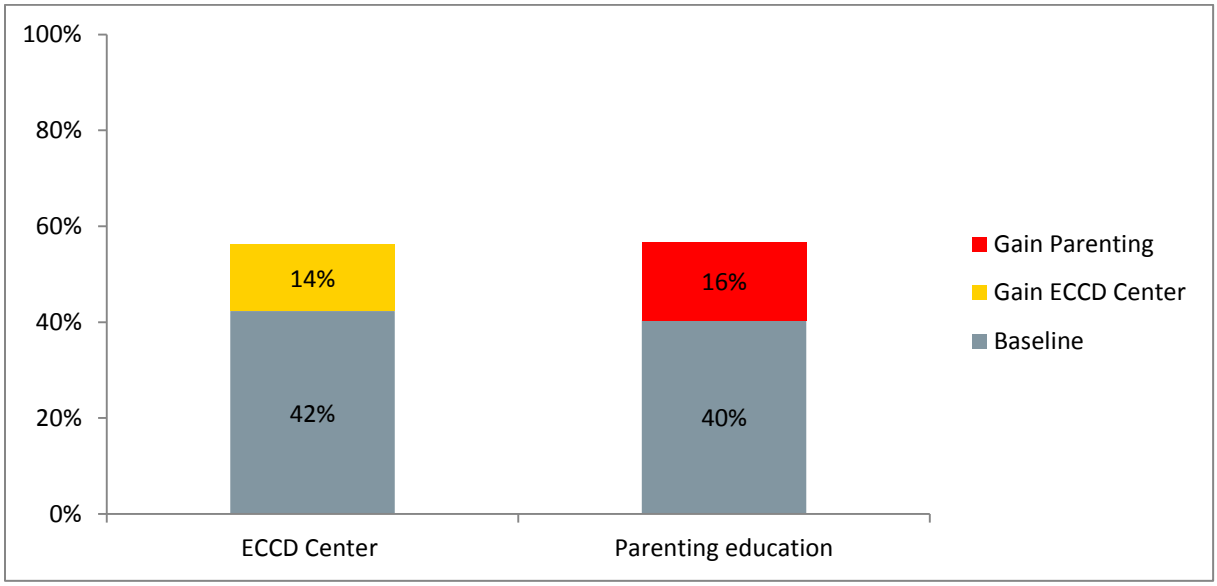
settings, IDELA has meaningfully captured variation in children's learning gains based on differing levels of programmatic inputs. In the case of Ethiopia shown in Figure 7, we compared the effectiveness of two different program packages. One group of children had no exposure to an ECCD program over the course of a school year (No ECCD), another group attended a typical ECCD program in the area (Standard ECCD), and the third group attended a higher quality ECCD program with a special focus on strengthening emergent literacy and math skills and additional inputs in terms of classroom resources (books, manipulatives) as well as teacher mentoring and supervision (High Quality ECCD). It is clear from Figure 7 that IDELA allowed the documentation of the varied gains across the two program packages.

Figure 7. Average IDELA baseline and gain scores by program input using IDELA, Ethiopia 2014



