

Urgency and Possibility Results of PRIDI A First Initiative to Create Regionally Comparative Data on Child Development in Four Latin American Countries Technical Annex

Aimee Verdisco
Santiago Cueto
Jennelle Thompson
Patrice Engle
Oliver Neuschmidt
Sebastian Meyer
Eugenio González
Beatriz Oré
Katelyn Hepworth
Alejandra Miranda

1. Introduction and Background

The Regional Project on Child Development Indicators, or PRIDI, for its acronym in Spanish, was formally launched in December of 2009 by the Inter-American Development Bank. The first initiative of its kind worldwide,¹ PRIDI seeks to generate high quality and regionally comparable data on child development outcomes. This report summarizes the implementation of the project and results it obtained. It also comments on the final results and the potential the PRIDI data hold for future evaluations of early childhood development in Latin America and the Caribbean. All PRIDI products can be found on its website: www.iadb.org/education.

The PRIDI Project

PRIDI was launched in response to the lack of comparable data across countries on child development outcomes. It took its inspiration from the standardized tests implemented across

¹ Although there are a few international studies on children in preschool, no international study generating comparable household data across countries on child development exists.

the Region and internationally (e.g., the Latin American Laboratory for Education Quality, LLECE; the OECD Program for International Student Achievement, PISA; the Third International Math and Science Study, TIMSS; Program from International Reading Literacy, PIRLS, among others) and the impact they have had on informing the policy dialogue on education quality. Indeed, prior to the release of the LLECE data in the late 1999, education quality had little empirical referent in the Region. LLECE provided a big step forward by giving ministers a tool – high quality data – for the monitoring and regional benchmarking of learning. LLECE, its successors and international variants proved important for the policy dialogue its data generated.

A decade later, PRIDI saw ECD as the next frontier. Countries were placing more policy and programmatic emphasis on ECD, but few tools existed for the systematic monitoring and benchmarking of the development of children prior to entering formal schooling. No cross-national, comparable data on child development outcomes existed.

Launched in December of 2009 by the Inter-American Development Bank, PRIDI sought to fill this void (IDB, 2009; Verdisco, 2010, Verdisco et al., 2013). Its objectives are to:

- 1) Generate high quality, population-based and regionally comparable and relevant data on child development in nationally representative samples, and
- 2) Identify gaps in child development between different groups of children.

PRIDI was financed through non-reimbursable technical cooperation. It received aUS\$1 million grant from the Bank's regional technical cooperation modality, US\$140,000 from Fund for Diversity and Gender, US\$200,000 from the Social Fund (which financed the participation of Costa Rica) and US\$240,000 from two separate Economic and Social Work projects.

PRIDI is the first study of its type in the Region and internationally and its data are the first of their kind for any region of the world. PRIDI data allow countries to benchmark progress on child development both within their borders and within the Region, prompting policy dialogue between governments on how best to address the needs of young children and their families.

PRIDI explicitly incorporates issues of equity and inclusion. Three pillars orient all its activities: (i) children develop in an integral manner and the data used to assess their development should capture this; (ii) children should be able to achieve basic developmental milestones and competencies before entering school, independent of their race, gender, socioeconomic background, origin, language, or any other discriminator; and (iii) detailed and nationally

comparative data on child development to inform and guide policies, particularly towards the most disadvantaged children, are largely absent in the Region.

PRIDI, from its start, included indigenous and other marginalized populations, thus differentiating itself from other regional studies (e.g., standardized tests of learning) that exclude, ex-ante, such populations, given the logistical difficulties, including language and cost, of reaching them. With financing from the Bank's Gender and Diversity Fund, PRIDI instrumentation was adapted to the specific contexts of indigenous and rural children (discussed below), thus ensuring, to the extent possible, that all children participating in PRIDI have an equal opportunity to show what they know and can do.

Four countries, each led by a key ministry or ministries with responsibilities for young children, participated in PRIDI: Costa Rica (Ministry of Education), Nicaragua (Ministry of the Family), Paraguay (Ministry of Education), and Peru (Ministry of Women and Social Development, and Ministry of Education).² Each lead ministry, in turn, appointed a national coordinator for the project; in some cases, various persons served at different times as national coordinators:

Costa Rica: Ana Isabel Cerdas González (2009-2013)

Nicaragua: Luis Alemán (2009-2010), Xiomara Bello (2010-2011), Carol Herrera Mejía (2012-2013)

Peru: Amparo Muguruza (2009-2011), Carmen Vásquez Velasco (2011), Vanetty Molinero Nano (2011-2012), Vanessa Sánchez Jiménez (2013)

Paraguay: Graciela Rojas (2009-2011), Lourdes Romei (2011-2012), Nelida López de Lezcano (2013)

The national coordinators were key players throughout PRIDI. They participated in the development of PRIDI instrumentation, ensuring that all items and materials were relevant and “politically acceptable” within their home ministry and within other national ministries with responsibilities for young children, provided important contacts in the national institutes of statistics for the construction of sampling matrices, and supported fieldwork by providing letters of introduction and facilitating other logistics for the consulting firms (discussed below).

² Haiti, Argentina and Ecuador expressed interest in participating in PRIDI, and were present at the time the project was first conceptualized. Given the breath of disaster wrought by Haiti's earthquake in January of 2010, the country dropped out. Argentina also dropped out, given competing internal priorities. Ecuador was involved until the beginning of 2012, at which time it stopped participating, given changes in the Ministry of Economic and Social Inclusion (the lead ministry) and internal priorities.

In addition to the country coordinators, a management team of world-renown experts provided technical assistance to the project, led by Patrice Engle (Calpoly University; UNICEF)³ and Santiago Cueto (GRADE, Peru), and supported by Beatriz Oré (Universidad Antonio Ruiz de Montoya, Peru), Fabiola Lazarte (Universidad Antonio Ruiz de Montoya, Peru) and María Estella Ortiz (Chile). The International Association for Educational Achievement, or IEA, which has supported a number of important international studies on results in education, provided technical assistance on issues related sampling, item validation, and the collection of data. Substantial contributions were made by Oliver Neuschmidt, Sebastian Meyer, Olaf Zuehlke, Eugenio González, Dirk Hastedt, Phamen Mirazchiyski, Andres Sandoval, and Hans Wagemaker. From the side of the Inter-American Development Bank, the project was led by the Education Division: Aimee Verdisco and Jennelle Thompson. A number of research assistants, including Katelyn Hepworth (IDB), Alejandra Miranda (GRADE), Mayli Zapata (GRADE), and Ismael Muñoz (GRADE) provided key assistance. Critical review and comments came from Hugo Ñopo (IDB), Ann Weber (Stanford University), and Lia Fernald (University of California Berkeley).

³ Patrice Engle made immeasurable contributions to ECD worldwide and to PRIDI. Without her technical and operational leadership, PRIDI would not have been possible. She died in September of 2012.

2. Conceptual Framework

Given the lack of antecedent, a first task was to create a Conceptual Framework that clearly defined the objectives, scope and methods of PRIDI in such a way that all resulting concepts were acceptable to and applicable in each participating country. To support this task, the national coordinator for each country convened an inter-sector working group comprised of all ministries with responsibilities for children aged 0 to 6 years. Each working group documented initiatives, past and present, related to the measurement and collection of data on the state of children in this cohort, summarizing the results of such initiatives and any other report that speaks to issues of child development. These reviews clearly revealed that all four countries involved in PRIDI were measuring child development outcomes, and that all had invested considerable time and effort in defining child development.

Building on this information and, where necessary, complementing it with international experiences and literature, the management team began elaborating the Conceptual Framework, which was then reviewed and approved by the participating countries. In the final version of the Conceptual Framework (see <http://www.iadb.org/es/publicaciones/detalle,7101.html?id=31906>), the definition of child development appears as:

An integral process which includes not only verbal skills and knowledge and intellect, but also social skills and motor development, and strategies for learning, such as attention and inhibition of impulsive behaviors, as well as basic notions of health and nutrition. It is the process through which a child is prepared for new levels of responsibility and progressively gains new levels of autonomy.

A number of other definitions follow. As noted above, the objective of PRIDI is to generate high quality data that are comparable across countries. The purpose of these data is not to screen children, but rather to identify gaps in child development between populations and areas, understand the magnitude of these gaps, and plan accordingly to mitigate them.

What PRIDI Measures

The Conceptual Framework recognizes child development as a holistic and integrated process that encompasses any number of domains: cognitive, emotional, health, social, motor, executive functioning, etc. Yet, given operational considerations, most importantly the need for simplicity and affordability, PRIDI could not measure all domains. Based on a detailed review of the literature, the expert opinion of PRIDI's management team, and the policy priorities of the

participating countries, four domains of child development were chosen to be included in PRIDI: cognition, communication and language, socio-emotional, and motor.⁴ Each finds theoretical justification in the Conceptual Framework, and is summarized in Table I.

Table I: Domains of Child Development Included in PRIDI

Domain	Definition	Justification
Cognition	Ability to solve problems, including abilities to categorize, sequence, pay attention, recognize relationships between numbers and relationships between parts and whole, and of executive functioning.	Basic abilities for learning in school. Associated with learning, test scores and later successes in life.
Language and Communication	Development of expressive and receptive language. Expressive language relates to the child's ability of articulate words and concepts. Receptive language relates to a child's comprehension of language. Relates to knowledge and interest in books and drawings.	Highly correlated with and predictive of learning in school Interest in books is an early learning skill.
Socio-emotional	Social abilities and abilities to confront and adapt to new situations.	Association with a child's ability to adapt to new situations. Has predictive validity.
Motor	Fine and gross motor skills, including coordination.	Through their motor skills, children experience new things. Motor skills are related to learning and to cognition.

PRIDI also places considerable emphasis on capturing and understanding the factors associated with ECD. Child development emerges from and is affected by the interaction of a number of contextual variables from the home, community, and parents. A wide breath of literature speaks to the impact these and other factors have on child development. Table 2 summarizes the associated factors included in PRIDI:

Table 2: Associated Factors Included in PRIDI

Factor	Definition	Justification
Child characteristics	Birth date, sex, maternal language, birth-order	Immutable characteristics of the child

⁴ The Conceptual Framework includes “emerging academic skills” as a domain to be included in PRIDI. The idea was that this domain would be constituted a-priori with items from the other domains. However, based on field results from phase II, it was difficult to identify such items, as they largely fell into the cognition or language domains. A decision was taken to drop the “emerging academic skills” domain from PRIDI.

Home characteristics	Socio-economic status of the home, presence of both parents in home, maternal language and education level of parents, access to basic services in the home, number of siblings, parental interaction with child, language-rich activities and materials in the home, child rearing strategies	Environment in which a child develops and grows. Socio-economic status strongly correlated with ECD. Poverty poses a serious risk to ECD and tends to occur concomitantly with other factors that detrimentally affect it, including inadequate nutrition, poor sanitation and hygiene, poor maternal education, and inadequate stimulation in the home.
Community characteristics	Urban or rural, distance from health post or hospital, availability of basic services	Services available to the family and child
ECD or early education programs	Participation and duration of child in such programs. Included here are ECD, early education, nutrition, and conditional cash transfers	Participation in these programs, if they are of quality, is likely to have a positive influence on child development

The Children Included in PRIDI

PRIDI had to make a decision regarding the age range of children. Although the early childhood period spans from year zero through eight, only children between 2 years and 4 years 11 months are included in PRIDI. This decision reflected the operational considerations noted above. Evaluations of children under the age of two tend to be less stable than evaluations of older children, and reliably evaluating young children (less than two years) requires highly trained enumerators (e.g., trained psychologists), which would have had cost implications beyond what the PRIDI budget could accommodate and likely beyond what country budgets could accommodate in any future iteration of PRIDI. Children five years of age and older, with high probability, have already entered formal schooling in most of the participating countries, thus posing analytical complications, such as controlling for the effect of schooling.

PRIDI evaluates children in their home. In this way, PRIDI considers all children in the 2 year to 4 year 11 month age range, and not just those in organized child care centers or schools. As will be discussed throughout this Annex, PRIDI's final data are nationally representative and include about 2,000 children in each country.⁵

PRIDI includes indigenous children in those countries sizable with indigenous populations (Nicaragua, the Atlantic Coast; Paraguay, Guaraní children in rural areas; Peru, Quechua speaking children in the Cusco area). In an effort to reduce biases inherent in instrumentation

⁵ It should be noted that original calculations called for nationally representative samples of 3,000 children in each country. This calculation was based on three age cohorts: 2, 3 and 4 year-olds. In November of 2011, IEA revised the sample size to 2,000, on the assumption there were only two cohorts, as determined by the scales: 24 to 41 months, and 42 to 59 months.

and its application, thus ensuring that all children participating in PRIDI have as equal a chance as possible to show what they know and are capable of doing, considerable time and effort were invested in the adaptation of all PRIDI instruments to the peculiarities of each participating population (detailed below).

These are important considerations that set PRIDI apart from other studies on ECD and preschool in the Region. Few evaluate development outcomes in young children in nationally representative samples. None evaluates children in the home or specifically adapts its instruments to indigenous populations.

3. Creating PRIDI Instrumentation

Instrumentation was a key concern in PRIDI, as the instruments bear a close relationship to what is measured and how it is measured. A potential avenue for PRIDI would have been to choose a test currently in use in the Region or elsewhere, adapt it and apply it. This course of action was rejected both by the countries and PRIDI's management team. From the management side, there was some anecdotal evidence that some scales and evaluations currently in use (e.g., the use black and white pictures) were biased against indigenous and rural children and underestimated their abilities. For example, a widely applied test based on black and white drawings includes a hot air balloon and an ornamental lamp in its opening sequences – objects that are not likely to be recognized by outside of urban areas.

Application of instrumentation matters as well. How an item is applied bears a direct relation to what a child or mother/caregiver does or how she reacts to a given prompt. For example, the application of an observational scale of development asks the mother how her child reacts when he sees his image in a mirror. Although this is a common practice in many parts of the world, in some indigenous communities the practice is frowned upon, given prevailing beliefs on the fate of the child's soul. The same competency – recognition of self – can be assessed in other, more culturally acceptable ways, such as use of personal pronouns.

From the national reviews (current as of 2010), it became evident that the participating countries were measuring child development outcomes and, in some cases, were using the same tools to do so (e.g., Escala Abreviada de Nelson Ortiz, Escala de Desarrollo Integral del Niño; see Conceptual Framework; and Annex 1). Many of these scales, however, were outdated and thus not informed by current advances in the field of early childhood development and brain research. The decision reached by the countries and PRIDI team was to develop a new test that built from what the countries were already doing, and to complement these initiatives with items found in more recent, international tests.

A three-phase plan was laid out for elaborating, validating and applying the instruments:

- 1) A formative phase (Phase I), in which the newly elaborated instruments and respective materials, manuals and forms would be piloted in small samples of children in two countries and adapted to different populations.
- 2) A validation phase (Phase II), in which the instruments, adapted through the formative experiences of phase I, would be applied in limited samples (200 children) in all four participating countries, and validated against two internationally normed tests: the Peabody Picture Test (in its Spanish version, Test de Vocabulario en Imágenes Peabody, TVIP) , a normed reference test for measuring receptive vocabulary in children that has

been applied in various countries in the Region; and height-for-age, an internationally used anthropomorphic test to measure physiological growth.

- 3) A phase of national application (Phase III), in which PRIDI instruments together with the TVIP and height-for-age would be applied in nationally representative populations (about 2,000 children) in each participating country.

The process of elaborating the PRIDI instrumentation was long and required attention to the smallest details. Theoretical considerations, particularly that of ECD being an integral child-centered process, had to be balanced with more operational issues, including budgetary limitations and country capacities for validating new instrumentation and applying it in nationally representative samples. The PRIDI team, in coordination with the national coordinators of each country, dedicated over a year to the design and validation of an initial set of instruments. Five general principles guided their actions:

- 1) Measure skills and abilities present prior to school entry and that can predict academic achievement at a population level (not individual, not diagnostic)
- 2) Use various indicators or scales; no global or composite indicator and no rankings
- 3) Define a series of items that capture PRIDI's dimensions in children from 24 to 59 months
- 4) Define items and concepts that are relevant to Latin America and are applicable to a wide range of socio-economic and cultural (e.g., indigenous) groups, with a minimum of inputs (e.g., prompts, toys, materials)
- 5) No professional training required for the application of any instrumentation; training for application will be short and application will occur within the home

The first step was an analytical exercise mapping theoretical constructs of PRIDI's intended developmental domains and associated factors against discrete test/survey items designed to capture such data. This exercise systematically defined each domain, justified its relevance for child development and PRIDI, and then amassed a number of items taken from a number of tests as possible candidates to measure the given domain. Items were taken and adapted from tests and surveys being applied in the participating countries (e.g., Nelson Ortiz, EDIN, TEPSI) and internationally (Bayley's, Woodcock-Muñoz, Denver, MICs) as of 2010 (see Annex 2). A similar process was followed for creating the associated factor surveys (see Annex 3, which provides the final mapping used in PRIDI).

The leap from this exercise to a scale or survey was not straightforward. In child development, age-appropriateness matters and has serious implications for how a given domain is assessed. A wide range of developmental outcomes occur between the ages of 2 years and 4 years and 11 months. Issues of how best to evaluate the domains included in PRIDI also entered into discussion. For example, whereas observation may be an appropriate means for evaluating

some competencies (e.g., language) in older children, it may not be for younger children. In this case, evaluating the same competency in a younger child may require a different method, such as mother response, for which a long line of literature finds validity (see Conceptual Framework).

With these considerations in mind, the PRIDI team met in Asunción, Paraguay in December of 2010 and January of 2011 with the objective of creating the preliminary instrumentation. The result of these efforts was the creation of the Engle Scale for Child Development (Escala Engle de Desarrollo Infantil), named in honor and recognition of Patrice Engle. Based on direct observation of the child, and designed to be easily administered by people with some knowledge of ECD and short, hands-on training, it measures cognition, language and communication, and motor skills. Given the large range of development which occurs in children from 24 to 59 months, the Scale consisted of two evaluations, referred to as Forms: one for children from the age of 24 to 41 months (Form A), and the other, more difficult, for children 42 to 59 months (Form B).

The materials necessary for the application of the Scale were of common-usage in the countries and within their various populations, of minimal cost, durable and easily transported in a backpack, and could be used for various items and in both scales: e.g., small wooden blocks, a puppet, pencils, drawing, and a ball. A detailed manual was created that included figures and instructions on the correct use the materials and correct application of each item.

In addition, PRIDI incorporated the use of mother/caregiver report for the evaluation of socio-emotional development, as well as a means for capturing factors associated with the home and child. To this end, two surveys were developed. The first instrument (survey of the family) inquired into the basic characteristics of the household and the environment within which the child was developing, including socioeconomic factors. The second (survey of the child) inquired into factors associated with the child, and included items related to the socio-emotional state of the child⁶, attendance in preschool or organized care, overall health condition, and disciplinary methods used in the home, among others. The ECD module of UNICEF's MICs4 questionnaire was distributed between the two surveys.⁷

⁶ Socio-emotional development is included in the Engle Scale. The only section of the Engle Scale evaluated via maternal or caregiver report is socio-emotional development.

⁷ As stipulated in a Memorandum of Understanding between the Bank and UNICEF in 2011, PRIDI would include the MICs4 module on ECD.

The PRIDI team also developed a number of forms to accompany the application of these instruments, including a consent form to be signed by the mother or caregiver prior to the application of the PRIDI tests, a registry of households to guide the sampling process in the field, and a feedback form to be filled out by the enumerator upon completing the application of all instruments in a given household to note any issues or concerns.

As with the Conceptual Framework, each item on each instrument and form required approval by the participating countries. Initial consent on the instruments described above was given by the countries in August of 2011. Consent at this time was conceptual, as the instruments had not yet been field tested. The basic structure of the Engle Scale at this point in time largely reflected the mapping illustrated in Annex 2, albeit with the addition of motor items (e.g., walk a straight line, jump on two feet or one foot, kick/catch a ball); the basic structure of the surveys followed the mapping found in Annex 3.

4. Phase I. Formative: Functionality and Cultural Adaptations

The objective of Phase I was to observe the functionality of the PRIDI instruments when administered to a small sample of children. Two countries were chosen to participate in Phase I based on their interest in, commitment to and ongoing work in early childhood development, and the presence of significant indigenous populations, which provided an opportunity to closely study issues related to instrument adaptation: Peru and Paraguay. The PRIDI team asked each country to review the Engle Scale and the two surveys and to adapt them to their national context. Each country was permitted to include additional items it thought relevant to its national context and to make any modification in the items necessary to accommodate each “population.” However, they were instructed not to eliminate any item from the Scale or the surveys. The preference of the PRIDI team at this point in time was to have longer, rather than shorter, instruments (Scale and surveys). Experiences gained in the field over phase I and II, and the analyses of the results, would be used to identify specific items within the Engle Scale and surveys that best captured the desired information and performed as expected in the field. No item would be eliminated from the Scale or a survey without being tested in the field.

From this process, two Engle Scales, each with two age-specific Forms, were created: a Paraguayan Scale and a Peruvian Scale. Each country also created a manual to support the application of its respective Scale.

In Paraguay, the resulting evaluation for children 24 to 41 months (referred to as Form 23, with “23” being a reference to children aged 2 – 3 years) included 105 items along the domains included in PRIDI, each with its own point scale. Form 34 (children 42 to 59 months, or 3 – 4 years, hence the reference to “34”) included 99 items, each with its own point scale. In the Peruvian case, Form A, analogous to Form 23 in Paraguay, included 33 items, each with its own point scale. Form B, the more difficult scale for the older children, included 36 items, each with its own point scale.

Differences between each country’s Forms were not as great as the figures in the preceding paragraph may suggest. In the Peruvian case, a number of items often were subsumed as sub-items under a single item. For example, in Form A (Peru), question 4 asks a child to identify different colored cubes: red, yellow and blue. In Form 23 (Paraguay), three separate questions (35-37) are used to evaluate the same competencies. Similarly, in Form B (Peru) a single question (6) asks the child to point to different geometric shapes: a square, triangle and circle. Three separate questions (30-32) on Form 34 (Paraguay) ask the child to do the same.

In general, each item on the Peruvian Forms was included on the Paraguayan Forms. The Paraguayan Forms included a number of additional items (e.g., riddles, like “what has four legs and is used for sitting?” on Form 23, and items related to musicality on Form 34, “differentiate between two tones” or “differentiate between short and long tones”) and repeated items at different points (for example, naming or pointing to everyday objects), varying a single characteristic (for example, quantity, in naming/pointing to 3 objects in one series of responses and then to 7 objects in the next series in Form 23, or height, in identifying the tallest block of 2 blocks, and then of 3 blocks in Form 34).

Scoring of the Scale

Among the objectives of this phase was that of acquiring as much information as possible about the performance of PRIDI children. For example, when applying an item, it was important to know not only whether or not a child correctly responded to an item, but also how many tries it took to get the correct response, which prompt elicited the correct response, and how much time was needed to respond to the item. To this end, scoring during the phase was formatted in such a way to capture this information. In subsequent phases, the formatting was changed and made simpler: enumerators indicated whether a child responded correctly or not (see below).