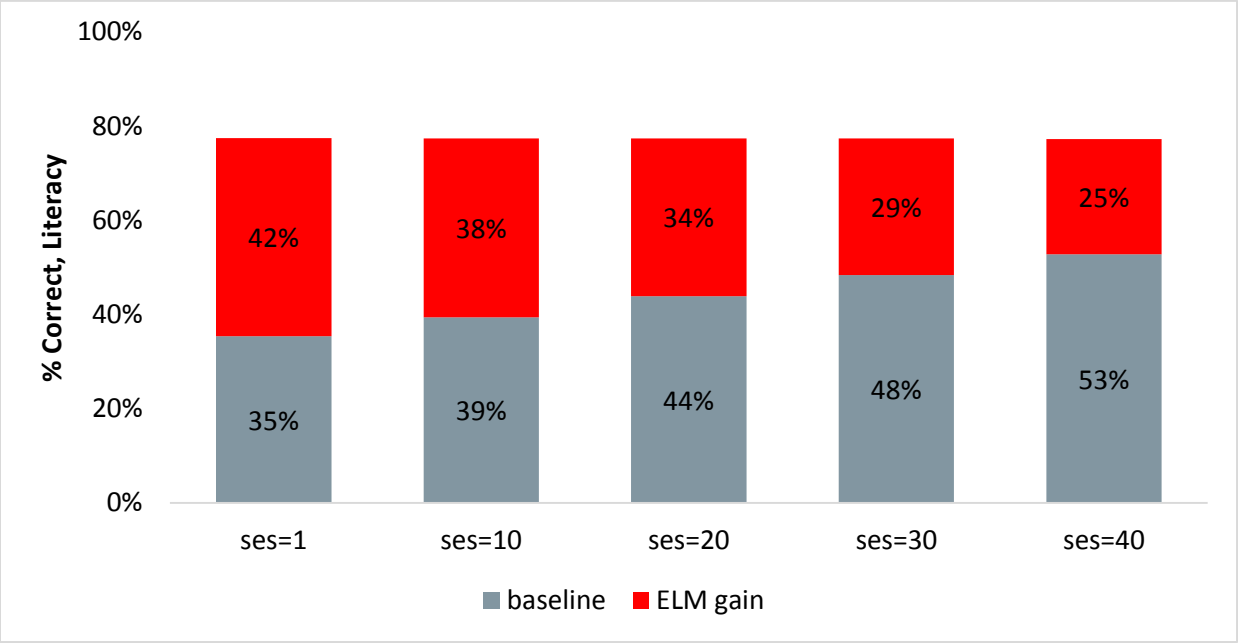
In Rwanda, we used IDELA to compare two completely different program approaches to ECCD services - a traditional ECCD center based program and a parent outreach education program (Figure 8). In this instance, IDELA's sensitivity to children's learning and development allowed local and national partners to have substantive discussions about the important role that parenting education can play for children who do not have access to ECCD centers.

Figure 8. Average IDELA baseline and gain scores by program input using IDELA, Rwanda 2014



Further, IDELA enables programs to reflect not only on the overall quality and effectiveness of an intervention, but also on the equitable outcomes of the intervention. For example, using IDELA as an evaluation tool in Ethiopia before Save the Children's Early Literacy and Math (ELM) intervention began allowed program teams to observe the differences in children's development related to socio-economic status at baseline (shown in Figure 3) and again in a follow-up study. Figure 9 displays the baseline scores and subsequent gains after 10 months of high quality intervention focused on strengthening emergent literacy and math and demonstrates the ability of this program to really serve as an equalizer for children from the poorest families.

Figure 9. Average IDELA baseline and gain scores by socio-economic status using IDELA, Ethiopia 2014



Note: Socio-economic status significantly predicts Emergent Literacy at $p < .05$, controlling for mother and father's education, child age and intervention status, clustered at the school level.

Next Steps with IDELA

The almost four year process of developing and validating IDELA has resulted in a rigorous, holistic, yet feasible and reliable international instrument for measuring early learning and development of children 48 to 72 months. With 22 items IDELA is one of the shortest assessments that provides a holistic picture of learning and development in the early years. IDELA has not been validated at the diagnostic level for the screening individual children. Currently IDELA does not set thresholds for "school readiness" and an IDELA score in a certain range cannot be taken as indicative of a serious problem. We believe IDELA's most important contribution globally is its ability to measure holistic early learning and development at the group or population level. For this purpose the psychometric properties of the instrument described in this paper are exceptionally strong and on par with, or exceeding, other available tools and questionnaires used to assess school readiness. One of IDELA's greatest strengths and what sets it apart from other available ECCD tools is the proven feasibility and adaptability to varied contexts. IDELA can be administered consistently, with high levels of fidelity in low resource settings and is sensitive to important elements of children's early learning and development.