Cultural Adaptations

Both countries culturally adapted their respective version of the Engle Scale. The process of cultural adaptation was understood as a process of adapting each item in such a way that each population – urban, rural and indigenous – understands it in the same way but that the competency to be evaluated remains the same across all populations. The process applied during Phase I included various steps to more fully understand the contexts within which PRIDI would be applied; the majority of this work took place prior to translation. Each country team met with representatives from units responsible for overseeing indigenous issues in the respective ministry leading PRIDI. They then collected qualitative information through observation on daily routines of children aged 2 to 6 in indigenous areas and met with education professionals working with kids in rural areas (Paraguay), and in Cusco (Peru) to have a clear understanding of the context within which they were working. Each team also included members who spoke Guarani and Kohn (Paraguay) and Quechua (Peru).

Information gained through these activities provided insight into the conditions and contexts in which PRIDI instrumentation would be applied. For example, distances between homes in rural and indigenous areas in both Paraguay and Peru were long and access was hard. Enumerators had to walk considerable distances and four-by-four vehicles were necessary to get the

communities. This called attention to the need to limit the number and weight of inputs needed to apply the Engle Scale and to use study backpacks to transport everything.

In addition, few homes had ideal places for the application of the Scale. Most homes had a single area that served multiple functions (e.g., sleeping and cooking; many had animals inside this area) and tended to be dark (not all had electricity; few received sufficient sunlight). The following photos are from the formative work in Andahuaylillas, in the Peruvian Altiplano.



Based on these observations, PRIDI would likely need to be applied on a bed or the patio in rural areas, with all the consequences this implies, such as interference from animals and other family members (children and adults), hot sun, etc. A number of recommendations about how to ensure successful application within such contexts thus were generated, including that enumerators carry a blanket to sit on if the Scale had to be applied on the floor, that they have a flat object with them, like a clipboard, for those items prompting kids to use blocks, if the floor or bed was uneven, etc. Again, as the following photos (from Lluto and Andahuaylillas, both in the Peruvian Altiplano, and San Mateo, a rural community outside of Lima) illustrate, the difficulties of fieldwork, particularly in rural and indigenous areas.



The Scale was then translated into Quechua (of Cusco) in the case of Peru. In this phase, Paraguay chose not to translate the Engle Scale into Guaraní.⁸ The process of translation

⁸ In Phases I and II, enumerators translated the items into Guaraní in the field as they applied the instruments. This approach generated a considerable lack of consistency in the terms and expressions used, and raised a number of

included translation of the scales from Spanish into the respective indigenous language, and from this language back to Spanish. Where there were differences, the translators met to determine the final wording to be used in the indigenous version.

Virtually no variation existed between Paraguay and Peru in the surveys. What little variation did exist reflected national adaptation. For example, materials used for floor construction in Paraguay included bamboo, but rock or brick in Peru. Each country included its own list of ECD programs, as was necessary. Neither Paraguay nor Peru translated the household and child surveys into an indigenous language.

Using the manuals and the instruments, the PRIDI team trained teams of enumerators in both countries in the application of the Engle Scale and surveys, prior to going to the field. In Paraguay, Guaraní speakers were included in this team; Quechua speakers from Cusco were included in the Peruvian team.

The countries were also asked to create a small, non-randomized sample of children that “over-sampled” indigenous children and was stratified to include children in the PRIDI age cohort in urban, rural and indigenous areas. In Paraguay, the sample consisted of 56 children (23 children for Form 23 and 23 children for Form 34); in Peru, of 41 children (23 children for Form A and 18 children for Form B).

Once the sample communities were identified, field teams were accompanied by Ministry staff familiar with each community (Ministry staff with responsibilities for non-formal early education from the Early Education and the Indigenous in Paraguay, and madres comunitarias from the WawaWasi Program in Peru). These professionals helped bridge communication between the enumerators and mothers asked to participate in Phase I. In some cases, particularly in indigenous communities in Paraguay, these professionals were key in securing the consent of community elders, who then “authorized” mothers to participate.

Phase I: Instrumentation

Instruments applied during Phase I included the following:

- Engle Scale of Child Development. Directly applied to children to evaluate their cognitive, language and motor development, with Form 23 in Paraguay and Form A in

concerns within the management team. Guaraní versions of the PRIDI instrumentation were not used in the field until Phase III.

Peru applied to children aged 24 to 41 months, and Form 34 in Paraguay and Form B in Peru applied to children aged 42 to 59 months.

- Family Survey. Mother/caregiver report to collect information associated with the home and environment within which the child is developing, including socio-economic factors.
- Child Survey. Mother/caregiver report to collect information on factors associated with the child, including the child's socio-emotional development, basic health, participation in ECD or early education programs, and different methods of discipline applied in the home.
- Informed Consent Form. Signed by the mother or caregiver prior to applying PRIDI instrumentation.
- Registry of home visits. To guide the process of sampling in the field and keep track of the homes visited.
- Preliminary interview. To guide the selection of children to participate in PRIDI.
- Feedback Form. Filled out by the enumerators after completing the application of all PRIDI instrumentation with the purpose of registering any difficulty or issue encountered.

The final breakdown of Scales applied during Phase I follows. The surveys and supporting documentation were applied in all cases (see Tables 3 and 4).

Table 3: Application of Forms 23 and 43

Paraguay				
	Form 23		Form 34	
	Female	Male	Female	Male
Urban	3	2	1	1
Rural	11	7	11	8
Indigenous	3	2	2	5
Total (56)				

Table 4: Application of Forms A and B

Peru				
	Form A		Form B	
	Female	Male	Female	Male
Urban	6	1	3	3
Rural	6	7	4	5
Indigenous	2	1	1	2
Total (41)				

Results: Phase I

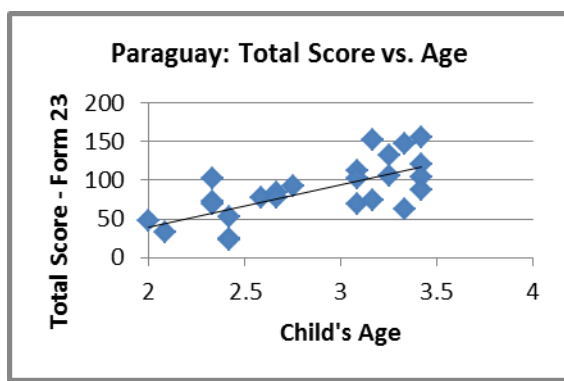
Results were analyzed by the PRIDI team, including IEA, in collaboration with the local consultants. The small sample size limited the types of analyses that could be undertaken, and raised the need for extreme caution in interpreting results. For example, correlations seen in these data cannot be generalized. In addition, the results of the Engle scale are analyzed as the sum of all items across all domains, rather than by sub-scale, a larger sample size, such as that

included in Phase II, is necessary to empirically confirm each domain. In general, the results of this phase were encouraging.

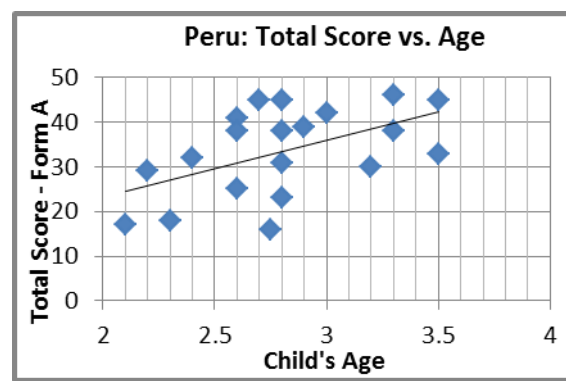
Results of the Engle Scale: Form 23 and Form A

Results from Form A indicated that the items were functioning as expected based on the literature and other tests. Scores were evenly distributed and strongly correlated with age of the child, thus creating a base and ceiling for the scale (see Graphs 1 and 2).

Graph 1: Form 23 by Age, Paraguay



Graph 2: Form A by Age, Peru



Correlations were less strong and uniform when discriminated by household education (not shown), a proxy for socio-economic status. In the case of Paraguay, a strong correlation emerged between each child's score and the average level of education of all adults living in the household. This, too, was to be expected. Yet the same was not found in Peru. Although a concern, the lack of a clear correlation in Peru was likely explained by the small sample size and the oversampling in indigenous areas which reduced variation in the levels of education of the mothers/caretakers within the sample.

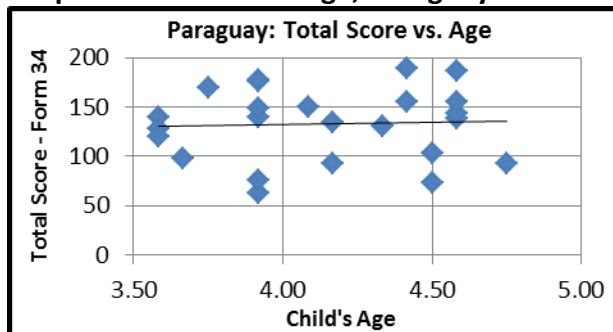
In general, the scores of indigenous children were considerably lower than scores of urban children or rural children. In Paraguay, Form 23 produced dramatically lower scores from children in rural areas and indigenous populations. This outcome was consistent with expectations: given a lack of educational resources, children living in rural areas and indigenous populations tend to score lower than children from urban areas. The magnitude of this disparity was dramatic. Although the small sample size limited any firm conclusion, such a magnitude suggested the need to ensure that the instruments were appropriate to indigenous children and not under-estimating their abilities. For example, field reports from the enumerators indicated that in indigenous populations, some children did not know what "shoulders" were, when asked to point to different body parts, but could identify all others when prompted. No

such disparities were observed in the case of Peru (Pachacámac is an urban area; Huarochirí, a rural province; and Cusco, which has a large indigenous population).

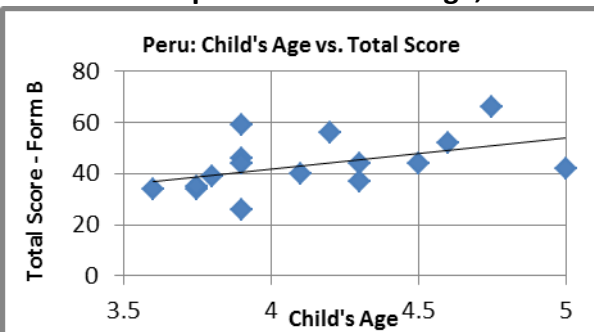
Results of the Engle Scale: Form 34 and Form B

Results from the scale created for older children differed in some important ways from those emerging from Form 23/A. Foremost among these differences was the lack of variation in the distribution of scores. On both Form 34/B, most all children completed all tasks successfully thus suggesting that there were too few difficult tasks and making it impossible to create a score ceiling. This issue was highlighted by a low correlation between a child's age and the score on the assessment; this situation was particularly noteworthy in the case of Paraguay where the correlation between age and scores was essentially flat (Graph 3). More of gradient exists for Peru (Graph 4).

Graph 3: Form 34 and Age, Paraguay



Graph 4: Form B and Age, Peru



As in Form 23/A, correlations with household education were weak (likely reflecting low variation in this indicator) and little difference emerged between children in urban and rural areas. Again, these results were likely pointing to a need for more difficult items to be included on this scale and the need to significantly revise Form 34/B.

Results of the Surveys for the Child and Household

In general, the surveys for the child and household functioned largely as anticipated in both countries. In the survey of the child, no real difference emerged between Paraguay and Peru in the items associated with the socio-emotional domain, all of which were posed to the mother or principal caretaker. In both countries, children in rural areas received higher scores than children in urban areas on a number of items (e.g., the child worries if someone is crying; if the child is doing something and makes a mistake, s/he persists and keeps on trying without getting frustrated or angry; the child can play for 15 minutes or more without requiring the attention of an adult).

The accompanying forms, including the consent form, household registry, the preliminary household interview (to determine if a child within the PRIDI cohort was present), functioned as envisioned.

Operational Issues during Phase I

Feedback from enumerators proved critical to the PRIDI team in understanding what worked and what did not in the field. The most often reported issue from their point of view related to the time required for the application of the battery of instruments, particularly the Engle Scale. Although the intention was to slowly and intentionally apply the Scale, thus allowing for the close observation of its functioning with different groups of children, the time required appeared to be far too long for the age-group at hand. Application tended to run well over an hour, often approaching two hours. Many children seemed to be fatigued by the end of the test or got bored during it, which influenced their performance. Indeed, as is known from the literature and the experience of the PRIDI team, if children lose interest in an assessment, their performance likely will suffer, especially on tasks presented towards the end. In these cases, the child may well be able to do the task correctly but does not due to fatigue or boredom.

The application of the surveys required anywhere from 9 to 55 minutes, time that many mothers or caregivers did not have available, given competing responsibilities for children, food preparation and other chores. In both countries, enumerators reported that it often required 2-3 visits to the same home to apply the complete set of instruments. In a number of cases, this created logistical difficulties, especially in terms of scheduling the re-visit/s around the work obligations of the mother; in some cases, it was impossible to align schedules, and the assessment could not be completed.

In addition, issues identified during the cultural adaptations (see above) came to pass. Paraguay reported difficulties in getting to rural villages, in some instances due to a lack of infrastructure, and in others due to rain and floods (see photos, below). Where roads were unpassable, enumerators had to either go by foot or postpone the planned visit, all of which increased the time and resource commitment necessary to carry out Phase I activities.



Enumerators in both countries often had to walk long distances to find home with children in the PRIDI age range. Once these homes were identified, some mothers or caregivers were reluctant to let enumerators into their home and provide details on their child. If they consented to participate in PRIDI, time was an issue. In a number of cases, mothers or caregivers simply stopped answering questions in order to attend to chores, their children or leave for work, thus creating incomplete cases and requiring the identification of new children and mothers in new homes and the start of the entire application process from zero. This, in turn, generated additional costs in terms of time and travel.

Most evaluations of children in rural and indigenous areas were done on the ground, or another less than ideal space (e.g., floor; see photos below, from Paraguay and Peru, respectively).



The enumerators also reported that, in a number of instances, the application of the scales was complicated by the presence of other kids or the intervention of adults. In large families, in particular, the enumerators often had difficulty getting the child to be evaluated into a situation where s/he could work individually. When parents or other adults did intervene, it seemed that they were doing so to help the child come up with the right response (again, see pictures, from Paraguay and Peru, respectively).



Kids were also interested in the whole kit of materials, and insisted in knowing “what else” the enumerator had with her. To deal with this, enumerators often had to reveal objects before using them in the evaluation.

The formative work also served to inform the PRIDI team about the appropriateness of the materials used for applying the development scales. For example, some of the drawings used to describe sequences on both scales appeared to confuse children and not all children could name items that were assumed to be “daily use” items across the board. In this last case, it was impossible to know whether the lack of identification stemmed from a lack of vocabulary or from simply not knowing what the object was.

Enumerators in both countries also reported difficulties in applying the field protocol for the Engle Scale and for scoring items. For example, instructions on how to score correct versus incorrect answers were confusing and led to significant levels error in scoring such items. Other problems were encountered included the calculation of the age of the child to be evaluated and the determination of when to repeat instructions if a child did not respond to a given prompt. In the surveys, difficulties arose in both countries with questions related to the mother’s health. Various terms, for example “depression”, were not readily understood by the mothers interviewed and generating considerable inconsistencies and delays in application as enumerators attempted to describe in their terms the intended meaning.

These issues were highlighted and included in the revised manuals and in the training of enumerators in Phase II (see below).

5. Phase II: Validation

On the basis of results from Phase I, a number of revisions and adaptations to the instruments were made. Foremost among these was an almost complete revision of Form 43/B. The PRIDI team modified the items to increase the level of difficulty of the scale. For example, in Phase I, a child was asked to count to five, in the revised version the task is to count as high as 20. In Phase I, a child was asked to differentiate between above and below, and in front of and behind; in the revised version, s/he was also asked to differentiate between right and left. In addition, the Paraguayan and Peruvian version of each evaluation (Form A/23 and Form B/43) were consolidated into a Form A and a Form B. These had 32 and 39 items, respectively, although a number of questions included sub-items (e.g., when asked to place a pencil above and below a plate; above and below are counted as separate sub-items). Confusing and non-functioning items (e.g., head and toes, an item designed to evaluate execute functioning skills and included in all Forms, in which the child was asked to touch her head when prompted to touch her toes and vice-versa) were eliminated.

The revised scales were applied to a sample of 12 children in Cusco, and performed much as was to be expected. Form B displayed much more variation in results, and the average application time for scales and surveys (combined) was reduced to about 40-60 minutes. With these results, the PRIDI team felt confident moving forward with these revised scales and surveys. Consulting firms were hired in each of the participating countries via competitive international bid to undertake the field work, including sampling, data collection and digitization, and the recruitment and training of enumerators:

Costa Rica: Leyden Consulting Firm
Nicaragua: CIASES and Gesaworld
Paraguay: Universidad Iberoamericana
Peru: SASE Consultores

From this point forward, a division of tasks emerged: conceptual approval at the ministerial level, and fieldwork by a consulting firm. It was thus indispensable that each firm maintain a close relationship with its respective leading ministry and national coordinator. The national coordinators facilitated contacts with their national statistical institutes for issues related to sampling (discussed below) and provided the firms with letters indicating that PRIDI was a national study supported by the respective ministry. Such actions were of critical importance for ensuring the implementation of Phase II and III of PRIDI.

In a meeting of national coordinators, lead investigators from each firm, and the PRIDI team in November 2011 in Lima, Peru, the countries approved, with minor changes, the Scale and surveys (post-Cusco) for application in Phase II. The list of materials needed to apply the Scale was also finalized and countries were given guidelines for the purchase of materials. In addition, IEA reformatted each instrument based on the feedback received from the formative work, and developed a software for the entry and management of data that uses consistent

coding across the scales and surveys. The PRIDI team provided training on the application of the Peabody test and height-for-age, both of which would be applied in Phase II to validate the PRIDI instruments. IEA returned to Lima in January 2012 to give a training seminar on the use of the software to the data managers of each form, and to prepare a manual to guide the use of this software (see Data Management Manual).

Field Logistics

The PRIDI team finalized the instruments and distributed them to the countries and firms at the end of January 2012. They also finalized and distributed a series of materials designed to help ensure that the application of all PRIDI instrumentation was as uniform as possible across the four participating countries. Included here was a manual for the application of the Scale (Form A and Form B), a DVD demonstrating the correct application of each item of the Scale, a manual and detailed curriculum for the training of enumerators in PRIDI instrumentation (discussed below), the TVIP and height-for-age, and a manual for the application of the surveys.

The manuals for the application of the Scale and surveys directly addressed issues raised in Phase I. For example, concrete details were given on how to calculate the age of the child participating in PRIDI, the number of times a given item could be repeated, how to deal with other children and adults who intervene in the application of the instruments, and how to determine the adequacy of the physical space for the application of the Scale. Point scales were made uniform across forms, and the format of the Scale and surveys was edited to allow for easier scoring and registry.

In addition, a number of innovations were introduced to the Scale. In both Forms A/23 and B/34 one of the first items prompted children to give the enumerator an exact number of cubes (2 and 1 in Forms A/23 and 4 and 3 in Forms B/43). In the consolidated Forms A and B used for Phase II, a puppet of a dog named Rocky was introduced. The competency to be evaluated – exact quantity – remained the same, but the manner in which the relevant items were applied changed along the following lines. The enumerator introduced children to Rocky and asked them to pet him. The enumerator then indicated that Rocky was hungry and that he wanted to eat. Children were also to give Rocky the indicated number of cubes, which represented dog food. The use of the puppet put children more quickly at ease with the enumerator and proved to be more fun for the children.

Building on the experiences gained during Phase I, IEA developed a protocol to account for idiomatic and linguistic differences between countries. The Form of National Adaptations (Formulario de Adaptaciones Nacionales, FAN) instructed countries to adapt the language used

in the distributed – international/Peruvian⁹ - instruments to national and sub-national contexts, including to indigenous populations. The international character of PRIDI had to be respected. PRIDI evaluates the same competencies in all participating countries, and all countries apply the same instruments. No change in the substance or intent of the item thus was permitted. Examples of the national adaptations made include the use of “vos”, as opposed to “tu” in Paraguay and different words to refer to the same object (e.g., “media” versus “calcetín” in Nicaragua; “barriga” versus “panza” in Paraguay; “vela” versus “candela” in Costa Rica, and “cordón” versus “pita” in Peru). Annex 4 provides additional examples of the national adaptations made in Forms A and B. Corresponding changes were made in the manuals as well.

In terms of the necessary cultural adaptation, the methodology followed in Phase I was replicated in Nicaragua and Peru. All instrumentation was adapted to the respective indigenous populations and translated into Miskitu (Region Autonoma del Atlantico Norte) and Quechua (Cusco). In Paraguay, field teams included Guaraní speakers who, as in Phase I, translated the items into Guaraní as needed. A Guaraní version of the PRIDI instrumentation was not used in the field until Phase III.

The scoring of the Scale was simplified during Phase II. Items were formatted in such a way that enumerators indicated whether the child responded correctly or not; no additional information was collected as was done in Phase I. In most cases, a child received a score of 0 when she did not answer correctly and a score of 1 when she did. In a few cases, scores included 2s and 3. For example, the “memory of words” item prompted the child to repeat a series of words, and was scored up to 2, in the following way:

Consigna		Puntuación					
11.1	Serie de 2 palabras	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 0	<input type="checkbox"/> 88	<input type="checkbox"/> 99	
11.2	Serie de 3 palabras	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 0	<input type="checkbox"/> 88	<input type="checkbox"/> 99	
11.3	Serie de 4 palabras	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 0	<input type="checkbox"/> 88	<input type="checkbox"/> 99	

Similarly, the “animal game” item, which was applied with a series of 12 cards, carried a maximum score of 3. In scoring this item, the enumerator calculated the number of correct answers and the number of errors:

⁹ Given that the lead consultants working on PRIDI and responsible for preparing the instruments for phase II and III were Peruvian, the international version of PRIDI was Peruvian.

N° de ACIERTOS ____

N° de ERRORES ____

Acierto: el niño coloca la mano cada vez que aparece la vaca.

Error: se considera error cada vez que aparece una vaca y el niño no coloca la mano o cada vez que aparece cualquier otro animal (perro o gallina) y el niño coloca la mano.

15. Puntuación:

4 aciertos y ningún error	3
4 aciertos y hasta 2 errores	2
3 aciertos y hasta 3 errores	1
2 ó menos aciertos	0
El niño se rehusó a realizar el ítem...	88
El examinador aplicó mal el ítem	99

In instances where the child did not respond to an item, enumerators scored the item in the following manner: 88 in cases where the child refused to answer (did nothing or did not speak) and 99 in cases where the item was not applied or there was an error in application.

Training of Enumerators

PRIDI was structured around the idea that the application of its instruments would not require highly specialized or trained personnel. The profile of enumerator that the PRIDI team looked for was for people with an interest in child development that were available for specific, project-based training. This profile was piloted in Phase II. The firms were asked to recruit women (under the assumption that young children and their mothers would respond best to women and respecting prevailing norms regarding visitors entering a home with women and children present), preferably university students in fields related to child development, and to recruit more than needed in order to have a pool of applicants from which to choose the best candidates.

Profiles of the enumerators varied a bit by country. In Paraguay, the university team partnered with the Department of Early Education of the Ministry of Education. The Ministry donated the time and per diem of a number of its technical staff, who were trained by the university team in

the application of PRIDI instrumentation and its supervision. Each enumerator had to be certified or in the process of being certified in ECD, psychology, education or nursing; all were women with the exception of one. Similarly, in Nicaragua, the Ministry donated professionals who were trained in PRIDI instrumentation and who supported field supervision. Enumerators were all female with the exception of one. In Costa Rica, enumerators were university students studying psychology (the majority) or social work; all were women. In Peru, criteria used to recruit enumerators included prior field experience evaluating small children and/or knowledge of ECD and evaluation of ECD. Again, all were women.

The PRIDI team generated a detailed training curriculum on the application of the PRIDI instrumentation, the TVIP and height-for-age to be implemented over the course of 10 days. The training curriculum describes PRIDI project, its objectives and the importance of ECD internationally. It provides precise instructions on how to apply each item of the Engle Scale (cross-referenced with the DVD and manuals), the surveys and the two tests for validation. It also includes classroom-based and child-based (e.g., in an ECD center, to practice the application of the scale on children) activities centered around the application of the Scale and the proper scoring of each item, as well as tests to ensure that candidates achieved the expected level of mastery in application. Role playing was used to practice applying the Engle Scale and the TVIP. Participants were filmed in order to allow for feedback and group evaluation, processes that were necessary to achieve the standardization of processes. Each recruit had to successfully complete six applications of Form A and Form B in practice sessions with children (in ECD centers) before being “qualified” for field work.

Key staff from each of the consulting firms were trained by the PRIDI team. These staff then replicated the training with their recruits. The training curriculum recommended that each firm recruit 10% more candidates than would be needed to carry out activities in Phase II. This 10% proved necessary. Each country reported some drop-out in their initial pools of recruits and some recruits were unable to successfully complete the course of training.

Sampling

Issues of sampling weighed heavily in PRIDI, and participating countries were required to collect all data using high quality standards. The sample determined the extent to which the collected data yield unbiased, representative, estimates of population characteristics and enable meaningful comparisons of children between and within countries.

A major contribution of IEA in PRIDI was the technical assistance it provided to countries and the management team in issues related to sampling. The sampling manual it elaborated for PRIDI calls for randomly selected nationally representative samples in each country for Phase III

created via a three-staged strategy (discussed in further detail in Chapter 8). In the first stage, a sample of primary sampling units (PSUs) was drawn.¹⁰ Then, a sample of secondary sampling units (SSUs) – households was selected within each of the sampled PSUs in a second stage.¹¹ Thirdly, and in an effort to prevent clustering, only one child per household could participate in PRIDI. In homes where more than one child fell within the 24 to 59-month age cohort, a rule based on the month and day of each child’s birthday was applied to randomly choose one child. In a similar vein, in instances where fieldworkers did not find anyone at an identified home, they were instructed to visit this same home up to three times before removing it from the sample.

The sample used in Phase II was small (200 children in each country) and non-random. However, every effort was made to draw these samples in the same manner as would be required in Phase III, and to test all necessary procedures and protocols. All four countries stratified by urban-rural (as defined by each country). In addition, Peru and Nicaragua included indigenous children (see below). Costa Rica does not have a sufficiently large indigenous population to merit such stratification, and in Paraguay, the rural-urban stratification was deemed to be sufficient to capture the indigenous population. Details of these samples follow in Table 5:

Table 5: Characteristics of the Phase II Sample

	Achieved Sample, Phase II						Total Homes	
	Urban	Rural	Indigenous	Form A	Form B	Total	Total # Visited	% Accepted
Costa Rica	194	0	0	69	125	194	3390	6%
Nicaragua	135	120	52	140	115	255	488	52%
Paraguay	114	77	18	100	91	191	224	85%
Peru	121	50	52	115	108	223	790	28%

As can be noted from the table above, a serious issue faced in all countries was the identification of homes with children in the PRIDI age cohort. For example, in Costa Rica, it was necessary to visit about 17 homes to get a single participant in PRIDI, and all of these were in urban areas. The Costa Rican team visited 313 homes in rural areas, but did not manage to incorporate a single child in PRIDI. This situation had obvious cost and time consequences for the project, elevating both. Among the reasons cited for not participating in PRIDI were: no

¹⁰ The first stage of sampling (PSU) consists of geographical areas or administrative divisions. PSUs must be non-overlapping and cover the entire area of a country (except for any excluded areas).

¹¹ The second sampling stage (SSU) is performed within the selected PSU. In the case of PRIDI, the SSU are most likely to be households.

child in the home in the PRIDI age group (the most-cited reason), no one at home, incorrect registry of the home, mother or child not at home, and refusal to participate.

Phase II: Instrumentation

The same PRIDI instrumentation, adapted on the basis on Phase I results, were used in Phase II.

- Engle Scale of Child Development. Directly applied to children to evaluate their cognitive, language and motor development. The Engle Scale consisted of two, separate evaluations: Form A and Form B, as in Phase I. Form A was applied aged 24 to 41 months, and Form B to children aged 42 to 59 months.
- Family Survey. Mother/caregiver report to collect information associated with the home and environment within which the child is developing, including socio-economic factors.
- Child Survey. Mother/caregiver report to collect information on factors associated with the child, including the child's socio-emotional development, basic health, participation in ECD or early education programs, and different methods of discipline applied in the home.
- Informed Consent Form. Signed by the mother or caregiver prior to applying PRIDI instrumentation.
- Registry of home visits. To guide the process of sampling in the field and keep track of the homes visited.
- Preliminary interview. To guide the selection of children to participate in PRIDI.
- Feedback Form. Filled out by the enumerators after completing the application of all PRIDI instrumentation with the purpose of registering any difficulty or issue encountered.

The objective of Phase II was to validate the PRIDI instruments adjusted at the end of Phase I against the internationally accepted instruments/measures. To this end, the following instrument and measure were also applied:

- Peabody Picture Vocabulary Test, in its Spanish version, Test de Vocabulario en Imágenes Peabody (TVIP).
- Height-for Age.

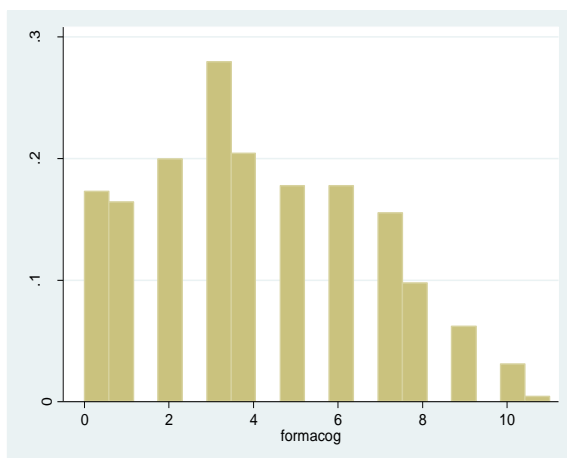
This entire battery of instruments was applied a sample of around 200 children in each of the four participating countries (see Table 5, above).

Results of Phase II

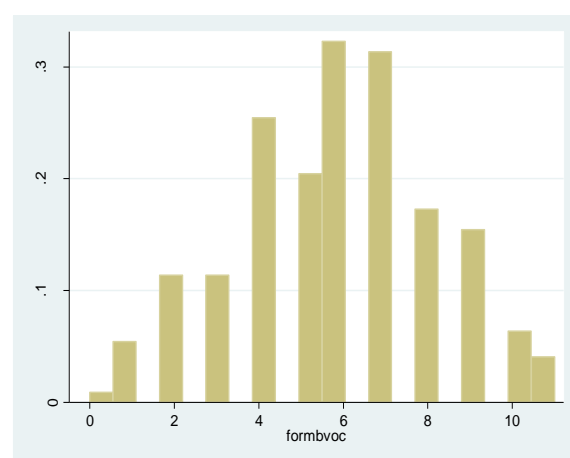
Phase II included far more and more sophisticated analyses than Phase I, which was limited to the analysis of the sum of all items across all domains. In particular, Phase II required that all data from each country be analyzed together. This was necessary to ensure the validity of each item, (e.g.,) that each item captured the same information in each country and within each country (across different populations), and that the PRIDI items bore an association with normed instruments, TVIP and height-for-age.

These expectations were confirmed with Phase II data. The distribution of scores from both Forms A and B were fairly evenly distributed (see Graphs 5 and 6). As with results from Phase I, the small size of the sample remains an issue in Phase II, thus limiting the ability to generalize or draw strong conclusions from the analyses.

Graph 5. Form A: Distribution of Scores

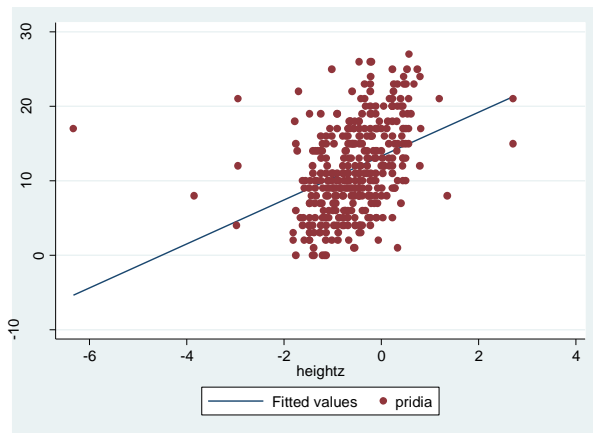


Graph 6. Form B: Distribution of Scores

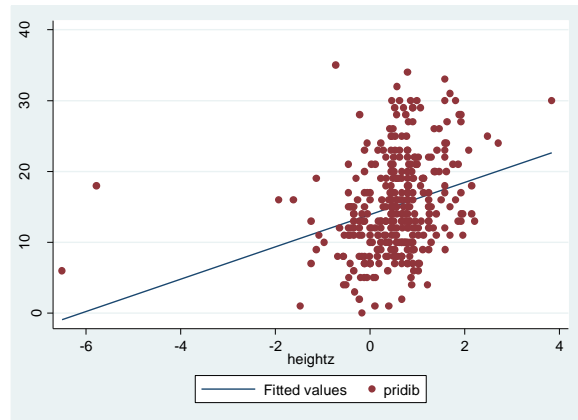


Both Forms correlated well with the TVIP and height-for-age, instruments that were administered precisely for the purposes of validation. Anthropometric data (height-for-age) were normalized to the World Health Organization scale. When correlated with the Engle Scale, the expected relationship – positive – appeared for both Forms (see Graphs 7 and 8).

Graph 7. Form A and Height Z-score
 $r=0.3827$

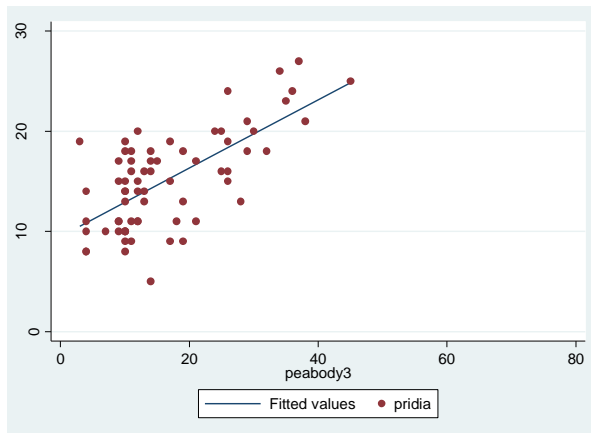


Graph 8. Form B and Height Z-score
 $r=0.295$

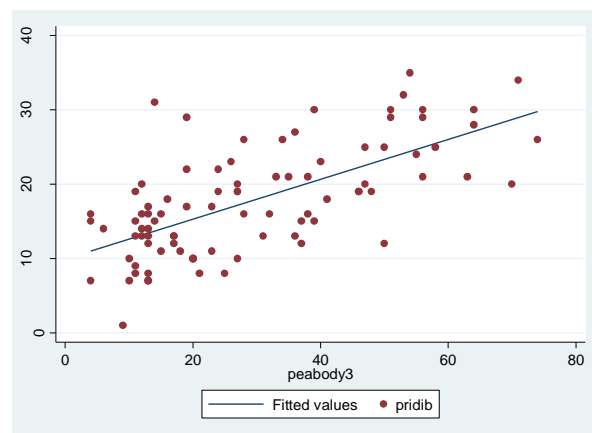


Results for the TVIP follow a similar trend, albeit with higher correlations (see Graphs 9, 10, 11 and 12).

Graph 9. Peru. Form A and TVIP
 $r=0.6811$

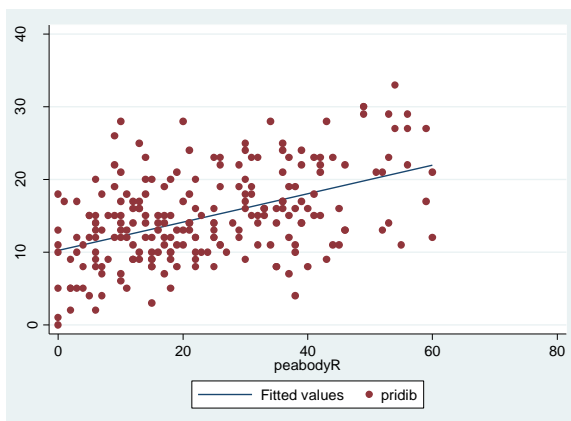
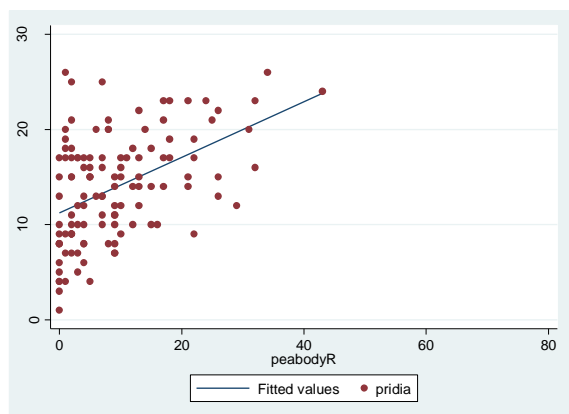


Graph 10. Peru. Form B and TVIP
 $r=0.6654$



Graph 11. Costa Rica, Nicaragua, Paraguay.
Form B and TVIP.
 $r=0.4685$

Graph 12. Costa Rica, Nicaragua, Paraguay.
Form B and TVIP.
 $r=0.6654$



Several details merit mentioning regarding the TVIP. This Test was applied across the Phase II sample, that is, to children receiving Form A and Form B, but not uniformly. Two different versions of the TVIP were administered. Peru applied the version validated in Spain¹² and the remaining three countries applied the version validated with a Latin American sample;¹³ in Phase III, all four countries applied the version used this latter version, which is older but of more common use in the region.

Prior to going to the field, several adaptations were made to the administration of the TVIP for use with PRIDI. Rather than starting with the age of the child and identifying the respective basal line and ceiling during application,¹⁴ in all PRIDI cases application of the TVIP started with item one and basal lines were not identified. This decision reflected operational considerations, including the difficulty faced in training enumerators how to establish a base line for different age groups. Starting with the first item simplified the application process and allowed for any necessary correction to be made afterward. Beyond this, the TVIP was applied using the original items and images according to instructions.

Item Validation: Cognition, Language and Motor Sub-Scales

IEA undertook Rasch analyses of the Engle Scale by item by country and by item across countries to examine issues of variability and item validation, independently of stratification. Items that appeared too hard or easy or that displayed low variation within and between countries were dropped, in order to ensure that the Scale was adequately discriminating

¹² See <http://web.teaediciones.com/peabody-test-de-vocabulario-en-imagenes.aspx>.

¹³ See <http://www.pearsonclinical.com/language/products/100000487/test-de-vocabulario-en-imagenes-peabody-tvip.html>.

¹⁴ The basal line consists of the highest group of eight consecutive correct answers. The ceiling is defined by the highest group of six correct answers within the previous eight items.

between children. Other items were dropped based on feedback from the field that indicated difficulties in administration, registering answers, and time required for application. For example, a number of items on both Forms prompted children to point and then (in a separate item) name a series of body parts (A17/A18), everyday objects (A21/A22 and B27/B28), activities (A23/A24 and B18/B19) or shapes (B8/B9). In all cases, items on both Forms A and B that prompted a child to point to these objects were found to be too easy and eliminated. Similarly, in cases where more than one item measured the same competency, items requiring inputs were eliminated (e.g., point to/name everyday objects versus was eliminated, but name body parts was kept). All in all, 14 items were dropped from Form A and 19 items from Form B (see Annex 5).

Some items were added to the Scale to ensure sufficient items across domains. Item A-18/B-18 (identify colors, a vocabulary item; used in various tests applied throughout the Region, including the TEPSI and Escala Abreviada Nelson Ortiz, and thus previously validated) was created and added to both Forms. Item B-22 (draw a person with 6 parts, a cognitive item widely applied as the “Draw a Person Test” and validated in the US and elsewhere) was created and added to Form B.

In some cases, remaining items were modified to ensure a better understanding by children or more efficiency in application. For example, on Form A, item 20 (memory of words) proved difficult. This was a cognitive, not a language item/competency; the intent was not to evaluate language or vocabulary. Enumerators said three sequences of two unrelated words, and children were prompted to repeat the sequence back to the enumerator. The enumerator then said three sequences of three words, which the children were prompted to repeat back. The children became fatigued by such repetition, and the item was changed to two sequences of two words, followed by two sequences of three words. A third sequence, of 4 words was added, so that the item would be identical in Form B and thus could be used as an anchor item (A/B-13, in the final Scale). Item 20 on Form B was modified to be the same as item 20 on Form A.

Item A23 and A18 prompted children to name a number of common activities (shown as drawings on cards), including petting a dog, milking a cow, washing, etc. The same drawings were used for applying both items. In Form A, the activity to be named was “brushing hair”; in Form B, it was “making a braid.” Analyses of results indicated that children did not always understand “brushing hair.” As a result, the prompt was changed to “making a braid,” and was used on both Forms. In much the same vein, children found “planting”, which appeared on both forms, confusing. It changed to “sweeping”.

In Form B, item 1 (exact quantity, with sub-items for 5, 10 and 20) was modified to reduce administration time: if a child failed on any sub-item, the item was discontinued. During the pilot phase all sub-items were applied. Same was modified in item 2, which prompted the child to count and add additional qualities (e.g., add two additional balls to the plate): if a child failed on the first sub-item, the item was to be discontinued.

Children were confused by items 13 and 14, which placed the puppet in front of or behind the child (item 13) or to the left or right of the child (item 14) and prompted the child to indicate in front/behind and right/left. The items were modified to use a pencil and asking the child to put it in front of/behind her or to right/left of her.

Changes in scoring were introduced in items B25 (animal card game) and B34 (sun and moon). In both cases, enumerators scored the number of correct answers. The modification introduced prompts enumerators to score correct and incorrect answers.

The application manuals for all such changes – elimination, modification or addition of an item – were modified accordingly.

Inclusion of Anchor Items

Twelve anchor items were included in each Form (5 in cognition, 4 in motor, and 3 in language) to allow for the eventual vertical equating of the two Forms, that is, to estimate how a child given Form A would have performed on Form B and vice-versa) and to allow for reporting as a single scale. Anchor items are the same in both Forms. To this end, small changes were introduced in both Forms, as indicated in Table 6.

Table 6. Anchor Items

		Form A	Form B	Modification
Cognition	Animal card game	A-10	B-10	Added to Form A
	Respond with coherence	A-11	B-11	-
	Memory of words	A-13	B-13	-
	Understanding temporal sequence, 2 actions	A-14	B-14	Added to Form B
	Organize sticks in color sequence	A-19	B-19	-
Motor	Imitate construction of bridge with 5 blocks	A-3	B-3	Added to Form A (previously was a bridge of 3 blocks)
	Walk in straight line, forward and backward	A-16	B-16	-
	Catch a ball	A-17	B-17	-

	Copy figures	A-21	B-21	Square and triangle added to Form A
Language	Differentiate concepts behind/in front	A-8	B-8	-
	Name actions	A-12	B-12	2 activities added to Form A (swing, knit)
	Identify colors	A-18	B-18	Added to both Forms

Scaling and Validation of Cognition, Language and Motor Domains

From these reduced Forms, factor and item response theory (IRT) analyses were undertaken to ensure that items theoretically mapped onto each domain hung together in an empirical analysis. Based on these results adjustments were made (e.g., IRT analyses suggested that some items thought fall into the language domain were found to be a better fit in cognition). Table 7 provides the final breakdown of items into domains:

Table 7. Engle Scale. Items by Domain

	Form A	Form B	Anchor Item
Cognition			
	A-1	B-1	
	A-2	B-2	
	A-4	B-4	
	A-6	B-5	
		B-7	
	A-10	B-10	X
	A-11	B-11	X
	A-13	B-13	X
	A-14	B-14	X
		B-15	
	A-19	B-19	X
		B-22	
Language			
	A-5		
	A-7	B-6	
	A-8	B-8	X
	A-9	B-9	
	A-12	B-12	X
	A-18	B-18	X
	A-20	B-20	

Motor			
	A-3	B-3	X
	A-15		
	A-16	B-16	X
	A-17	B-17	X
	A-21	B-21	X

Scaling and Validation of the Socio-Emotional Domain

The socio-emotional scale was included in the survey of the child and thus applied via report of the mother or caregiver. The same scale was applied to all children regardless of whether they were administered Form A or Form B. The scale consisted of 27 items of relevant behaviors, including social skills, emotional stability and autonomy. In an exercise much like that described above to validate the cognitive, language and motor domains, Rash/IRT and factor analyses were carried out to determine the extent to which these items empirically hung together in a single scale. These analyses suggested that, of the 27 items originally included in the socio-emotional scale, 15 hung together and should be maintained. The PRIDI team decided to include an additional item not included in the factor (child likes to play with other children, even if she doesn't know them), given that it was the only item that spoke to how the child interacts with her peers (theoretically an important aspect of socio-emotional development). The resulting, reduced socio-emotional scale consisted in 16 items.

Estimation of Internal Consistency of All Domains

The internal consistency of each domain (Cronbach's Alpha Reliability Coefficient) was calculated. With the exception of the motor and language domains in Form B, which fell marginally below the threshold defined by the PRIDI team ($> .6$), the resulting coefficients were acceptable (Table 8).

Table 8. Internal Consistency of the Engle Scale

Form and its Dimensions	Cronbach's Alpha
Form A	
Cognition	0.6904
Motor	0.6679
Language	0.7098
Form B	

Cognition	0.7840
Motor	0.5523
Language	0.5768
Socio-Emotional	.7407

Results and Adjustments of the Surveys

Items in both the survey of the child and the family were analyzed by item by country and by item between countries for variation. In instances where there was no variation, the item was dropped. For example, a number of the items on the socio-emotional scale showed little or no variation within or between countries (e.g., child is happy; child is sad, worried or pensive). In those cases where enumerator feedback indicated difficulty or confusion in application or scoring, the item was dropped, particularly if no improvement in the scoring could be found. Items inquiring into the mother's health appeared to be often misunderstood, particularly in rural areas, and were eliminated. In instances where an enumerator could verify information from a health certificate or birth certificate, instructions were changed. Enumerators asked for the relevant document and copied the necessary information, rather than asking the mother or caregiver to provide such information.