The analyses to date also raise a number of questions for future research. We will to continue to monitor item performance, internal consistency and inter-rater reliability of the tool in new settings, with a strong focus on the socio-emotional development and approaches to learning domains. Additionally, we will continue to monitor links between the core IDELA instrument and measures of executive function with the hope of bolstering the tools sensitivity to health-related issues for young

children. Another area that warrants further attention is the predictive validity of the tool. We look forward to the opportunity to investigate the links between specific school readiness skills and performance in a primary education setting.

To date the instrument has been used in over 16 countries with many more planning for its adaptation and use in 2015 alone. Numerous partners, including national governments, academic institutions and local and international implementing organizations are supporting the further use and learning from IDELA. In 2015 and beyond, we are looking at the application of IDELA in nationally representative populations in a number of settings. The international applicability of the assessment offers exciting opportunities for building a global evidence base for ECCD and contributing to national and international conversations enabling children all over the world to fully access their right to healthy development and quality education. Finally, IDELA was developed with the goal of being widely used as an open source tool. Save the Children is working on effective ways to share our experiences to date, including manuals, reports and supporting materials with interested partners without losing the main goals of the assessment and its rigorous application.

References

Bowey, J. A. (1995). Socioeconomic status differences in preschool phonological sensitivity and first grade reading achievement. *Journal of Educational Psychology*, *87*(3), 476.

Cameron Ponitz, C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology, 45*, 605–619.

Dowd, A., Pisani, L., (2013). Two wheels are better than one: The importance of capturing home literacy environment in large-scale assessments of reading. *Research in Comparative and International Education, 8,* 359-372.

Engle, Pat L, Lia C.H. Fernald, Harold Alderman, Jere Behrman, Chloe O'Gara, Aisha Yousafzai, et al. "Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries." *The Lancet.* 378 (9799), 1339-1353. 8 October 2011.

Fleiss J. The Design and Analysis of Clinical Experiments. New York: John Wiley & Sons, 1986.

Fleiss, J. L. (1971) "Measuring nominal scale agreement among many raters." *Psychological Bulletin*, *76* (5), 378–382.

Friedlander, E., Dowd, A.J., Borisova, I. & Guajardo, J. (2012). *Life-wide learning: Supporting all children to enjoy quality education*. NY: UN Women & UNICEF. http://www.worldwewant2015.org/node/283236

George and Mallery (2003). SPSS for Windows step by step. *A simple guide and reference. 11.0 update* (4th ed.). Boston: Allyn & Bacon.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. National Academies Press.

Lonigan, C. J., Schatschneider, C., & Westberg, L. (2008). Identification of children's skills and abilities linked to later outcomes in reading, writing, and spelling. *Developing early literacy: Report of the national early literacy panel,* 55-106.

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: evidence from a latent-variable longitudinal study. *Developmental psychology*, *36*(5), 596.

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *IEA's progress in international reading literacy study in primary schools in 40 countries*. Chestnut Hill, MA: TIMMS and PIRLS International Study Center, Boston College.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. *Specific reading disability: A view of the spectrum*, 75-119.

United Nations Educational, Scientific, and Cultural Organization (UNESCO). (2005). *The State of the World's Children: Childhood under threat.* Paris: France.

United Nations Educational, Scientific, and Cultural Organization (UNESCO). (2014). *Children of the Recession: The impact of the economic crisis on child well-being in rich countries.* Paris: France.

Wagner, D. A., Lockheed, M., Mullis, I., Martin, M.O., Kanjee, A., Gove, A. & Dowd, A.J. (2012): *The debate on learning assessments in developing countries*, Compare: A Journal of Comparative and International Education, 42:3, 509-545.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child development, 69(3),* 848-872. Chicago, IL.

Young Lives, (2010). Early Childhood Care and Education as a Strategy for Poverty Reduction: Evidence from Young Lives. Young Lives Policy Brief 9.