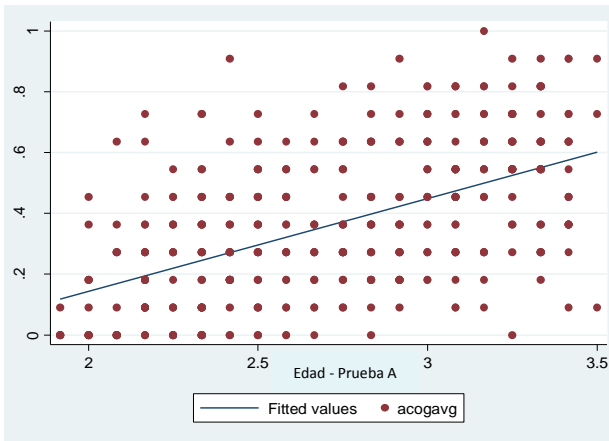
Both surveys were shortened and consolidated into a single survey. Changes in formatting across the board were introduced to facilitate more efficient scoring of items, for example, for items that were not applied or that had an error in application (combined into a single code). Changes introduced into the surveys intended to reduce duplication and ensure that each question was clearly understood. For example, a paragraph was added instructing enumerators what to do if the child was disabled. The UPM (primary sampling unit) was added to all forms. In items inquiring into the highest level of education achieved, a single category was created for graduate-level education (previously, there were options for master and doctorate). Annex 6 details these and the other changes made in the surveys and forms.

Analyses of Domains and Key Associated Factors

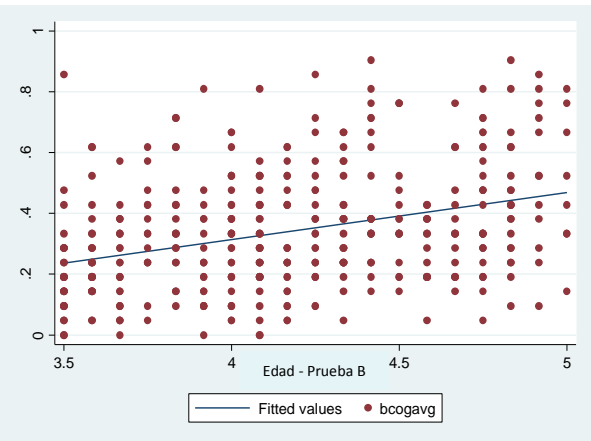
Several analyses were undertaken to ensure that results of the Engle Scale discriminated by key factors identified in the literature, including age, maternal education, the socio-economic characteristics of the home, and the nurturing environment (described below). In each case, the expected trends appear. Correlations are positive, if somewhat low, likely reflecting the small sample size.

Insofar as ECD refers to the ordered emergence of skills and abilities, older children perform better (Graphs 13, 14, 15, 16, 17, 18 and 19). In all domains evaluated by direct observation, Form A discriminates better than Form B. The correlation with socio-emotional skills is weaker.

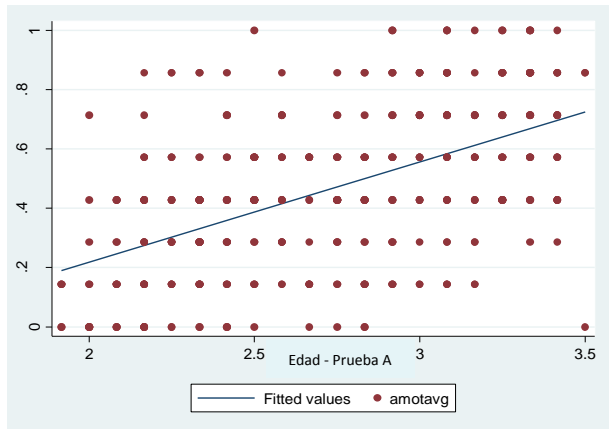
Graph 13. Form A. Cognition and Age
r=0.5468



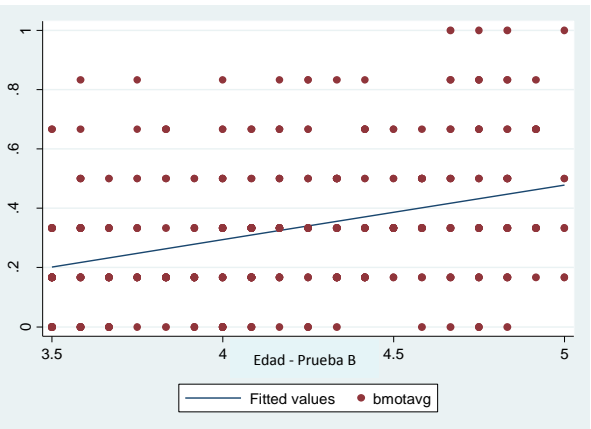
Graph 14. Form B. Cognition and Age
r=0.3412



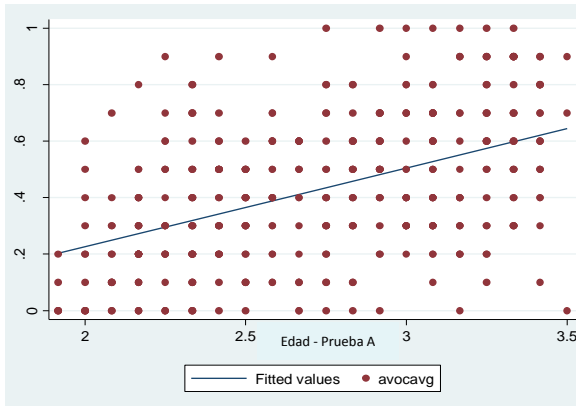
Graph 15. Form A. Motor and Age
r=0.5374



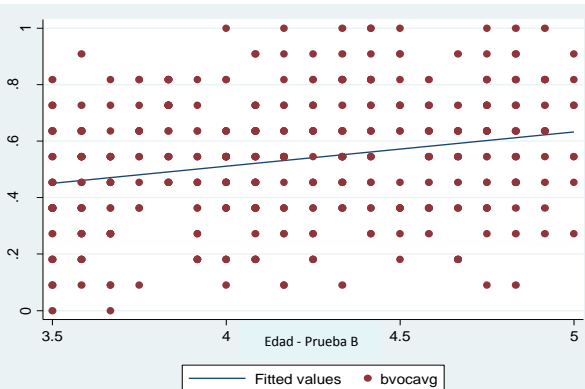
Graph 16. Form B. Motor and Age
r= 0.3453



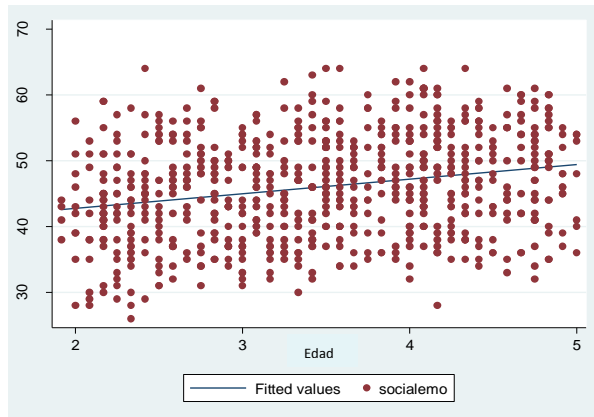
Graph 17. Form A, Language and Age
 $r=0.4724$



Graph 18. Form B, Language and Age
 $r=0.2441$

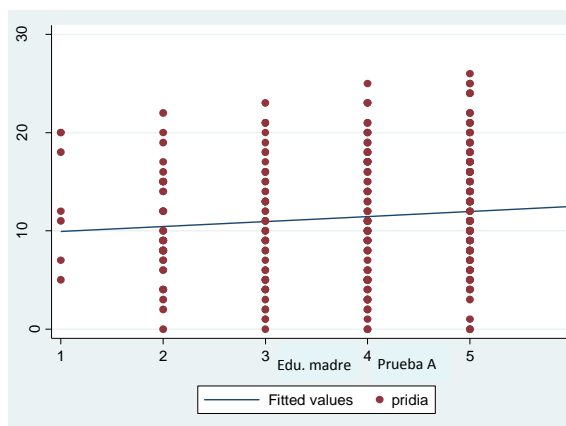


Graph 19. Socio-Emotional and Age of Child
 $r=0.2410$

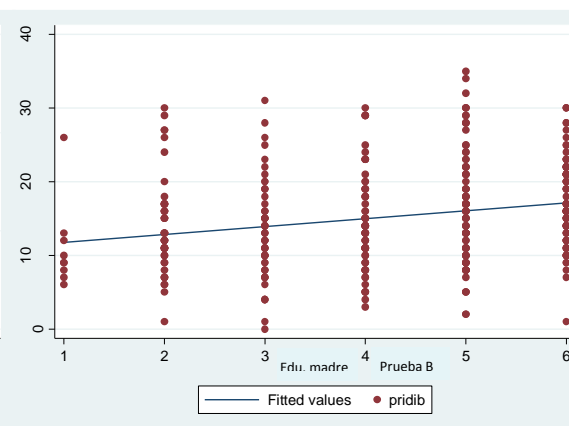


Similar trends also appear for maternal education (Graphs 20, 21 and 22). Again, the low correlations are likely to be a result of the small sample size.

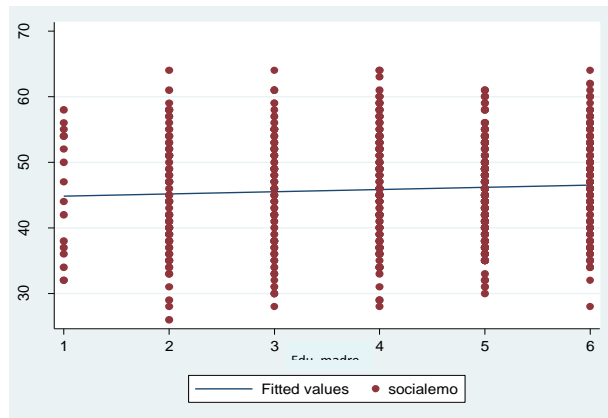
Graph 20. Form A and Mother's Education
 $r=0.1136$



Graph 21. Form B and Mother's Education
 $r=0.2273$

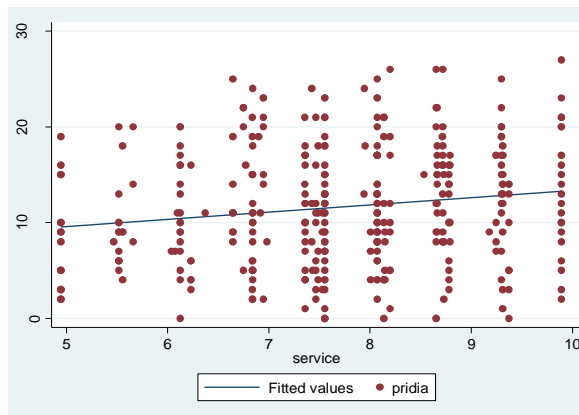


Graph 22. Socio-Emotional and Mother's Education
 $r=0.0603$



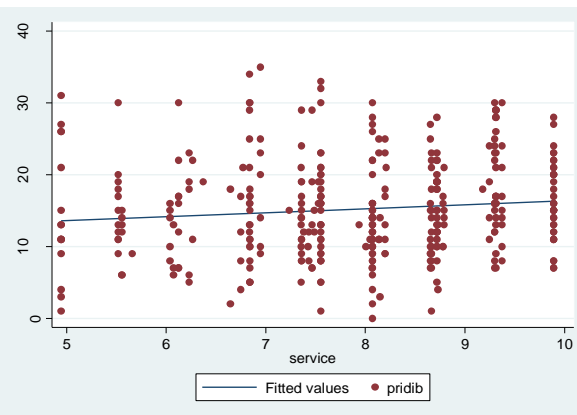
Much the same can be observed for the correlations with services in the home, used in this instance as a proxy for the socio-economic level of the home (Graphs 22, 23 and 24). The correlations are positive but weak for Forms A and B, and stronger for the socio-emotional domain.

Graph 22. Form A and Services in Home
 $r=0.1543$



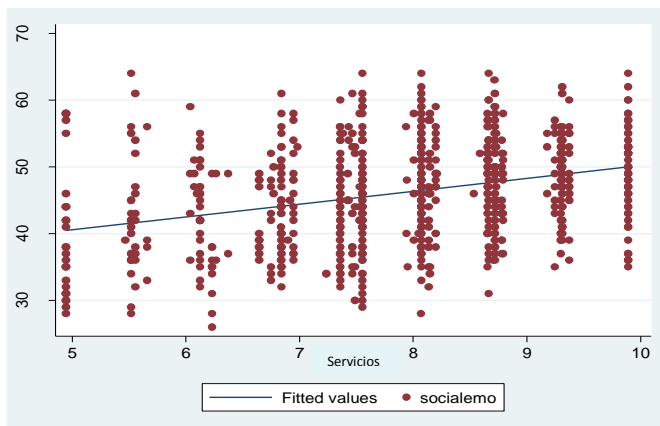
Servicios Prueba A

Graph 23. Form B and Services in Home
 $r=0.1084$



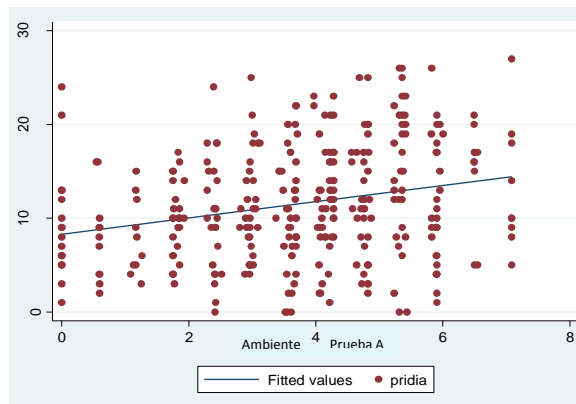
Servicios Prueba B

Graph 24. Socio-Emotional and Services in Home
 $r=0.3218$

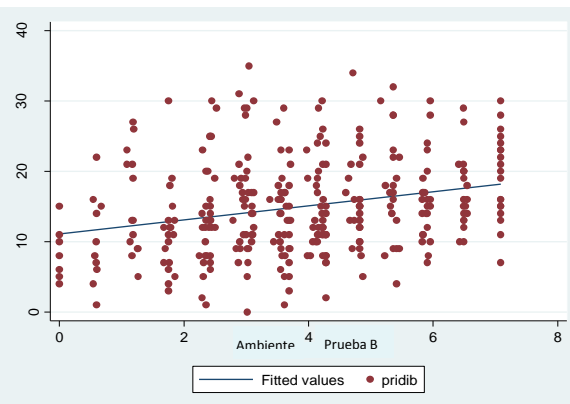


The final correlations are with the nurturing environment, an index created using information on the number of books available in the home and on the type of interaction between the child and other members of the household, Graphs 25, 26 and 27). Correlations are positive, if low.

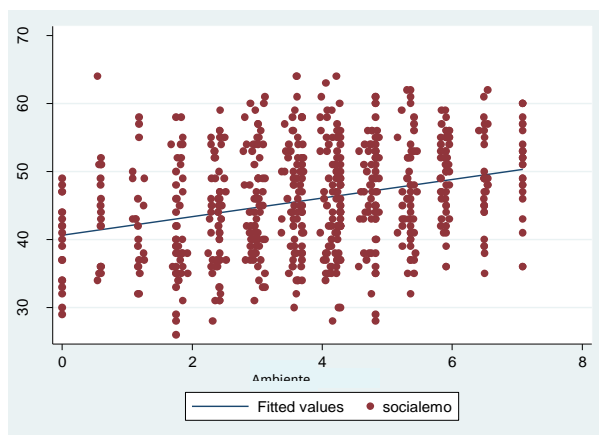
Graph 25. Form A and Nurturing Environment
 $r=0.2505$



Graph 26. Form B and Nurturing Environment
 $r=0.2614$



Graph 27. Socio-Emotional and Nurturing Environment
 $r=0.3011$



Operational Issues during Phase II

In addition to permitting a more rigorous analysis of results, Phase II shed light on a number of operational issues related to the application of instrumentation. Foremost among these was the time required to apply the entire battery of tests (PRIDI instrumentation, TVIP and height-for-age). The Scale and surveys during Phase II were kept purposively long, with the idea that no item would be eliminated without a field test; limited field tests during Phase I indicated a total of 40 minutes for the Scale and surveys. Yet the time required for the application of all Phase II instrumentation proved to be excessive, running up to 90 minutes for Form A and to 120 minutes for Form B,¹⁵ in addition to the application of the surveys, which required an additional 20-40 minutes each, the TVIP and height-for-age. Nicaragua and Paraguay, citing the need to establish a warm relationship with the child prior to starting the evaluation, reported application times of 3.5 to 4 hours. Costa Rica, the only country with a field protocol in which a pair of enumerators – one to evaluate children; the other to apply the surveys – reported 45 minutes for the full battery of instruments applied to Form A children, and up to two hours for Form B children.

Such considerations reduced the number of evaluations an enumerator could complete per day. All countries calculated their Phase II budgets on 4 evaluations per day. Field realities were

¹⁵ The enumerators in Phase I, particularly in the Peruvian case, had considerable experience in evaluating young children and were closely involved in the creation of Forms A and B, including in the design of the manuals and training for their application. These enumerators thus were “expert” and “efficient” in the application of the PRIDI instruments when compared to their counterparts (e.g., in Paraguay). Other differences accounting for a longer application time in Phase II included the TVIP and height-for-age measurements that were not done in Phase I.

different. On average, enumerators completed less than 3 evaluations per day (often 2) in Nicaragua, and 3 in Paraguay and Peru; each team in Costa Rica could complete a single evaluation per day. In many cases, PRIDI was applied early in the morning or late in the evening, given that these were the only time mothers had available. Each of the four countries reported applying PRIDI on weekends.

As with Phase I, ideal conditions for the application of instrumentation rarely existed. In a large majority of cases, the Scale was applied on the floor and the surveys either in the doorway, patio or outside (see photos below, the first from Nicaragua; the other two from Peru).



Peru, interestingly, calculated average application times for the different regions in which it applied PRIDI. As can be seen from the table below, application time of the Engle Scale in Metropolitan Lima was about half of that required in rural and indigenous areas (see Table 9).

Table 9. Estimated Time of Application in Peru, Phase II

	Selva	Sierra	Costa	Metro Lima
Form A	45 - 60 min	35 min	35-45 min	40-50 min
	1 h 10 min - 1 h	45 min - 1	60 min - 1 h	
Form B	30 min	h 15 min	15 min	50-60 min
TVIP	10 -15 min	10-30 min	n/d	15-25 min
Height-for-Age	5-10 min	5-10 min	5-7 min	3-8 min
Survey Child	30-35 min	31-35 mn	30-40 min	15-25 min
Survey Home	20-25 min	21-25 min	16-22 min	15-25 min

Such results called attention to different aspects of application, including the difficult conditions in which the Scale was applied (e.g., on the floor, other family members wanting to watch). Differences in application times for the surveys reflected the use terms either unfamiliar (e.g., emotions) or confusing (e.g., difference between programs; or between rules, routines and limits) in rural or indigenous areas.

Within this context, a series of problems arose in relation to the TVIP and its application in the field. Foremost among these was an extremely and unacceptably high level of administrator error in scoring (see Table 10).

Table 10. Issues in Application. TVIP

	Peabody Test Administered	Test Discontinued Too Late	Test Discontinued Too Early	Error in Administration
Costa Rica	156	53 (34.0%)	3 (1.9%)	35.9%
Nicaragua	116	17 (14.6%)	9 (7.8%)	22.4%
Peru	172	6 (3.5%)	32 (18.6%)	22.1%
Paraguay	147	49 (33.3%)	54 (36.7%)	70.0%

Administration errors included terminating the test either before or after the appropriate cut-off point (determined by the child's performance). Instances where the TVIP was stopped after the appropriate cut-off point (after the child had already answered six or the preceding eight questions incorrectly) were salvageable. Scores could be manually corrected by carefully reviewing the results and recalculating the correct cut-off point, although at considerable cost in terms of time of the reviewer. This was done during Phase II by the PRIDI team. However, in tests that were discontinued too early, children were not administered a sufficient number of items to calculate a ceiling. These cases had to be (and were) considered void, as there was no way to calculate the correct score with the data collected.

In addition, the TVIP did not function well with the younger children, that is, those to whom Form A was administered. Many of these children failed to progress beyond the example and practice questions, meaning that the Test, per se, was not administered.

No country provided standard TVIP scores in Phase II. Standard scores are age-based. For example, three year-olds would not be expected to perform at the same level as five year-olds.

The high rate of administration error and errors in calculating scores from all countries was troublesome moving into Phase III. In Phase II, all of TVIP scores from all countries were confirmed and, where possible, corrected manually. In the case of Paraguay, the country team did not input data for the raw scores, necessitating the manual recalculation of all scores by the PRIDI team.

In Phase III, the country teams were asked to provide standard as well as raw scores; no manual confirmation of scores would be possible, given the size of the nationally representative

samples and budgetary limitations. The PRIDI team thus placed a high priority on reviewing and adapting the training curriculum for Phase III, particularly for the TVIP, and on monitoring compliance with its implementation.¹⁶

Issues of sampling also arose during Phase II. The fact that a large number of homes visited did not, in fact, have an eligible child, reflected the use of outdated census data. In Costa Rica, the sample for Phase II was drawn based census data from 2000. In the case of Nicaragua, the sample was drawn from 2005 census data and in Peru, where the results were better, the data came from the 2007 census. Paraguay, with the support of IEA, drew its sample without census data. It identified 3 departments (Asuncion, Central and Guaira) and three municipalities within each. Technical staff from the Department of Initial Education in the Ministry of Education provided the team with lists of homes with children in the PRIDI age cohort. This allowed for a considerably higher acceptance rate in Paraguay. In each of the four countries, these issues were re-visited in preparation for Phase III. As will be discussed below, the PRIDI team and IEA in particular, worked directly with the countries to ensure more efficient sampling frames for Phase III.

A number of delays affected the execution of Phase II. Although the fieldwork in all countries started around the same time, four months separated the receipt of the first database from the last. Reasons for such delays had to do with many of the issues mentioned above, including the rainy season and, in the case of Paraguay, the need to coordinate with the Ministry of Education, which directly supported the firm responsible for data collection with staff (trained by the firm) who served as enumerators. Given the international character of PRIDI and the need to analyze all data together, Phase II could only move as quickly as the slowest participant. These issues arose again during Phase III, during which 11 months elapsed between receipt of the first and last database (see below).

¹⁶ In the case of Paraguay, the Phase II training curriculum was not fully implemented; enumerators received less than half the mandated training (4.5 of the 10 mandated days), a likely cause for the high rate of administration error (70%), especially for tests that were discontinued too early.