Appendix A. Internal Consistency of IDELA, by country

Table A1. Internal consistency of IDELA, by country

	Motor	Literacy	Numeracy	Socio-emotional	IDELA
Mozambique (N=131)	0.85	0.68	0.71	0.71	0.84
Zambia (N=262)	0.87	0.85	0.85	0.82	0.94
Pakistan (N=473)		0.74	0.76	0.75	0.85
Bhutan (N=99)	0.85	0.71	0.76	0.76	0.89
Ethiopia (N=387)		0.89	0.92	0.81	0.95
Mali (N=1260)	0.84	0.79	0.79	0.79	0.91
Bangladesh (N=600)	0.88	0.73	0.77	0.68	0.89
Indonesia (N=148)	0.83	0.81	0.72	0.68	0.88
Malawi (N=748)	0.82	0.73	0.78	0.75	0.91
Rwanda (N=722)	0.81	0.82	0.81	0.79	0.92
Egypt (N=444)	0.84	0.75	0.80	0.74	0.91

Appendix B. Associations between IDELA domains

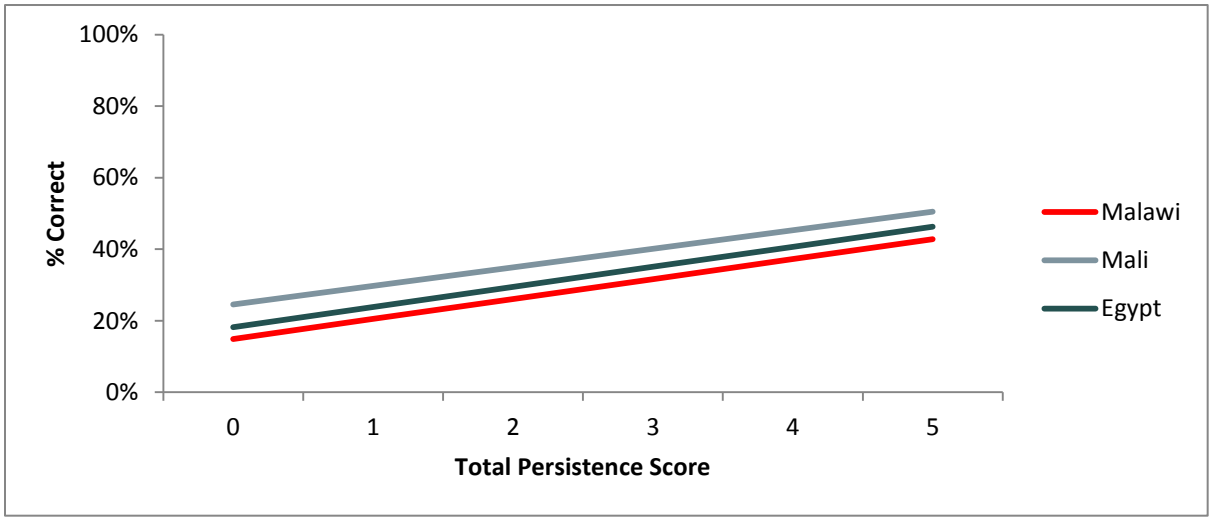
Table B1. Correlation between the emergent literacy domain and other IDELA domains

	Motor	Emergent Numeracy	Socio-emotional
Zambia	0.62	0.68	0.69
Pakistan		0.64	0.35
Bangladesh		0.60	0.51
Egypt	0.54	0.60	0.50
Malawi	0.51	0.63	0.58
Mali	0.50	0.51	0.55
Indonesia		0.63	0.59
Ethiopia		0.60	0.51
Bhutan	0.56	0.60	0.39
Mozambique		0.44	0.35
Rwanda	0.67	0.74	0.70
Average	0.57	0.61	0.52

Approaches to learning

The original IDELA measure of approaches to learning is centered on assessor's observations of children's persistence in answering complex assessment questions. We have found this simple measure of approaches to learning to be significantly correlated with IDELA scores across countries. Examples from Malawi, Mali and Egypt are shown below in Figure B1. Since observing the success of these persistence items we have expanded IDELA to include a more nuanced measure of approaches to learning including assessor observation of children's persistence, attentiveness and motivation. Data on these measures will be available in 2015.