6. Phase III: Application in Nationally Representative Samples

A concrete result expected from Phase II was the validation of streamlined instruments and, with this, a significant reduction in the time required for their adaptation.¹⁷ As a result of the analyses described below, Form A was reduced to 21 items and Form B was reduced to 22 items; the Surveys were consolidated. When tested, application time for the Engle Scale, the Survey and the entire battery of accompanying instrumentation ranged between half an hour and 40 minutes. Given the issues raised during Phase II (e.g., Form A children often not proceeding beyond the practice questions), the PRIDI team decided to limit the application of the TVIP to children receiving Form B in Phase III. To ease application and organization in the field, scoring and instructions for height-for-age and for scoring the TVIP were added to both Forms. Other adaptations in scoring and registry were introduced to facilitate a more efficient in field application. In addition to those mentioned above, in homes where more than one child was eligible to participate in PRIDI, the rule on how to select the child was modified. Rather than using a child's birth to randomly select the child, enumerators used the first letter of his or her name.

Building on the experience of Phase II, some modifications were introduced into the profile of enumerators and the training manual. In line with Paraguay's objective to install capacities for the evaluation of young children, it added the requirement that all enumerators be functionaries of the Ministry of Education. The criterion of recruiting only women as enumerators was removed, as both Paraguay and Nicaragua demonstrated that the sex of the enumerator made little difference. What did make a difference in recruiting enumerators was whether or not they had prior field experience and had worked with small children in the past. The training manual and curriculum were modified to reflect these changes, as well as the need to include more structured training on the application and scoring of the TVIP and practice sessions for applying the surveys to mothers. An additional day of training (for a total of 11 days) was mandated. Regarding the TVIP, enumerators would be required to confirm TVIP scores prior to inputting data and include standard scores (as a reference for age-adjustment of raw scores) in their respective databases for Phase III.

Results from Phase II were presented to the country coordinators and representatives from each of the firms responsible for the field work in San Jose, Costa Rica, in January of 2013. At the request of the countries, minor changes were introduced into the items of the Scale and survey, and the application and training manuals and the training DVD were revised and

¹⁷ More comprehensive training could increase the efficiency of enumerators, but the main technique to decrease the time needed was to shorten the instruments.

updated accordingly. All PRIDI instrumentation was finalized in February 2013 and distributed to the countries for application in Phase III (nationally representative samples of around 2,000 children). Prior to going to the field, each country reviewed the instruments and introduced any new national adaptation via the FAN, as in Phase II, and recruited and trained the necessary enumerators. The PRIDI team emphasized the need for adhering to the PRIDI training curriculum and duly implementing all its activities.

Given that issues related to sampling assumed a paramount importance in Phase III, an entire chapter (8) is devoted to these issues; Chapter 9 discussed sampling weights and variance. In total, 7,710 children participated in PRIDI, between Forms A and B (see Table 11).

Table 11: Number of Children by Form and Country

	Costa Rica	Nicaragua	Paraguay	Peru	Total
Form A	886	881	720	1289	3,776
Form B	918	954	784	1278	3,934
Total	1,804	1,835	1504	2567	7,710

Phase III: Instrumentation

Phase III saw some changes in instrumentation. All PRIDI instrumentation was in its final, streamlined version and format.

- Engle Scale of Child Development. Form A, 21 items, for children 24 to 42 months. Form B, 22 items, for children 43 to 59 months.
- The Family and Child Survey. A single, consolidated survey.
- Accompanying Forms:
 - Informed Consent Form.
 - Registry of home visits.
 - Preliminary interview.
 - Feedback Form.
- Peabody Picture Vocabulary Test, applied only to Form B children (version validated in the Latin American sample).
- Height-for Age.

Results of Phase III

Chapter 10 details the scaling and scaling methodologies used to ensure the reliability and validity of the scales. These processes confirmed the sub-scales both for Form A and Form B

(see Tables 20, 22 and 28 in Chapter 10, where the Cronbach's Alpha scores are reported for each sub-scale) and the definitive version of the Engle Scale.

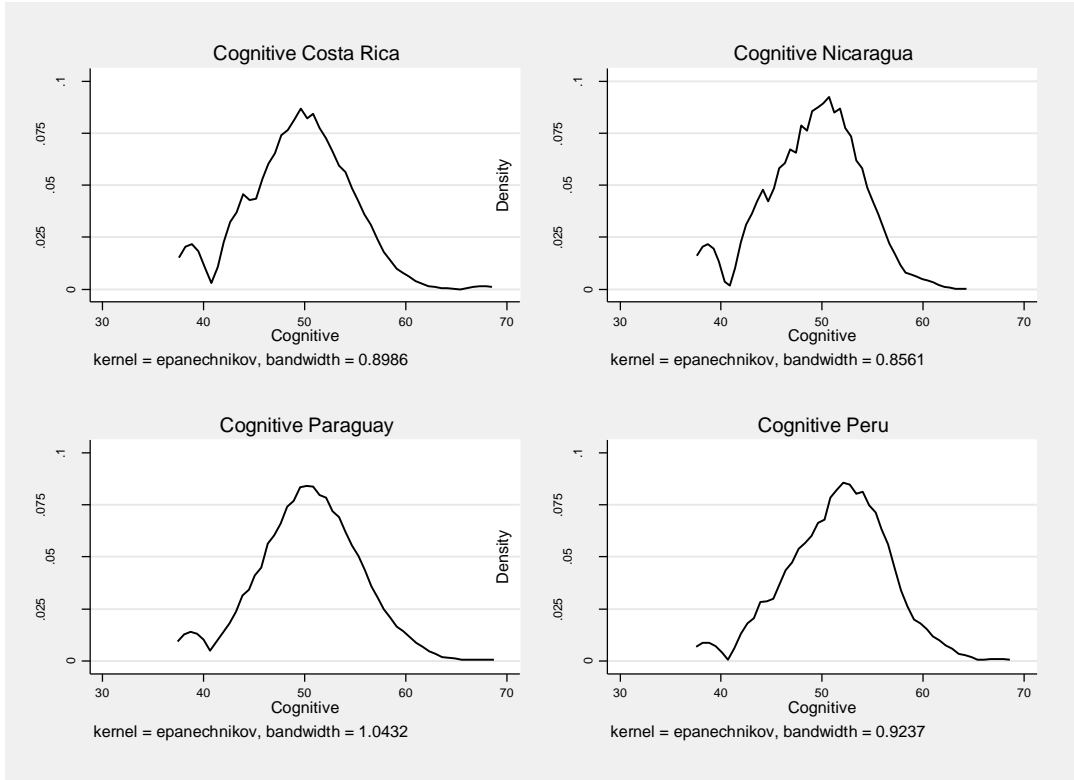
These analyses also prompted some recoding of items. For example, item 21, an anchor item, was re-coded in the following manner: 21A (draw a line), 0-1 (not correct/correct), and 21B, 21C, and 21D (circle, square and triangle, respectively) were recoded into a single item with 0 for not correct, 1 for 1 or 2 correct, and 2 for three correct. The justification for this change rested on the idea that, compared to drawing a line (the simplest of the tasks prompted), the three later items are complex designs that require that children join or complete lines into a geometric figure. With a minor change – the elimination of item G from the socio-emotional scale (child cries when mother/caregiver leaves), an item that did not hold together in the scale upon analysis – the Scale is the same as that used in Phase III; no changes in the survey were introduced.

Based on these results, a number of descriptive analyses were undertaken. Most are reported in the PRIDI Final Report, and are not reproduced here. Except where noted, scores are presented on a single, vertically equated scale, where the international average is 50 and the standard deviation is 5.

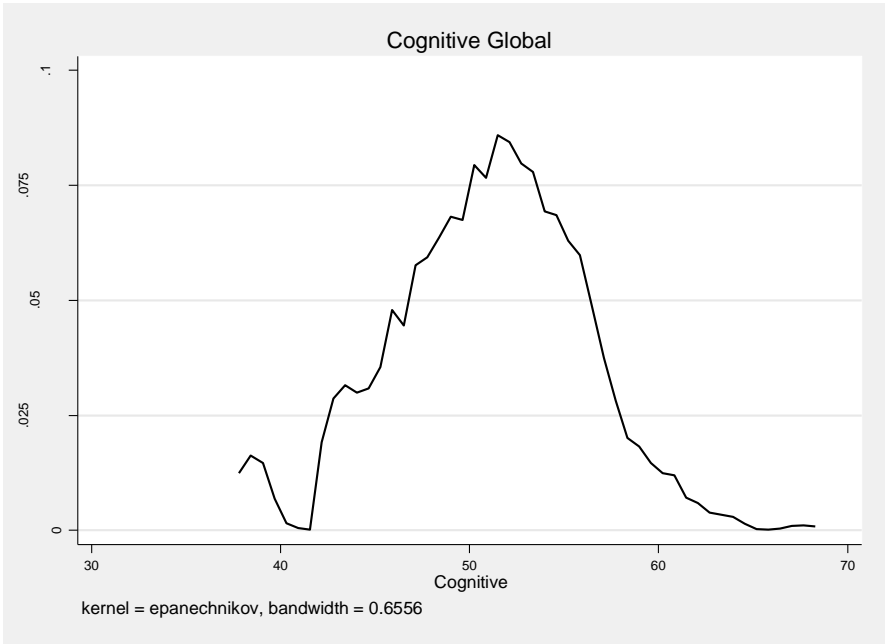
Distributions of Scores

One of the first analyses done was a kernel density plot of each domain to determine if its distribution was normal. As can be seen from the graphs below, although the distributions are not bimodal, there is a peak that appears at the lower end of the cognitive, language and motor sub-scales (see Graphs 28, 29, 30, 31, 32 and 33).

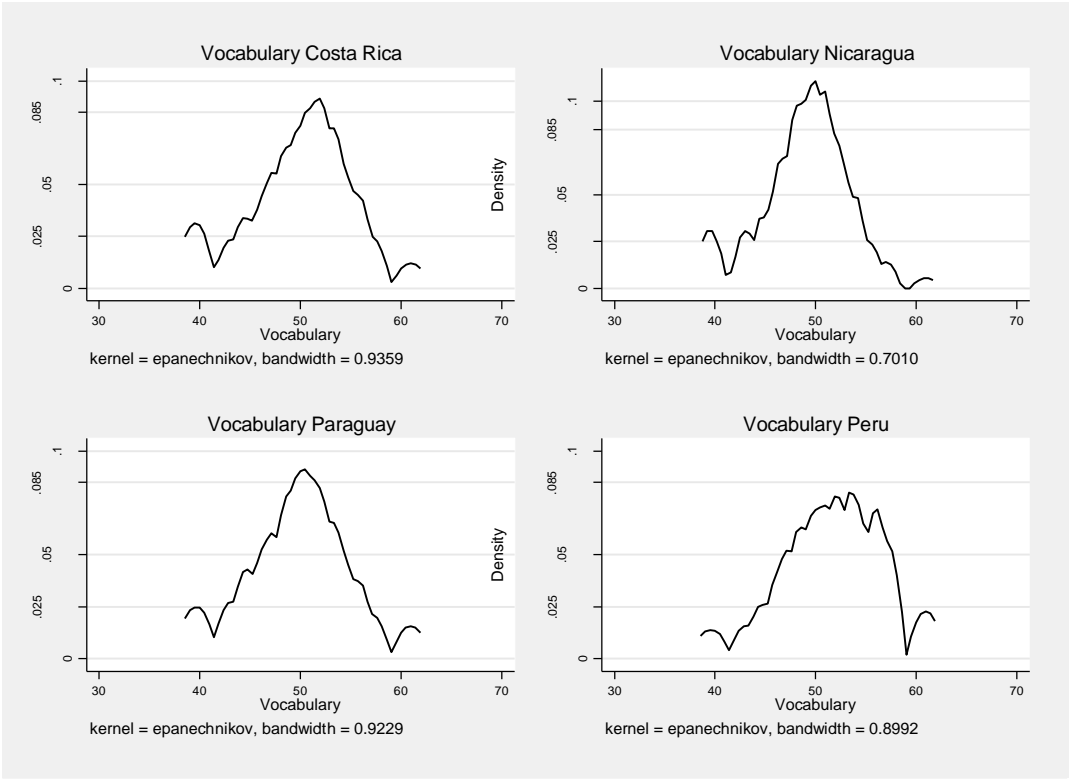
Graph 28. Distribution of Scores in the Cognitive Domain, by Country



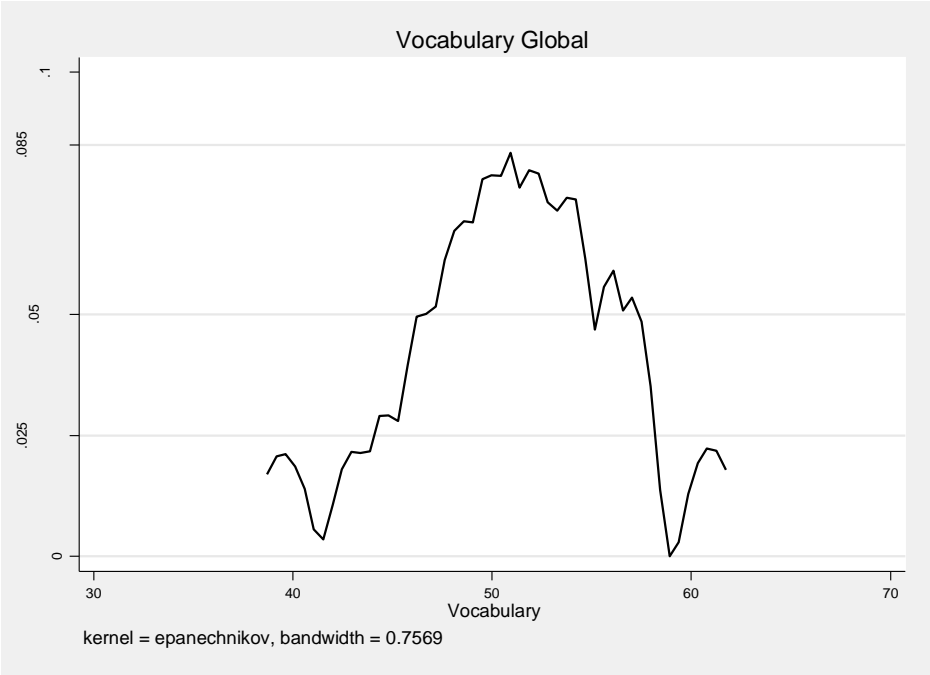
Graph 29. Distribution of Scores in the Cognitive Domain, Global



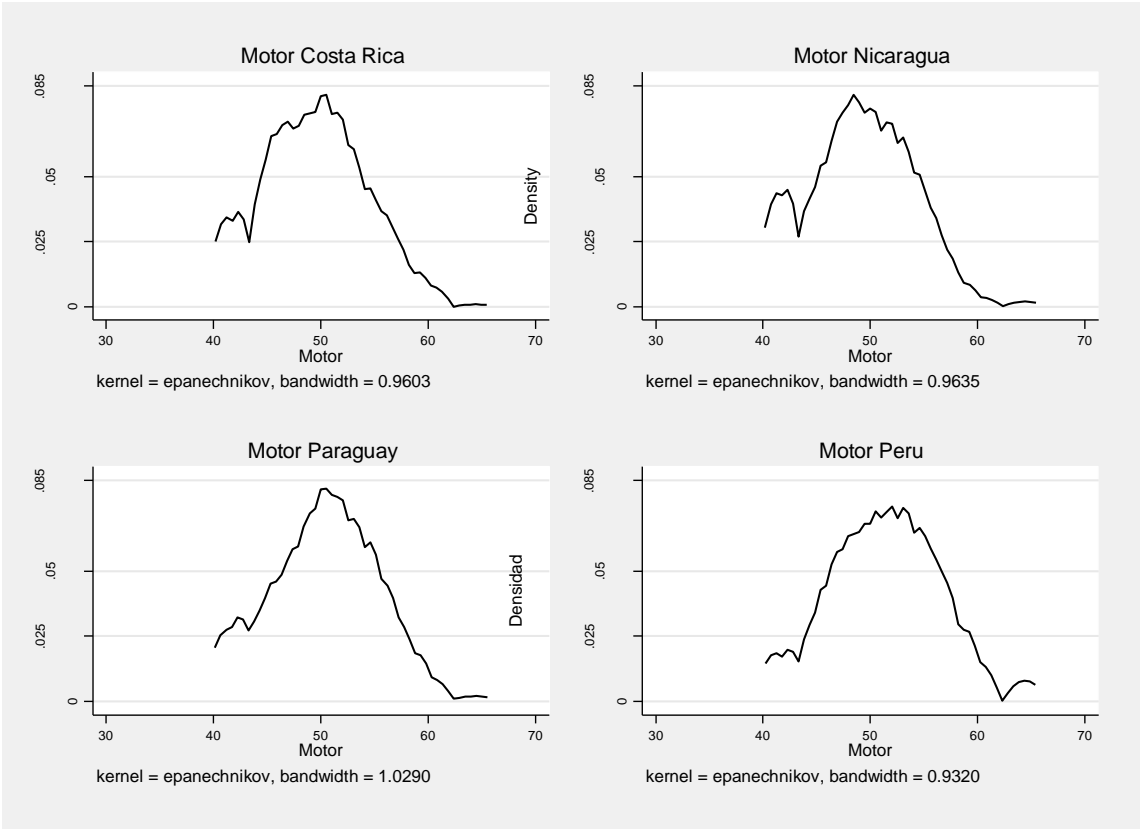
Graph 30. Distribution of Scores in the Language and Communication Domain, by Country



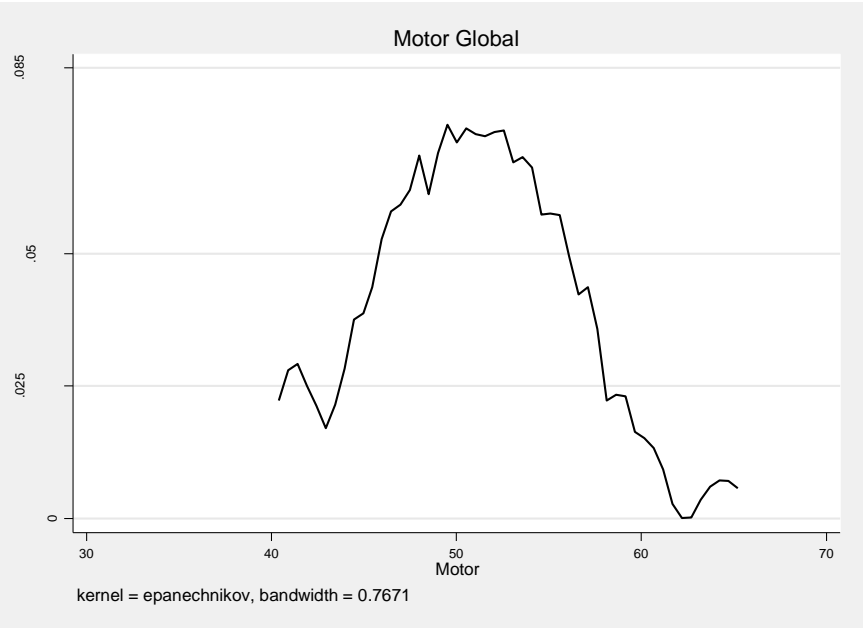
Graph 31. Distribution of Scores in the Language and Communication Domain, Global



Graph 32. Distribution of Scores in the Motor Domain, by Country



Graph 33. Distribution of Scores in the Motor Domain, Global



Extreme scores on both ends of the distribution merit comment. Regarding those on the left-hand side, all items on the Engle Scale were converted to a score of either 0-1 or 0-1-2. For example, an original raw score of 31 in Form A cognitive was recoded to a maximum of 14. Thus, any recoded-zero – that is, an extreme score at the bottom end of the distribution - reflects that a child could not reach a minimum threshold, as defined by the converted score, not necessarily that a child could not get a single item correct in the original format.

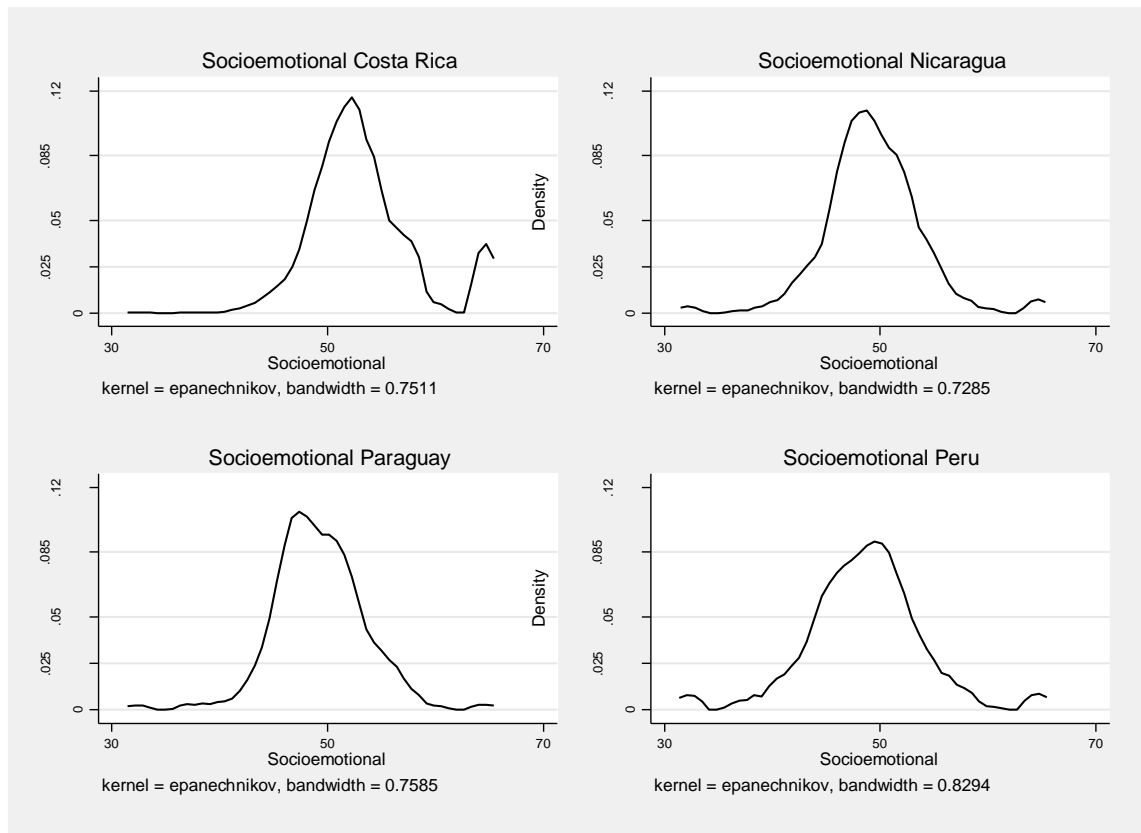
Insofar as it was suspected that they majority of these children would be PRIDI's youngest participants, the percent of 2 years old also was calculated. This exercise was done for results on Form A only. On a vertically equated scale, the probability of Form B children getting Form A items right can be assumed to be high (see Table 12).

Table 12. % Children with No Item Correct, by Domain

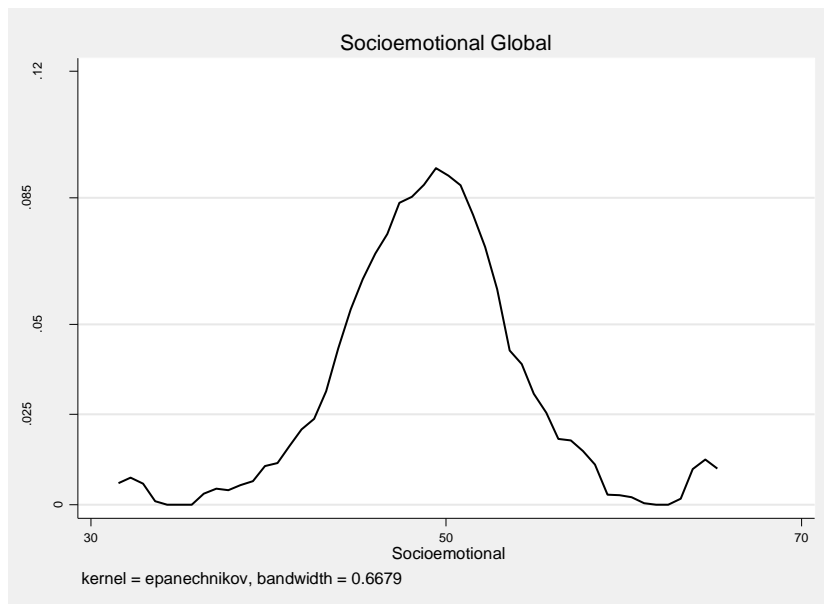
Domain	Country	% of children with no item correct	Among children with no item correct, % 2 year-olds
Cognitive	Costa Rica	4.98%	92.26%
	Nicaragua	5.04%	91.58%
	Paraguay	3.37%	92.39%
	Peru	2.32%	96.42%
Language	Costa Rica	8.59%	92.76%
	Nicaragua	6.52%	87.78%
	Paraguay	6.54%	83.54%
	Peru	3.62%	96.59%
Motor	Costa Rica	8.76%	91.49%
	Nicaragua	10.80%	90.59%
	Paraguay	7.67%	90.36%
	Peru	4.99%	92.87%

The opposite trend appeared in the same analysis of the socio-emotional sub-scale. Although not bimodal, a second peak appears at the higher end of the distribution, particularly in Costa Rica (see Graphs 34 and 35).

Graph 35. Distribution of Scores in the Socio-Emotional Domain, by Country



Graph 35. Distribution of Scores in the Socio-Emotional Domain, Global



To better understand the extreme scores on the right-hand tail of the distribution, the percent of four year-olds about whom their mothers or caregivers answered with the highest scoring option on all items was calculated (see Table 13).

Table 13. % of Mothers Answering With the Highest Scoring Option on All Items, Socio-Emotional Scale

Socio-Emotional	% of mothers answering yes to all items	% four year-olds about whom the mother responded
Costa Rica	8.10%	39.73%
Nicaragua	1.61%	81.80%
Paraguay	0.43%	21.96%
Peru	2.06%	52.14%

Results for all domains were analyzed by the PRIDI team. Reasons underlying the zero score in the cognitive, language and motor domains, and high-end score in the socio-emotional domain are not known. In the case of the zero scores, children could have refused to do the item, or did not know the answer, or there was an administrator effort in applying the Scale. For the high scores in the socio-emotional domain, mothers may not have understood the question, or there was an administrator effect, or some other factor. But none of these potential explanations could be verified. Administrator effects, for example, could bias the entire range of scores, not just those at the two extremes.

Based on these considerations, and the fact that the Engle Scale emerged from a series of reasoned decisions taken over the course of the three phases (and described here), the PRIDI team decided to maintain all cases in the final analyses of Phase III data. That said, results for PRIDI's youngest participants, i.e., two year olds, should be interpreted with some caution, as should results in the socio-emotional domain for four year olds.

Calculation of Indices

To further explore whether PRIDI, and the Engle Scale in particular, maintain the expected correlations with key associated factors identified in the literature, two indices were calculated using exploratory factor analysis: wealth index and the nurturing environment. The wealth index was computed following Schady et al. (2014) using characteristics of the home found in the Survey. The nurturing environment index was calculated following the Family Care Indicators and

Hamadani et al (2010). Table 14 includes the items and their ranges used in the construction of each index.

Table 14. Items and Ranges Used in the Construction of the Wealth and Nurturing Indices

Index	Survey Item	Range
Wealth	Infrastructure in the home in good condition (terminado); includes wall, roof and flooring	0-3
	Number of assets in the home; includes refrigerator, gas or electric stove, electric iron, bicycle, motorcycle, car or van, radio, television, computer, landline phone, cell phone, and internet;	0-12
	Access to services in the home; includes access to electricity, potable water and sanitation	0-3
	Ratio of number of household members to bedrooms	0 upwards
Nurturing Environment		
	Number of books in home	0-3: 0=0, 1=1-5, 2=6-10, 3=11+
	People who interact with the (PRIDI) child; includes father, mother, other relatives, and friends	0-4
	Frequency of interaction	0-4: 0 = 0 times/week, 1 = 1-2 times/week, 3 = 3-4 times a week, and 4 = every day
	Number of activities undertaken with (PRIDI) child; includes read books or look at pictures in a book, tell stories, sing songs, go outside, play and name things, and draw things	0-6
	Rules in the home; includes types of foods eaten, family meal times, bed times, and household chores;	0-4
	Hygiene practice by child; includes washing of hands before eating, brushing teeth after eating, and washing hands after using the bathroom	0-3

Table 15 indicates that the wealth and nurturing environment indices, while related, are capturing different factors and vary in their relation to different subdomains. The wealth or the socio-economic status of the household thus does not appear to be a determinant variable in and of itself of child development.

Table 15. Correlations among key variables

	Socio-emotional	Cognitive	Motor	Language	Height-for-age (z)	Wealth Index
Cognitive	0.21*					
Motor	0.17*	0.69*				
Language	0.19*	0.75*	0.69*			
Height-for-age (z)	0.02	0.02	0.02	0.05*		
Wealth Index	0.14*	0.18*	0.10*	0.21*	0.19*	
Nurturing Environment	0.25*	0.24*	0.17*	0.27*	0.09*	0.35*

* Correlations statistically significant at 5%

These results similarly indicate that the developmental domains measured by the Engle Scale are related, albeit different. This is consistent with the literature and desirable. Each domain contributes to healthy child development.

The final PRIDI report details all further results of from the Engle Scale; these results were reviewed and approved by each of the participating countries (see Verdisco et al., 2014). Chapter 11 details the creation and validation of the PRIDI database.

Operational Issues during Phase III

All four countries reported difficulties in complying with the intended sample (see Chapter 8 for details). With high levels of internal migration in all countries, census data were not always accurate. Even when PSUs were correctly identified, long distances often separated one home from another. In Peru, enumerators often walked between 1.5 and two hours between homes in the Center-East region. Mountainous or otherwise difficult territory complicated travel in the rest of the zones, with the exception of the Costa (Lima and Callao). In Libertad, enumerators traveled 16 hours to reach the selected district and, once there, walked long distances to find homes with PRIDI children. In Nicaragua, enumerators walked between two and five kilometers between homes in the departments of Rio San Juan, Chontales, Boaca and Madriz. In this regard, having both the backpack of PRIDI materials and the talio-meter for measuring height-for-age were cumbersome. Both Costa Rica and Peru reported issues with the safety of their enumerators working in areas characterized by drug trafficking, thus requiring extra security measures to be taken (e.g., use of radios to maintain contact between supervisors and enumerators).

In the case of Paraguay, inaccuracies in the preliminary census data used to draw the sampling frame had enormous cost and time implication: the team visited 15,257 homes to comply with a reduced sample of 1,500 children. In many of these visits, the team noted the absence of the State. The PRIDI visit was the first made to the home by any official or affiliated agency. This complicated the efficiency with which the team was able to carry out its tasks. Mothers wanted to talk to the enumerators about their children, the state of their home, and themselves

(among other issues). Many expected some type of compensation for their time; none was given.

The comparison of the Paraguayan experience with the Nicaragua experience is interesting and offers an important lesson learned. Nicaragua based its sampling frame on 2005 census data. These data, like those used in Paraguay, were outdated. Prior to initiating the field work in each PSU, the respective application team in Nicaragua met with local authorities. These meeting proved critical, in that they yielded important information on the demographic composition of the PSU. Such meetings increased the efficiency with which the teams were able to comply with the necessary sample size. In total, in Nicaragua, the application teams visited 7,749 homes. Of these, less than 1% (0.74%) refused to participate in PRIDI.

No country reported problems related to the time required to apply PRIDI instrumentation. Application of time of the PRIDI instrumentation was around half an hour.

7. Sampling and Sampling Weights

Issues of sampling loomed large in preparation for Phase III. Unbiased estimates of population characteristics and meaningful comparisons of children between and within countries depend on sound and defensible selection procedures.

Each country was responsible for creating a nationally representative, random sample of 2,000 children (1,000 from 24 to 41 months; 1,000 from 42 to 59 months) based on a three-staged strategy, outlined below. Selection probabilities of sampled units were known at each step, thus allowing for the calculation of sampling weights and correct variance estimates. IEA reviewed and approved sample designs and supervised the selection process. This process is detailed below.

Target Population

The target population consisted of children between 24-59 months of age (after rounding). This target population did *not* include children that speak languages other than those used to apply the PRIDI instrumentation (Spanish in all four countries, in addition to Guaraní in Paraguay, Miskito in Nicaragua, and the Cusco variant of Quechua in Peru). children living out-of-country, children living in institutions and children with serious disabilities.

For test administration, two sub-groups were defined:

- Children between 24 and 41 months (Form A);
- Children between 42 and 59 months (Form B).

Each country maintained several exclusions for operational reasons:

- Costa Rica: small census sectors; child exclusion rate: <0.1%
- Nicaragua: Regions Región Autónoma del Atlántico Norte (RAAN, non-Miskitu children) and Región Autónoma del Atlántico Sur (RAAS), small communities; child exclusion rate: 16.4%
- Paraguay: El Chaco area; child exclusion rate: 1.2%
- Peru: small departments (Madre de Dios, Moquegua, Tacha, Tumbes, Amazonas, Apurímac, Huancavelica) and other indigenous languages; child exclusion rate: 8.6%

Sampling Stages

The basic outline for sample selection in each country followed a three-stage sample design. The Primary Sampling Units (PSUs) in the first stage of sampling consisted of geographical areas or administrative divisions. PSUs were non-overlapping and covered the entire country area (except for any areas or regions excluded ex-ante). PSUs in Costa Rica, Paraguay and Peru were census sectors; in Nicaragua, communities (in rural areas) and neighborhoods (in urban areas) were used as PSUs.

The Secondary Sampling Units (SSUs) in the second sampling stage consisted of households within the selected PSUs. In Costa Rica, Nicaragua and Paraguay, all SSUs in a sampled PSU were selected. In Peru, a sample of SSUs was randomly selected after an enumeration of residences in sampled PSUs.

As a third step of sample selection, one child was randomly selected within a sampled SSU in case more than one child of the PRIDI target population was found in a home. If more than one PRIDI-eligible child was found in a household, the child sample selection was performed randomly, using a table of selection numbers (*Kish grid*). This procedure replaced the use of birthdays applied in Phase II, as feedback from the field indicated that the birthday rule proved confusing.

Stratification

In order to improve the efficiency of the sample design and to ensure adequate representations of specific groups of interest in the sample, stratification was used during PSU sampling. Strata are groups of units that share some common characteristic which are likely to be linked to levels of child development. Independent samples of SSUs were selected from each stratum. The following explicit strata were formed used in PRIDI countries:

- Costa Rica: Area (Valle Central / rest of country), urbanization (rural / urban);
- Nicaragua: Departments and Regions, urbanization (rural / urban);
- Paraguay: urbanization (rural / urban);
- Peru: Oversampling area (Cusco / rest of country), natural regions (Sierra, Costa, Selva, Lima), proportion of Spanish speakers (high / low), urbanization (rural / urban).

Intended and Achieved Sample Size

An intended sample size of at least 2000 children was targeted in each country (in Peru, the target sample size was 2300, given oversampling in Cusco). This sample size met international standards and its precision required that the appropriate number of children were selected from a sufficient number of different PSUs. Where there was interest in a particular segment of the population, sample size was increased (oversampled) in areas where such segments were found. In the case of PRIDI, indigenous children were oversampled in the RAAN (Región Autónoma del Atlántico Norte in Nicaragua, Miskitu-speaking children) and in Cusco (Peru, Quechua-speaking children). In the case of Paraguay, initial results from the field indicated that the sample was smaller than expected. As a result, a second sample of PSUs was selected to compensate for the shortfall.

The intended and achieved sample sizes for PSUs and for children in each of the participating countries follow (Table 16):

Table 16. Intended and Achieved Sample Sizes

Country	PSUs		Children	
	Sampled	Achieved	Intended	Achieved
Costa Rica	150	150	2000	1804
Nicaragua	57	57	2000	1835
Paraguay	310	297	2000	1504
Peru	416	416	2300	2567

Annex 7 provides additional detail.

Sample Selection

The technical implementation of the actual sample selection differed from country to country, depending on the type of frame and the country-specific circumstances. For each country, every sampling unit had a known probability of selection and clearly defined sample selection procedures were followed. Non-random sample selection was not allowed at any stage of the PRIDI Phase III sampling process.

The first stage of sample selection concerned the selection of the PSUs. The procedure for selecting PSUs was systematic sampling from lists sorted by PSU size. The selection probabilities of the PSUs were proportional to their size (PPS sampling). This method increased the number of highly populated PSUs in the sample, thus reducing travel costs, while at the same time ensuring similar selection probabilities for selected children. The following criteria were used in defining the size of a PSU:

- Costa Rica: number of children per PSU according to census (0-2 years in 2011)
- Nicaragua: number of households per PSU
- Paraguay: number of households per PSU
- Peru: number of occupied households per PSU

SSU sampling in Peru was performed with equal selection probabilities. In the other three countries, all households in a PSU were selected.

Support from the National Statistical Institutes and National Coordinators

A number of issues surfaced in relation to sampling. In all cases, the consulting firms responsible for the data collection had difficulties in obtaining the most recent census data from the National Statistical Institutes to aid in the creation of the sampling frames. The national coordinators helped, using their good offices with the Statistical Institutes to get projections and/or more updated data. IEA provided targeted technical assistance directly to the countries.

In Costa Rica, census data for 2011 were used to draw the Phase III sampling frame, increasing the efficiency of implementation (to 12%, up from 6% in Phase II). In the case of Paraguay, the sampling frame was drawn by the National Statistical Institute, but the source data came from the preliminary results of the 2012 census. These data did not contain the level of disaggregation necessary (household members) to correctly identify households with children in the PRIDI age-range. Enumerators had to use maps to find potential PRIDI homes, which led to considerable delays and additional costs in identifying the children necessary to comply with the intended sample. As a result, and in coordination with IEA, the sample size in Paraguay was reduced to 1,500 children and two different samples were used, both of which were drawn following the methodology of primary and secondary sampling units outlined above. The first sample included 900 children and the second, the remaining 600.

Despite these issues, nationally representative samples, with the exclusions mentioned above and acceptable sampling errors were generated in each country. IEA validated each sample.

8. Sampling Weights and Variance Estimation

IEA created the sampling weights to be used in the analyses of all data. These weights are important for obtaining accurate, precise and internationally comparable estimates of population characteristics. Un-weighted population parameter estimates present an incorrect picture of the situation in the participating countries as they over-emphasize the impact of the following sub-regions on national estimates:

- the Cusco Region in Peru,
- the Miskitu population in Nicaragua,
- the urban areas in Paraguay,
- the rural areas outside of the Valle Central in Costa Rica.

Calculating sampling weights

The PRIDI child weight is a product of base weights and non-response adjustments. Base weights reflect the selection probabilities of PSUs and SSUs and, at each level of sample selection, are the inverse of the selection probability of a sampled unit. Non-response adjustments compensate for potential bias due to non-participation of sampled units.

PSU base weight (WGTFAC1)

The first stage of sampling for PRIDI was the selection of Primary Sampling Units (PSUs) in each country. The PSU base weight reflects the selection probabilities of this sampling step. The PSUs samples were selected independently in each explicit stratum h , with $h=1, \dots, H$.

Systematic sample of PSUs were drawn, with the selection probability of PSU i being proportional to its PSU size (*PPS sampling*). The PSU base weight is defined as the inverse of the PSU's selection probability. For PSU i in stratum h , the PSU base weight is given by:

$$WGTFAC1_{hi} = \frac{M_h}{n_h^s \times M_{hi}}$$

where n_h^s is the number of sampled PSUs in stratum h , M_h is the total measure of size in the PSUs of explicit stratum h , and M_{hi} is the measure of size of the selected PSU i .

PSU non-response adjustment (WGTADJ1)

In Paraguay, some PSUs dropped out of the sample because no child from the geographical area participated in the study. Therefore, it was necessary to adjust the PSU base weights to account for the sample size loss. Adjustments were calculated within non-response groups

defined by the explicit strata. Within each explicit stratum, a PSU non-response adjustment was calculated for each participating PSU i in stratum h as:

$$WGTADJ1_{hi} = \frac{n_h^s}{n_h^p}$$

where n_h^s is the number of sampled eligible PSUs and n_h^p is the number of participating PSUs in explicit stratum h . For the other three countries, the value of WGTADJ1 equals 1.

SSU base weight (WGTFAC2)

In each participating PSU in Peru, a number of Secondary Sampling Units (SSUs) - usually residences (*viviendas*) - were randomly selected using a random selection method with equal selection probabilities for each SSU. The SSU base weight is the inverse of the selection probability in this sampling step.

For each sampled SSU j , the SSU base weight is given by

$$WGTFAC2_{hij} = \frac{V_{hi}}{v_{hi}^s}$$

where V_{hi} is the total number of SSUs and v_{hi}^s is the number of sampled SSUs in PSU i in stratum h . For the other three countries, the value of WGTFAC2 equals 1, since all SSUs in sampled PSUs were selected for the study.

SSU non-response adjustment (WGTADJ2)

If a sampled SSUs did not participate in PRIDI, a non-response adjustment was computed. The SSU weight adjustment was calculated for each participating SSU j as

$$WGTADJ2_{hij} = \frac{v_{hi}^s}{v_{hi}^p}$$

where v_{hi}^s is the total number of sampled SSUs and v_{hi}^p is the total number of participating SSUs in PSU i in explicit stratum h .

Child base weight (WGTFAC3)

In case two or more PRIDI-eligible children were found in SSU k , one of them was randomly selected for the study. For weight calculation, the weights were multiplied with the number of PRIDI-eligible children in the SSU c_{hijk}

$$WGTFAC3_{hijk} = c_{hijk}$$

Correction factor (CORRFAC)

In few cases, the proportions regarding some key variables did evidently not match with population statistics. In these cases, it was decided to apply a correction factor to compensate for the imbalance. In two cases, the estimates were adapted based on census data and information from the statistical agencies:

- In Costa Rica, the preliminary weights showed an under-representation of the “Valle Central / Urban” category and an over-representation of “Rest of Country / Rural category”, which was corrected.
- In Nicaragua, the weights of Miskitu children increased to reflect the proportion of this language group in the population. Further, the “Managua/Urban” region was under-represented in the original weights, while “Chinandega/Urban” and “Chontales/Urban” were over-represented; this was corrected.

Final child weight (TOTWGT)

The final child weight of each child k in SSU j of PSU i in stratum h is the product of the six weight components:

$$TOTWGT_{hijk} = WGTFAC1_{hi} \times WGTADJ1_{hi} \times WGTFAC2_{hij} \times WGTADJ2_{hij} \times WGTFAC3_{hijk} \times CORRFAC_h$$

Estimating Sampling Variance

In PRIDI, the sample of children is not a simple random sample, but the result of a stratified multi-stage cluster sample. The use of this sampling design has an impact on the sampling variance:

- The children in PRIDI were selected using clustered samples. PSUs were not individuals, but geographical areas. Due to socio-demographic characteristics, children were more likely to be similar to each other within PSUs than across PSUs.
- Stratification limits the possibility for atypical samples and decreases sampling variance. All samples in PRIDI were drawn using explicit stratification variables. Within each country, independent samples were drawn among strata. Systematic sampling from a list sorted by PSU size further reduced the overall standard errors.

Both of these effects must be considered for estimating the standard errors to be used for forming confidence intervals and hypothesis testing.

Statistical software packages do not take the effects of clustering and stratification into account. The standard errors they calculate thus are too small. The Jackknife Repeated Replication (JRR) technique, implemented through the application of the IEA IDB Analyzer (see PRIDI User Guide), takes the effects of clustering and stratification into account and provides correct estimators for standard errors of the population parameter estimates.

The use of JRR requires systematically assigning pairs of PSUs to sampling zones, and randomly selecting one of these PSUs to have its contribution doubled and the other to have its contribution zeroed in corresponding sets of replicate weights. The statistic of interest is computed once for the original sample, and once again for each replicated sample. The variation between estimates for each of the replicated samples and the original sample estimate is the jackknife estimate of the sampling error of the statistic.

To apply the JRR technique used in PRIDI, the sampled PSUs had to be paired and assigned to a series of groups known as sampling zones. This was done by working through the list of sampled PSUs in the order in which they were selected and assigning the first and second PSUs to the first sampling zone, the third and fourth PSUs to the second zone, and so on. A maximum of 100 zones were formed in total. When 100 zones were completed, the process was continued by assigning the next pair of PSUs to the first sampling zone, then the next pair to the second sampling zone, and so on.

To compute a statistic t from the sample for a PRIDI country, the formula for the JRR variance estimate of the statistic t is

$$Var_{JRR}(t) = \sum_{h=1}^{100} [t(J_h) - t(S)]^2$$

The term $t(S)$ corresponds to any statistic for the whole sample using full sample weights; the element $t(J_h)$ denotes the same statistic using the h^{th} set of replicate weights.¹⁸

¹⁸ This formula holds for most statistics, and it holds for regression coefficients. All coefficients from the regression model will have to be "Jackknifed." The regression model has to be estimated for all 100 replicates; the standard errors of the coefficients then have to be calculated using this formula.

9. Scaling the PRIDI Data

The Cognitive, Motor and Language Sub-Scales

To achieve its goal of coverage of two age groups, the PRIDI instruments included a range of items that covered a range of tasks that the child was asked to perform. Depending on the age, the child was administered one of two forms. These forms have overlapping items. PRIDI relied on Item Response Theory (IRT) scaling to combine the children's responses and provide accurate estimates of proficiency on each of these scales. This section describes the process and the outcomes of scaling the data from the cognitive, motor and language scales.

The Engle Scale focuses on three major domains of skills in children's development: cognitive, motor and language skills. Each of the two forms (Form A and Form B) contains a set of common items, as well as a set targeted to the corresponding age group. When scaling the items, the common items were used to create a single scale for both forms, hence their denomination as anchor items.