Figure B1. Persistence predicting overall IDELA score, Malawi, Mali and Egypt 2014



Executive function and other IDELA domains

Significant relationships between direct assessment of executive function and overall IDELA scores as well as between observational measures of children's learning approaches with IDELA scores have been observed in multiple settings. Both measures of executive function (inhibitory control[7] and short-term memory) are found to significantly predict IDELA scores, with effect sizes ranging from .21 to .54 across IDELA domains and the overall IDELA score. Figure B2 displays an example of the predictive power of the inhibitory control and working memory tasks on overall IDELA score in Egypt and Table B2 provides a full description of the correlation with all IDELA domain scores.

Figure B2. Inhibitory control and short-term memory predicting overall IDELA scores, Egypt 2014

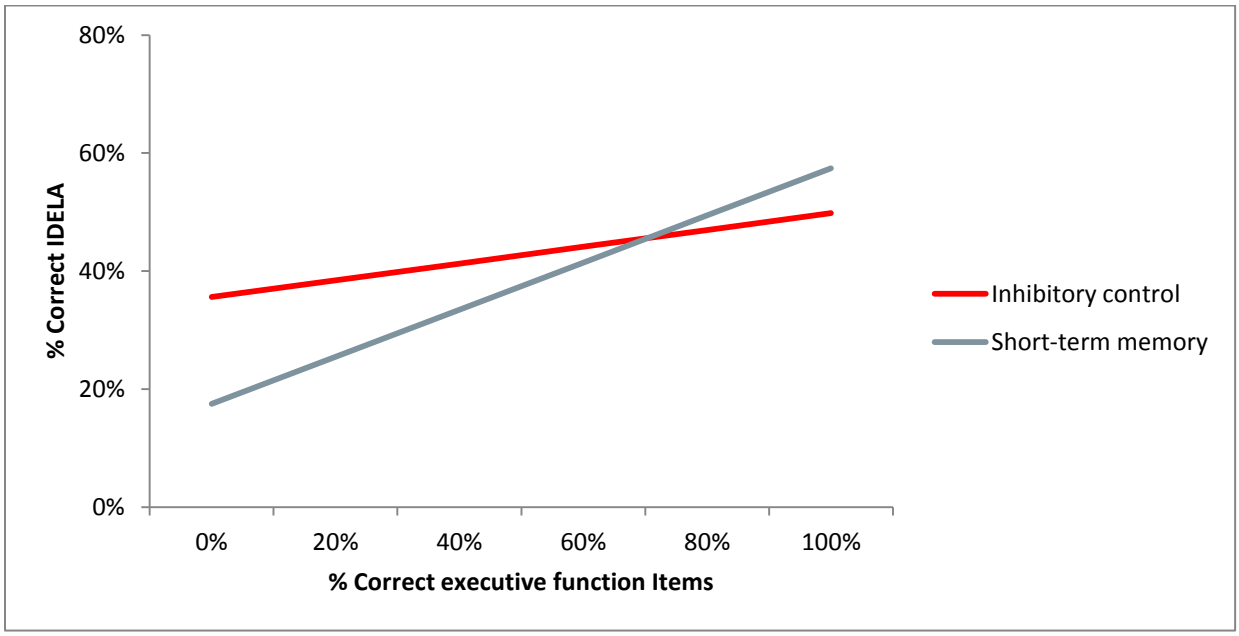


Table B2. Correlation between executive function items and IDELA domains

		Motor	Emergent Literacy	Emergent Numeracy	Socio-emotional	IDELA
Mali	Working memory	0.30	0.38	0.28	0.34	0.40
	Inhibitory control	0.35	0.42	0.39	0.39	0.48
Egypt	Working memory	0.42	0.42	0.48	0.39	0.54
	Inhibitory control	0.35	0.33	0.33	0.26	0.39
Malawi	Inhibitory control	0.26	0.21	0.28	0.24	0.40

[7] The inhibitory control item used in IDELA was developed and is reported on in Cameron Ponitz, C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology, 45*, 605–619.