The answers to the items were scored correct or incorrect, in the case of dichotomous items, or correct, partially correct, or incorrect in the case of polytomous scored items. Tables 17 and 18 provide a list of the items in test forms A and B, and their scoring scheme. Common items are identified across forms because they have the same number. For example, items PNA03 and PNB03 are the same item, whereas items PNA01Z and PNB01Z are not the same item.

Table 17: Items in Form A and Scoring Categories

Item Name	Source items	Domain	Common item	Categorized Sum of the scores of combined items	Final Score Categories
PNA01Z	PNA01A, PNA01B	Cognitive		2; 0-1	1; 0
PNA02	-	Cognitive		1; 0	1; 0
PNA04Z	PNA04A, PNA04B, PNA04C	Cognitive		3; 2; 0-1	2; 1; 0
PNA06	-	Cognitive		1; 0	1; 0
PNA10Z	PNA10A, PNA10B	Cognitive	X	10-12; 6-8; -12-4	2; 1; 0
PNA11Z	PNA11A, PNA11B, PNA11C	Cognitive	X	3; 1-2; 0	2; 1; 0
PNA13Z	PNA13A, PNA13B, PNA13C	Cognitive	X	5-6; 2-4; 0-1	2; 1; 0
PNA14	-	Cognitive	X	1; 0	1; 0
PNA19Z	PNA19A, PNA19B	Cognitive	X	2; 1; 0	2; 1; 0
PNA03	-	Motor	X	2; 1; 0	2; 1; 0
PNA15Z	PNA15A, PNA15B, PNA15C	Motor		2-3; 1; 0	2; 1; 0
PNA16Z	PNA16A, PNA16B	Motor	X	3-4; 2; 0-1	2; 1; 0
PNA17Z	PNA17A, PNA17B, PNA17C	Motor	X	2-3; 1; 0	2; 1; 0
PNA21A	-	Motor	X	1; 0	1; 0
PNA21Z	PNA21B, PNA21C, PNA21D	Motor	X	3; 1-2; 0	2; 1; 0

PNA05Z	PNA05A, PNA05B, PNA05C, PNA05D, PNA05E	Language		3-5; 1-2; 0	2; 1; 0
PNA07	-	Language		1; 0	1; 0
PNA08Z	PNA08A, PNA08B	Language	X	2; 0-1	1; 0
PNA09Z	PNA09A, PNA09B	Language		2; 0-1	1; 0
PNA12Z	PNA12A, PNA12B, PNA12C, PNA12D, PNA12E, PNA12F, PNA12G, PNA12H	Language	X	6-8; 2-5; 0-1	2; 1; 0
PNA18Z	PNA18A, PNA18B, PNA18C	Language	X	2-3; 1; 0	2; 1; 0
PNA20Z	PNA20A, PNA20B	Language		2; 1; 0	2; 1; 0

Table 18: Items in Form B and Scoring Categories

Item Name	Source item	Domain	Common item	Categorized Sum of the scores of the combined items	Final Score Categories
PNB01Z	PNB01A, PNB01B, PNB01C	Cognitive		4-6; 1-3; 0	2; 1; 0
PNB02Z	PNB02A, PNB02B, PNB02C, PNB02D, PNB02E, PNB02F	Cognitive		5-6; 2-4; 0-1	2; 1; 0
PNB04	-	Cognitive		1; 0	1; 0
PNB05Z	PNB05A, PNB05B	Cognitive		2; 1; 0	2; 1; 0
PNB07Z	PNB07A, PNB07B	Cognitive		14-16; 8-12; -16-6	2; 1; 0
PNB10Z	PNB10A, PNB10B	Cognitive	X	10-12; 6-8; -12-4	2; 1; 0
PNB11Z	PNB11A, PNB11B, PNB11C	Cognitive	X	3; 1-2; 0	2; 1; 0
PNB13Z	PNB13A, PNB13B, PNB13C	Cognitive	X	5-6; 2-4; 0-1	2; 1; 0
PNB14	-	Cognitive	X	1; 0	1; 0
PNB15	-	Cognitive		1; 0	1; 0
PNB19Z	PNB19A, PNB19B	Cognitive	X	2; 1; 0	2; 1; 0
PNB03	-	Motor	X	2; 1; 0	2; 1; 0
PNB16Z	PNB16A, PNB16B	Motor	X	3-4; 2; 0-1	2; 1; 0
PNB17Z	PNB17A, PNB17B, PNB17C	Motor	X	2-3; 1; 0	2; 1; 0
PNB21A		Motor	X	1; 0	1; 0
PNB21Z	PNB21B, PNB21C, PNB21D	Motor	X	3; 1-2; 0	2; 1; 0
PNB22Z	PNB22A, PNB22B, PNB22C	Motor		2-3; 1; 0	2; 1; 0
PNB06Z	PNB06A, PNB06B, PNB06C, PNB06D, PNB06E, PNB06F	Language		3-6; 1-2; 0	2; 1; 0
PNB08Z	PNB08A, PNB08B	Language	X	2; 0-1	1; 0
PNB09Z	PNB09A, PNB09B	Language		2; 0-1	1; 0
PNB12Z	PNB12A, PNB12B, PNB12C, PNB12D, PNB12E, PNB12F, PNB12G, PNB12H	Language	X	6-8; 2-5; 0-1	2; 1; 0
PNB18Z	PNB18A, PNB18B, PNB18C	Language	X	2-3; 1; 0	2; 1; 0
PNB20Z	PNB20A, PNB20B, PNB20C	Language		2-3; 1; 0	2; 1; 0

As part of the quality control for the scaling, several steps were undertaken. First we computed descriptive statistics for each of the items¹⁹. Tables 19 and 20 present descriptive statistics for the items in form A and B. The tables also contain the minimum and maximum scores for each item. The means for dichotomous items represent the proportions of the children that answered the items correctly. The means for the polytomous items represent the average score on the item. A mean closer to the maximum score indicates an easy item. A mean closer to zero indicates a difficult item.

Table 19: Descriptive Statistics of the Items in Form A

Item Name	N	Minimum	Maximum	Mean	Std. Deviation
PNA01Z	3995	0	1	.24	.425
PNA02	3968	0	1	.56	.496
PNA04Z	3993	0	2	.79	.899
PNA06	3986	0	1	.38	.485
PNA10Z	4000	0	2	.61	.824
PNA11Z	3986	0	2	.49	.608
PNA13Z	3968	0	2	.50	.556
PNA14	3966	0	1	.27	.443
PNA19Z	8000	0	2	.06	.272
PNA03	3993	0	2	.44	.644
PNA15Z	3979	0	2	1.19	.923
PNA16Z	3994	0	2	.48	.695
PNA17Z	3984	0	2	.85	.876
PNA21A	3945	0	1	.25	.434
PNX21Z	7881	0	2	.33	.577
PNA05Z	3995	0	2	.80	.739
PNA07	3971	0	1	.45	.497
PNA08Z	3989	0	1	.24	.426
PNA09Z	3989	0	1	.47	.499
PNA12Z	3989	0	2	.45	.582
PNA18Z	3982	0	2	.50	.760
PNA20Z	3976	0	2	.84	.676

¹⁹ Statistics presented in this chapter are weighted so that each country contributes equally to the estimates, and the sum of the weights is 4,000 by form. Total sample size across all 4 participating countries was 7710 cases.

Table 20: Descriptive Statistics of the Items in Form B

	N	Minimum	Maximum	Mean	Std. Deviation
PNB01Z	3995	0	2	.59	.772
PNB02Z	3971	0	2	.59	.787
PNB04	3974	0	1	.60	.490
PNB05Z	4000	0	2	.69	.740
PNB07Z	4000	0	2	.66	.837
PNB10Z	4000	0	2	1.28	.900
PNB11Z	3996	0	2	1.08	.648
PNB13Z	3992	0	2	1.15	.595
PNB14	3984	0	1	.40	.490
PNB15	3975	0	1	.22	.412
PNB19Z	3993	0	2	.59	.756
PNB03	3999	0	2	1.26	.788
PNB16Z	3984	0	2	1.30	.764
PNB17Z	3975	0	2	1.54	.735
PNB21A	3938	0	1	.71	.455
PNX21Z	7881	0	2	.33	.568
PNB22Z	4000	0	2	.47	.758
PNB06Z	3994	0	2	1.27	.762
PNB08Z	3993	0	1	.62	.484
PNB09Z	3993	0	1	.35	.478
PNB12Z	3997	0	2	1.22	.616
PNB18Z	3997	0	2	1.32	.855
PNB20Z	3991	0	2	1.27	.832

Tables 21 and 23 present the Cronbach's Alpha reliability coefficient for each of the scales. Tables 22 and 24 present the respective item discrimination indices. Both are measures of the internal consistency of a given scale and provide an indication of how closely related a set of items are as a group. The Cronbach's Alpha, for all scales, are greater than 0.6, an international benchmark for acceptability.

The corrected item-total correlation provides an indication of how strong the relationship is between a single item and the scale formed by all other items in the test. It is also called "discrimination index" because it shows how well the item can discriminate between respondents of different abilities. A correlation value less than 0.2 indicates that the corresponding item does not correlate very well with the scale overall and should be dropped.

Table 21: Cronbach's Alpha Reliability Coefficients for Scales in Form A

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	Number of Items
------------------	--	-----------------

Cognitive	.684	.694	9
Motor	.660	.676	6
Language	.756	.759	7

Table 22: Item-Total Correlations by Scale in Form A

Cognitive Scale	Corrected Item-Total Correlation
PNA01Z	.336
PNA02	.269
PNA04Z	.461
PNA06	.464
PNA10Z	.369
PNA11Z	.447
PNA13Z	.422
PNA14	.228
PNA19Z	.330

Motor Scale	Corrected Item-Total Correlation
PNA03	.420
PNA15Z	.483
PNA16Z	.473
PNA17Z	.424
PNA21A	.336
PNX21Z	.293

Language Scale	Corrected Item-Total Correlation
PNA05Z	.560
PNA07	.477
PNA08Z	.367
PNA09Z	.388
PNA12Z	.573
PNA18Z	.466
PNA20Z	.525

Table 23: Cronbach's Alpha Reliability Coefficients for Scales in Form B

	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	Number of Items
Cognitive	.761	.766	11
Motor	.597	.605	6
Language	.700	.702	6

Table 24: Item-Total Correlations in Form B

Cognitive Scale	Corrected Item-Total Correlation
PNB01Z	.546
PNB02Z	.586
PNB04	.410
PNB05Z	.423
PNB07Z	.388
PNB10Z	.290
PNB11Z	.436
PNB13Z	.390
PNB14	.231
PNB15	.376
PNB19Z	.503
Motor Scale	Corrected Item-Total Correlation
PNB03	.335
PNB16Z	.349
PNB17Z	.257
PNB21A	.295
PNX21Z	.416
PNB22Z	.368
Language Scale	Corrected Item-Total Correlation
PNB06Z	.571
PNB08Z	.366
PNB09Z	.276
PNB12Z	.521
PNB18Z	.516
PNB20Z	.381

Items are calibrated onto a single scale using a one parameter IRT model where the probability of a response is modeled as a function of the difficulty of the item, and the ability of the person. The discrimination parameter was fixed at 1, and scores were calculated using a weighted maximum likelihood estimation procedure. Calibration and initial assignment of scores was conducted using Parscale 4.0. For more information about the software used, and a description of IRT models (see du Toit,2003).

For the purpose of estimating the item difficulties, sampling weights were used in such a way that each country contributed equally to the difficulty of the items, and within each country, students from each of the two age groups (Form A and Form B) also contributed equally. The estimated difficulty parameter

estimates and their standard errors are presented in the following table, together with step parameters in the case of 2 point items (Tables 25, 26 and 27).

Table 25: Item Parameters in the Cognitive Scale

Item Name	Difficulty	(s.e.)	step1	step2
PNA01Z	0.484	0.024		
PNA02	-0.591	0.022		
PNA04Z	-0.138	0.014	-0.350	0.350
PNA06	-0.019	0.022		
PNB01Z	0.974	0.015	0.051	-0.051
PNB02Z	0.957	0.015	-0.026	0.026
PNB04	0.102	0.021		
PNB05Z	0.873	0.015	0.400	-0.400
PNB07Z	0.841	0.014	-0.184	0.184
PNB15	1.353	0.025		
PNX10Z	0.087	0.010	-0.369	0.369
PNX11Z	0.406	0.012	0.745	-0.745
PNX13Z	0.355	0.012	0.890	-0.890
PNX14	0.555	0.016		
PNX19Z	1.084	0.013	0.115	-0.115

Table 26: Item Parameters in the Motor Scale

Item Name	Difficulty	(s.e.)	step1	step2
PNA15Z	-0.730	0.014	-0.531	0.531
PNB22Z	1.237	0.016	-0.245	0.245
PNX03	0.255	0.011	0.260	-0.260
PNX16Z	0.183	0.011	0.239	-0.239
PNX17Z	-0.280	0.011	-0.088	0.088
PNX21A	0.078	0.016		
PNX21Z	1.370	0.015	0.528	-0.528

Table 27: Item Parameters in the Language Scale

Item Name	Difficulty	(s.e.)	step1	step2
PNA05Z	-0.113	0.016	0.600	-0.600
PNA07	-0.295	0.022		
PNA09Z	-0.354	0.022		
PNA20Z	-0.143	0.018	0.868	-0.868
PNB06Z	0.034	0.016	0.410	-0.410
PNB09Z	0.967	0.023		
PNB20Z	0.095	0.015	0.061	-0.061
PNX08Z	0.255	0.016		
PNX12Z	0.342	0.012	0.792	-0.792
PNX18Z	0.156	0.011	-0.158	0.158

After calculating scores for the children, using the item parameters presented above, the scores were standardized and placed on a more useful metric with mean 50 and standard deviation 5. The mean 50 and standard deviation 5 is obtained when combining all countries (4), and age groups (2), and using sampling weights that take into account the sampling within the country, but also equalize the contribution of each of these 8 groups. This standardization was done separately for each scale.

The standardization was done applying the following formula:

$$SC_i = \frac{X_i - A}{B} \times 5 + 50,$$

where

SC_i is the scaled score for an individual scale (Cognitive, Motor or Language)

X_i is the original child's score on the scale

A is the average of the original scores on the scale

B is the standard deviation of the original scores on the scale

Values for A and B were the following:

Scale	A	B
Cognitive	0.03275	0.79141
Motor	0.03322	0.87222
Language	0.00027	0.97741

Scaling the PRIDI Socio-Emotional Scale

The PRIDI Socio-emotional background scale is composed of multiple-choice questions regarding the children's social and emotional development. The questions are answered by the parent or guardian, and are related to children's relationships with other children and adults, willingness to participate in different activities with others, independence in certain daily routines and the way they react in certain situations.

The questions had four response categories: almost never (1); sometimes (2); often (3); and almost always (4). Preliminary statistical analysis showed that respondents did not differentiate the categories "almost never" from "sometimes." For the purpose of analyzing this scales categories 1 and 2 were collapsed into a single category.

In addition, preliminary statistical analysis showed that question number 7 (PNG26G) asking whether the child cries when the parent or guardian leaves, did not correlate with the rest of the items, and consequently did not correlate with the total score on the scale. This item was therefore removed for the purpose of scaling and computing scores for this scale.

As was done with the achievement items, the data for the socio-emotional scale was scaled using Partial Credit Modeling, where the discrimination parameters for all items are fixed to 1. Table 28 presents descriptive statistics for the items in the socio-emotional scale.

Table 28: Descriptive Statistics for the Socio-Emotional Scale Items

Item Name	N	Minimum	Maximum	Mean	Std. Deviation
PNG26A	7950	1	3	2.29	.825
PNG26B	7969	1	3	2.37	.776
PNG26C	7920	1	3	2.02	.886
PNG26D	7930	1	3	1.91	.879
PNG26E	7930	1	3	2.17	.861
PNG26F	7910	1	3	2.08	.868
PNG26H	7919	1	3	2.16	.859
PNG26I	7957	1	3	2.00	.916
PNG26J	7953	1	3	2.17	.866
PNG26K	7942	1	3	2.30	.818
PNG26L	7954	1	3	2.18	.868
PNG26M	7913	1	3	1.99	.901
PNG26N	7950	1	3	2.07	.867
PNG26NE	7959	1	3	2.00	.868
PNG26O	7959	1	3	2.37	.804

Following the review of the descriptive statistics, reliability analysis was conducted. Tables 29 and 30 show item discrimination and reliability coefficients for the scale.

Table 29: Cronbach's Alpha Reliability Coefficient for the Socio-Emotional Scale

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.846	.847	15

Table 30: Item-Total Correlations in the Socio-Emotional Scale

Item Name	Corrected Item-Total Correlation
PNG26A	.512
PNG26B	.491
PNG26C	.504
PNG26D	.396

PNG26E	.530
PNG26F	.378
PNG26H	.467
PNG26I	.367
PNG26J	.572
PNG26K	.599
PNG26L	.480
PNG26M	.481
PNG26N	.509
PNG26NE	.400
PNG26O	.448

Items were then calibrated onto a single scale using a one parameter IRT model where the probability of a response is modeled as a function of the difficulty of the item, and the ability of the person. The discrimination parameter was fixed at 1, and scores were calculated using a weighted maximum likelihood estimation procedure. Calibration and initial assignment of scores was conducted using Parscale 4.0. For more information about the software used, and a description of IRT models, please refer to du Toit (2003).

For the purpose of estimating the item difficulties, sampling weights were used in such a way that each country contributed equally to the difficulty of the items, and within each country, students from each of the 2 age groups of interest also contributed equally. The estimated difficulty parameter estimates and their standard errors are presented in the following table, together with step parameters (Table 31).

Table 31: Item Parameters in the Socio-Emotional Scale

Item Name	Difficulty	(s.e.)	step1	step2
PNG26A	-0.39019	0.0097	-0.05369	0.05369
PNG26B	-0.51179	0.0102	0.07829	-0.07829
PNG26C	-0.06093	0.00919	-0.16815	0.16815
PNG26D	0.06892	0.00928	-0.14284	0.14284
PNG26E	-0.24018	0.00938	-0.10313	0.10313
PNG26F	-0.12940	0.00934	-0.07516	0.07516
PNG26H	-0.21889	0.00940	-0.07457	0.07457
PNG26I	-0.03260	0.00896	-0.37744	0.37744
PNG26J	-0.23277	0.00933	-0.12489	0.12489
PNG26K	-0.40225	0.00977	-0.02681	0.02681
PNG26L	-0.23967	0.00931	-0.14575	0.14575
PNG26M	-0.02343	0.00909	-0.26597	0.26597
PNG26N	-0.12257	0.00932	-0.06767	0.06767
PNG26NE	-0.03858	0.00933	-0.05558	0.05558
PNG26O	-0.48954	0.00993	-0.08506	0.08506

As with the other scales, after calculating scores for the children using the item parameters presented above, the scores were standardized and placed on a more useful metric with mean 50 and standard deviation 5. The mean 50 and standard deviation 5 is obtained when combining all countries (4), and age groups (2), and using sampling weights that take into account the sampling within the country, but also equalize the contribution of each of these 8 groups.

The standardization was done applying the following formula:

$$SC_i = \frac{X_i - A}{B} \times 5 + 50,$$

where

SC_i is the scaled score for a child

X_i is the original child's score on the scale

A is the average of the original scores on the scale

B is the standard deviation of the original scores on the scale

Values for A and B were the following:

Scale	A	B
Socio-Emotional	-0.03774	0.588388

.

10. Creating and Confirming the PRIDI International Database

This chapter describes the procedures for checking the PRIDI data files and creating the PRIDI International Database that were implemented by the IEA Data Processing and Research Center (DPC) and the national centers of participating countries. The main purpose of these procedures was to ensure that

- All information in the database conformed to the internationally defined data structure
- The content of codebooks and documentation reflected the national adaptations to the background instruments
- All variables used for international comparison were comparable across countries

Confirming the integrity of the PRIDI International Database

The quality assurance of the national data files comprised several steps. First, the IEA DPC checked the database files submitted by each country. Standard cleaning rules were applied to verify the accuracy and consistency of the data, its structure and its accompanying documentation. Deviations of the international structure were documented and queries were addressed to the national centers.

Modifications to the data files were made as necessary. Once all modification had been applied, all data files were processed and checked again. This process, referred to as 'data cleaning' was repeated as many times as necessary until all data were consistent and comparable within and across countries. When the national files had been checked, the IEA DPC created national and international statistics for background variables and test items.

After reviews of item-statistics, univariates, and certain analysis result (such as factor and reliability analyses) had been completed, the IEA DPC calculated the PRIDI language, motor, and numeracy scores as well as an international PRIDI average score. Weights and scores, as well as certain indices derived from the original data, were included into the final PRIDI International Database (see Chapters 8 and 9).

Data Checking at the IEA Data Processing and Research Center

As described in Chapter 6 in Part A of this report, national center staff members in the four participating countries were responsible for entering their national PRIDI data into the appropriate data files and submitting these files to the IEA DPC. Staff at the IEA DPC then subjected these files to a comprehensive process of checking and editing. To facilitate the data cleaning process, the IEA DPC asked the national centers to provide them with detailed documentation of their data together with their national data files. The data documentation included the national versions of the test booklets (A & B) and questionnaires, as well as their final national adaptation forms (NAFs) containing the documentation on all national adaptations.

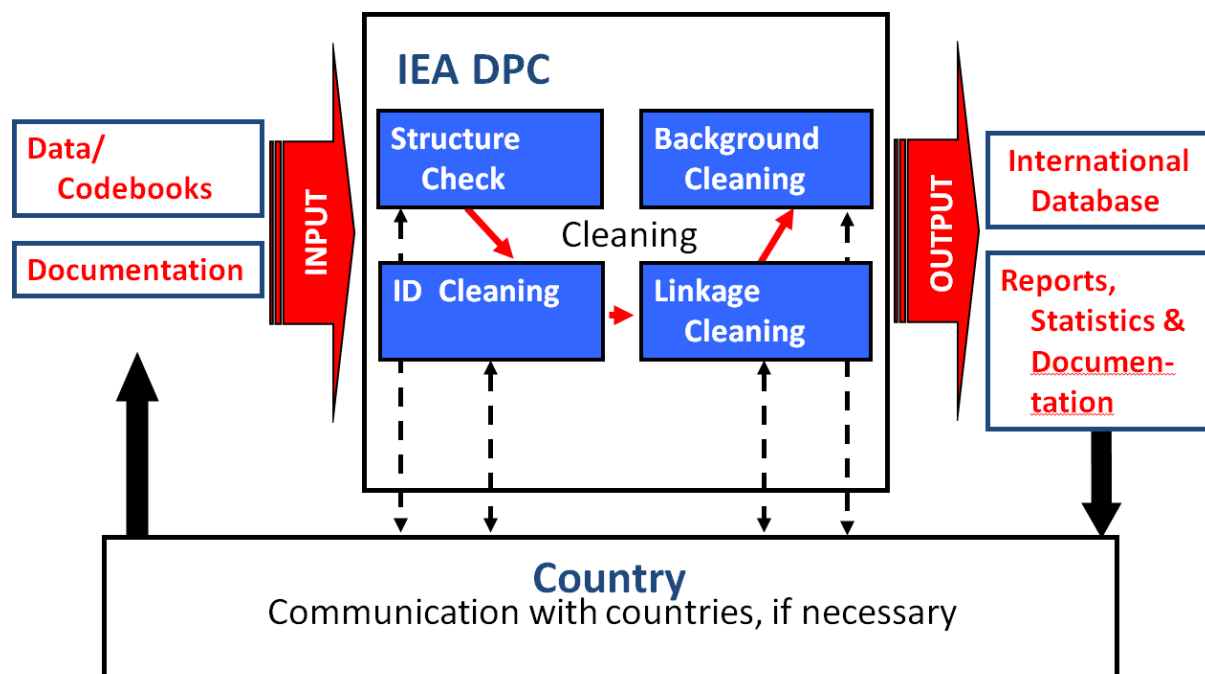
Preparing National data files

The main objective of the data-cleaning process was to ensure that the data adhered to the international format, that family and child information could be linked between different survey files, and that the data collected within each country in an accurate and consistent manner.

The SAS-based data cleaning consisted of the following steps that will be explained in more detail in the following sections:

- Documentation and structure checks
- Identification variable (ID) cleaning
- Linkage cleaning
- Background questionnaire cleaning

Figure 1 Overview of Data Processing at the DPC



Documentation and Structure Checks

For each country, data cleaning began with an exploratory review of its data-file structure and its data documentation. At the beginning of the process, the data files from the DME software containing the survey instrument data were imported to the SAS based cleaning system.

The first checks implemented identified differences between the international file structure and the national file structures in order to detect any structural adaptations (such as adding national variables or omitting international questions) implemented by the participating countries. To keep track of adaptations, the IEA DPC asked the national centers to complete national adaptation forms (NAFs) while they were adapting their national codebook structure. Where necessary, the IEA DPC modified the

structure of the country's data to ensure that the resulting data remained comparable between countries.

After each data file matched the international standard, a series of standard cleaning rules were applied to the files using a SAS software program developed by IEA DPC staff. This software could identify and, in many cases, correct inconsistencies in the data. Each problem was listed in a country specific report identified by a unique problem number.

Where problems could not be corrected automatically, they were reported to the responsible national team so that the original data collection instruments and tracking forms could be checked to trace the source of the errors. If a national center could not solve problems through verification of instruments with the forms even with the assistance of the IEA DPC team, IEA DPC applied a general cleaning rule to rectify these errors. After all of the automatic updates had been applied, IEA DPC used SAS recoding scripts to directly apply any remaining corrections to the data files.

Identification Variable (ID Cleaning)

While it already was assured by the data entry software DME that each record in the PRIDI data files had a unique identification number, the PRIDI ID cleaning rather focused on resolving inconsistencies between child participation status, data availability in the different survey instruments, and exclusion status information. In addition, the IEA DPC calculated the child's age based on child's birth dates and testing dates and checked for the correct assignment of the child test forms (Form A for children between 2 and 3.5 years and Form B for children between 3.5 years and 5 years) and the TVIP test (to be assigned for children between 2.5 and 5 years). Few cases of wrong assignments of test instruments that could not be corrected with help of the national centers were removed from the final International Database.

Linkage Checks

In PRIDI, data about children and their families appeared in different datasets, so it was crucial that the records from these files link together correctly to provide meaningful data for analysis and reporting. The linkage was implemented through a hierarchical ID numbering system incorporating a Primary Sampling Unit (PSU), family, and child component and it was assured that child's entries in the test data sets and background data sets were matched correctly and that children were assigned correctly to their families within their primary sampling units.

Engle Scale Results: Data Checks

In some items (for example PNA10A), number of correct responses and number of failures should add up to the number of trials. If inconsistencies between variables could not be solved in collaboration with the national center, both variables were coded to "Omitted".

For some questions (for example PNA01A) test administrator instructions requested the test administrator to skip an additional, more difficult part of an item if the child did not manage to answer correctly to the first part. In those cases the second part of the cases automatically was coded to 'incorrect' (0 score points) during the cleaning process.

Survey: Data Checks

The number of inconsistent and implausible responses in the background files varied, but none of the national datasets was completely free of inconsistent responses. Treatment of these responses was determined on a question-by-question basis, using available documentation to make an informed decision. Implausible values or values out-of-range that could not be rectified with help of the national centers were recoded to "Omitted" for the final PRIDI International Database.

National Cleaning Documentation

The IEA DPC provided National Centers with a report of all problems that were identified in their data and listed records of deviations from the international data collection instruments. Additionally, the IEA DPC provided the PRIDI Management Team with revised data files incorporating all agreed-upon edits, updates, and structural modifications, as well as list of new variables (such as the child's age) that could be used for analytic purposes.

Handling of Missing Data

Two types of entries were possible during the PRIDI data capture: valid data values, and missing data values. Missing data can be assigned a value of omitted (I don't know and refused), or not administered (including errors in administration) during data entry. The IEA DPC applied additional missing codes to the data to facilitate further analyses. This process led to five different types of missing data being distinguished in the International Database:

- Refused: the respondent had a chance to answer the question but refused to do.
- Don't know (background data only): the respondent had a chance to answer the question but answered that he doesn't know the response.
- Not administered/ Error in Administration: the respondent was not administered the actual item or question or the item was not applicable to the respondent.
- Not reached (only used in the test files): This code indicates those items not reached by the children due to a lack of time (assigned during data processing only).

As a preparation for the scaling process, several items have been combined to so-called derived variables which can be identified by having letter 'Z' at the end of the variable name. For all these variables the following rule concerning coding of missing variables was applied: If at least one of the included item-parts showed valid values (including the refused missing code), then it was assumed that

the complete item was administered to the children and consequently a score was assigned to the corresponding derived variable.

The PRIDI International Database

The PRIDI International Database incorporated the national data files from the four participating countries. The data files are available in SPSS format and are accompanied by data documentation and international variable almanacs.

Data processing at the IEA DPC ensured that:

- Information coded in each variable was internationally comparable.
- National adaptations were reflected appropriately in all variables.
- Questions that were not internationally comparable were removed from the database.
- All entries in the database could be linked to the appropriate respondent
- Sampling weights and child achievement scores were available for international comparisons.

More information about the PRIDI International Database is provided in the PRIDI User Guide for the International Database.

Annex I
Instruments Commonly Used in PRIDI Countries to Evaluate Child Development (2010)

Test	País	Institución	Año de construcción	Edades	Propósito y tipo de medida	Esferas	Descripción Y aplicación	Validación
EDIN	Costa Rica	Caja Costarricense de Seguro Social (al principio Salud Mental)	1987 primera edición	0-6 años	Monitoreo individual – por cada escala mide si el niño está por debajo de, igual a, o por encima de su edad (no está en el manual)	6: socio afectiva, motora fina, hábitos y higiene, cognoscitiva, lenguaje	Techo y piso; manual, requiere entrenamiento; número de ítems/área depende; formas para cada mes (0-1 ^a), cada 6 meses (1-3), y cada año (3-6). Mucho más detalla para los primeros años (sólo 18 ítems para 5-6 años); cada edad utiliza ítems diferentes	No hay una descripción de validación en el primero. Parece que está en la versión revisada
EDIN	Nicaragua	Universidad Nacional Autónoma de Nicaragua en la Licenciatura de Psicología General Aldeas SOS		0-72 meses		Áreas de Motora Gruesa, Motora Fina, Cognoscitivo, Lenguaje, Socio Afectiva, Hábitos de Salud y Nutrición		
Denver II	Costa Rica	Caja Costarricense de Seguro Social	1990	0-6 años	Diagnóstico – Prueba de Tamizaje; 3 niveles (normal, sospechosa, o no puede aplicar la prueba). Cada ítem es normal, avanzado, caución, o atrasado	Lenguaje, Motora Gruesa, Motora fina, Social-Adaptativo	125 ítems – empieza a la edad del niño, techo y piso (3 cada dirección), hace una suma de tipo de ítem. Requiere entrenamiento	Tiene confiabilidad; una revisión después del DDST; no tiene validez independiente;
Test de Denver	Nicaragua	Universidad Nacional Autónoma de Nicaragua en la Licenciatura de Psicología General		0-6años		Desarrollo de la psicomotricidad		

Test	País	Institución	Año de construcción	Edades	Propósito y tipo de medida	Esferas	Descripción Y aplicación	Validación
Escala CLAP	Nicaragua	Ministerio de Salud	Si es de CLAP del Uruguay, es el Denver			Motricidad Gruesa, Motricidad Fina, Personal-Social, Audición-Lenguaje		
Escala Abreviada de Desarrollo Infantil (Nelson Ortiz)	Ecuador	Asociación VELNC-RHV (2004) Secretaría Técnica del Frente Social Fondo de Desarrollo Infantil – FODI (2007) Ministerio de Coordinación de Desarrollo Social – Dirección de Gestión de Análisis e Investigación del Sector Social INFA – MIES (2008) El Ministerio Coordinador de Desarrollo Social	(hecho en 1990- informe del 1999)	<5 años	Una valoración global y general – y diagnostico primero Un puntaje total en cada área, y normas por edad para comparar	Motricidad Gruesa, Motricidad Fina- Adaptativa, Audición-Lenguaje, Personal-Social	13 grupos de edad (3,6,9,12,18,24,36,48,60,72, 84, 96) 4 dominios, 3 ítems/dominios = 148 (1 ítem/dominio para <1 mes) 15 – 20 minutos. Techo y piso como el Denver. Ítems viene del Denver y Griffiths. (Vuori-Christensen, Ortiz et al., 1974)	Normas con 16,180 niños <4 en Colombia – no sé si hay normas para lo niños mayores
Escala Abreviada del Desarrollo (Ortiz?)	Nicaragua	Escuela Preescolares		0-5 años		Motricidad Gruesa, Motricidad Fina- Adaptativa, Audición-Lenguaje, Personal-Social	“...si trata fundamentalmente de registrar para cada uno de los ítems si el repertorio en cuestión ha sido observado o no.” (Nelson Ortiz)	
Test Abreviado	Peru	MIMDES		0-4 años		0-2 Años: Lenguaje, Motricidad, Coordinación y Social; 3-4 Años: Lenguaje, Coordinación y Motricidad	50 ítems, 5 ítems por cada Grupo Etario. Los 40 Primeros ítems para Niños Menores de 2 años y los 10 otros para Niños y Niñas de 3-4 Años	
Test Abreviado Peruano	Peru	Ministerio de Salud	2009	0-5 años				Norma de evaluación está en proceso de impresión
Test de	Ecuador	Ministerio de	1986	2-14+	Evaluar el desarrollo	Lenguaje y	125 ítems para evaluar el	Muy

Test	País	Institución	Año de construcción	Edades	Propósito y tipo de medida	Esferas	Descripción Y aplicación	Validación
vocabulario en imágenes Peabody (TVIP)		Coordinación de Desarrollo Social – Dirección de Gestión de Análisis e Investigación del Sector Social El Ministerio Coordinador de Desarrollo Social	Adaptación hispanoamericana	años	de lenguaje de niños y niñas. Tiene promedio de 100 y D.E. de 15. Normas de niños de México y de Puerto Rico.	conceptos	vocabulario de niños y niñas – niño debe escoger entre 4 alternativos.	utilizado
Test Woodcock - Muñoz	Ecuador	Ministerio de Coordinación de Desarrollo Social – Dirección de Gestión de Análisis e Investigación del Sector Social El Ministerio Coordinador de Desarrollo Social	Version R - 1996; Bateria III COG y Bateria III APROV revisión en 2005	2-95 años	Medir el Desarrollo Cognitivo y el Aprovechamiento	Desarrollo Cognitivo	Pruebas de Habilidad Cognitiva Consiste en y Pruebas Estándares y 14 Pruebas Suplementarias 20 pruebas para niños pre-escolares empezando a 2.8 años. El el Habilidad intelectual breve son 3: 1 (comprensión verbal), 5 (formación de conceptos) y 6 (pareo visual 1).	Normas con 1400 niños hablantes de español; confiabilidad y validez
TEPSI (Haeussler, de Chile)	Peru	Ministerio de Salud Mimes	1985 primero; 10me edición	2-5 años	Tamizaje pero tiene un puntaje continuo también con escore continuo estandarizado (Punteo T)	Coordinación y Psicomotricidad Fina, el Lenguaje y la Motricidad Gruesa	16 Ítems de Coordinación y Psicomotricidad Fina, 24 Ítems del Lenguaje, 12 Ítems de la Motricidad Gruesa; tiene que administrar todos los ítems a cada niño;	Con Stanford Binet y también hizo análisis de confiabilidad
Escala Observacion de Desarrollo	Paraguay	Centro de estudios en derechos humanos, niñez y juventud (CENIJU) y Ministerio de Educación y Cultura	TEA ediciones (España), 2008	0-5ª11mes	Medir los estándares – medida continuo; para ver cambios a nivel de una población	Emocional y social; lenguaje; motor; cognición y razonamiento	24 -36ma: 59 items; 37-48m : 70 items; 49-59m: 74 items	No está discutido
Pruebas sin más información y de un país solo Pueden ser para diagnosticar retrasos								
Guía de Preguntas Rápidas	Nicaragua	Ministerio de Salud			Identificar Alteraciones Tempranas en el Desarrollo de Niños		Preguntas a las madres	

Test	País	Institución	Año de construcción	Edades	Propósito y tipo de medida	Esferas	Descripción Y aplicación	Validación
					y Niñas			
Hoja de Seguimiento Atención Integral a la Niñez (AIEPI)	Nicaragua	Ministerio de Salud			Evaluar el Desarrollo de Acuerdo a la Edad e Identifica Alteraciones Fenotípicas		Hoja	
Panel Interactivo	Peru	MIMDES		0-5 años	Evaluación Continua del Progreso del Niño o Niña en Wawawasi		Evaluación Regular del Progreso del Niño y Niña	
Lista de Cotejo	Peru	MIMDES	2007	Menores de 4 Años (0-3 años)	Mide el Plan Curricular del Componente Aprendizaje Infantil Temprano del Programa Nacional Wawa Wasi		Indicadores Distribuidos por Rangos de Edad y por Áreas Evaluadas. Los Indicadores son Dicotómicos	Creado en 2001 y Validado en 2002, 2004, 2007 y Adaptaciones
Guía de Observación de Prácticas de Crianza	Peru	MIMDES			Mide Logros Anualmente para Mejorar la Instrucción del Wawawasi en su componente cultural	Componente de Cultura de Crianza	Mide la Calidad de la Interacción Afectiva entre Padres e Hijos en las Prácticas de Cuidado Integral	
Evaluación del Desarrollo del Lenguaje - Estimulación de las Familias	Peru	Mesa Intersectorial del Desarrollo Infantil	2009	15 a 56 Meses	Unas medidas para incluir en encuestas nacionales	Lenguaje, Calidad del Ambiente	Para incluir en una encuesta en la casa con preguntas para padres solo	No – en proceso
Evaluación de Logros de Aprendizaje	Peru	Unidad de Mediciones del Ministerio de Educación		0-5 años				
EEDP	Peru	Ministerio de Salud		0-2 años	Diagnóstico	Área Motora, Área de Coordinación, Área de Lenguaje, Área Social	150 ítems repartidos en grupos de 10 para cada uno de 15 grupos	
Escala del	Costa Rica	Centro de Estudios	2005	0-12	No tengo			(en proceso)

Test	País	Institución	Año de construcción	Edades	Propósito y tipo de medida	Esferas	Descripción Y aplicación	Validación
Desarrollo de 0 a 12 años		del Niño y Asoc. Roblealto		Años	información			de validación)
Prueba PRUNAPE Nacionalizada	Ecuador	SENPLADES-INEC	2005 y versión de Ecuador 2010	0-6 años	Identificar problemas de desarrollo	Área Motriz Fina y Gruesa, Lenguaje y Personal-Social	79 ítems	Sí El nuevo versión está En proceso
Test Pegboard	Ecuador	Ministerio de Coordinación de Desarrollo Social – Dirección de Gestión de Análisis e Investigación del Sector Social El Ministerio Coordinador de Desarrollo Social				Motora fina?		
Medidas de otros países								
Papalote	México	CONAFE – educación inicial	2003		Evaluación continuo		La promotora marque sí o no cada mes; observaciones	no
ECD readiness	Jamaica	Comisión of ECD	2009				Lista de ítems para padres de contestar	No dice
Secuencias de Desarrollo Infantil Integral	Venezuela	Chilina Leon de Viloria, Universidad Católica Andrés Bello Caracas	2008 (segunda edición)	0-12 años	No tiene puntaje; para describir y para monitoreo de progreso	Motora, física, sexual, afectiva, social, moral, lenguaje, cognitiva; tiene manual y instrucciones	Aplicación individual; persona muy capaz; Los resultados son presentados en un grafico con 8 lados	No dice
Escala de Desarrollo	Guatemala	Lic Cadoret y Hernández para evaluar Hogares Comunitarios	2009 Revisión de EDIN (1984), UNICEF (1987), otros	0-6 años ^a	Monitoreo de niños en Hogares y grupo control; continuo	6: socio afectiva, motora fina, hábitos y higiene, desarrollo del pensamiento, lenguaje	2 ítems por cada edad (o mes) en cada categoría; reglas para techo; manual de instrucciones;	No dice
Medidas del ambiente								
Estudio Socioeconómico de le	Costa Rica	Ministerio de Salud, Dirección General de Nutrición y	2004					

Test	País	Institución	Año de construcción	Edades	Propósito y tipo de medida	Esferas	Descripción Y aplicación	Validación
Familia		Desarrollo Infantil						
Evaluación del Desarrollo del Lenguaje y Estimulación de las pFamilias	Peru	Mesa Intersectorial del Desarrollo Infantil	2009	15 a 56 Meses		Calidad del Ambiente	Preguntas para una encuesta nacional sobre el ambiente	En proceso
Household survey	Jamaica	Early Childhood Commission	2010	0-8				En proceso
Guías y pruebas para niños mas grandes								
Guías curriculares								
Guías para Consejería a Padres y Madres con Deficiencia Visual, Motora, Auditiva, Intelectual	Nicaragua	Asociación de Padres de Familia con hijos con Discapacidad			Información	Causa, Conceptos, Identificación Temprana y Estimulación de Áreas Relacionados		
Guía Mira como de Desarrollo PAININ	Nicaragua	Ministerio de la Familia, Adolescencia y Niñez			Estimulación por Áreas de Desarrollo del Niño y de la Niña		Guía de Actividades	
Valoración del Desarrollo de Niños y Niñas de 6 a 14 Años	Costa Rica	Universidad de Costa Rica, Escuela de Enfermería		6-14				
Guía Portage de Educación Preescolar	Nicaragua	Proyecto de Rehabilitación Basado en la Comunidad – Juigalpa CBM		0-5		Estimulación al Bebe (Visual, Táctil, Auditivo, Motora, Afectiva), Socialización,		

Test	País	Institución	Año de construcción	Edades	Propósito y tipo de medida	Esferas	Descripción Y aplicación	Validación
						Lenguaje, Autoayuda, Desarrollo Motriz		

Annex 2

Original Item Mapping by Age Range (Phase I):

Cognition, Communication and Language, and Socio-Emotional

Dimensión	Indicadores y Fuente	Indicadores Propuestos PRIDI
1. Capacidad de resolver problemas, funcionamiento ejecutivo y conceptos básicos de números, tamaño, forma, etc.		
24-36 meses	<ul style="list-style-type: none"> Coloca un cubo sobre la mesa y debajo de la mesa (o silla).^{2 6} Hace torre de cuatro a seis cubos^{2 6} Construye una torre de 8 o más cubos (de 12)^{5 6} Agrupar objetos por una dimensión¹ Señala la torre de cubos más alta (6) y más baja (4)⁶ Coloca un cubo dentro y fuera de una taza⁶ Se sirve de un bastón para acercar un juguete (M)⁷ “ Stroop de forma” con ítems como una fruta dentro de línea de otro fruta y debe identificar las frutas adentro¹ ¿Suele prestar atención, por un tiempo largo, objetos que le interesan, como juguetes, libros con dibujos o personas que le son de su agrado?^{1 8} 	<ol style="list-style-type: none"> Hacer una torre de 4 o más cubos (con modelo) Señala la torre de cubos más alta y más baja (6 y 4 cubos). Agrupar objetos en una dimensión El concepto de “uno”¹⁰ Puede hacer el “ Stroop” de Forma (o Prueba de Día y Noche) Juega con una muñeca lavando su cara, limpiando su nariz, y dando de comer con cuchara¹⁰ Distingue entre un objeto pesado y uno que es más liviano¹⁰
36-48 meses	<ul style="list-style-type: none"> Agrupar por una dimensión y después cambia y agrupa de otro dimensión¹ Construye un puente con 3 cubos con modelos presente (6 cubos)^{5 6 7} (2-3ª) Coloca todas las piezas del tablero al rotarlo⁶ Dice que falta cuando se quita un objeto de un grupo de 3 (memoria)⁶ Arma rompecabezas de 6 piezas sin tanteos⁷ Rehace un rompecabezas simple⁷ 3-4- Repite tres dígitos⁴ Poner objetos en categorías y después hacerlo cambiando dimensiones (1) (también en cognitivo)¹ Sacrifica satisfacciones inmediatas ante la promesa de un beneficio mayor (M)⁷ Indica su edad mostrando 3 dedos⁷ 	<ol style="list-style-type: none"> Agrupar objetos en una dimensión y después otro dimensión Construir un puente con 3 cubos con modelo presente Repite 3 y 4 dígitos “ Cabeza y pie” prueba (número correcto) Tiene el concepto de “más”¹⁰ Rompecabezas de figura humana Indica su edad con sus dedos o palabras Cuenta hasta 10 Copiar patrones con colores con 6 objetos¹⁰

	<ul style="list-style-type: none"> • Usa adjetivos numerales ⁷ • Cuenta automáticamente hasta 10 ⁷ • ¿Suele prestar atención, por <i>varios minutos</i>, en objetos que le interesan, como juguetes, libros con dibujos o personas que le son de su agrado, logrando jugar con ellos un buen rato antes de cambiar de actividad?⁸ 	
48-59 meses	<ul style="list-style-type: none"> • Agrupa por color y forma (2 dimensiones)⁴ • Distingue el tamaño entre tres tiras de papel (larga y corta)⁶ • Entre 5 figuras geométricas, distingue cual es distinta⁷ • Cuenta 3 o más o números con 1-1 correspondencia¹ • Cuenta de dos y después 3 cubos (de 5 cubos)⁶ • Entrega 3, 4, o 5 objetos según se le pidan⁷ • Reconoce los números y pueden encontrar otro ejemplo de un número⁷ • Puede sumar hasta 5 (3+2, 4+1)⁷ • Poner en categorías basados en una tercer dimensión^{1 10} 	17. Agrupa objetos por 2 dimensiones al mismo tiempo 18. Agrupar objetos en 2 dimensiones que depende de una tercera dimensión 19. Entre 5 figuras geométricas, distingue el distinto (modelo Woodcock- Muñoz) 20. Cuenta con 1-1 correspondencia hasta 5 (incluyendo conceptos de cuál es el quinto; que numero va después) 21. Entrega 3 y después 2 objetos (suma a 5) 22. Entrega 2 y más 1 para hacer 3 23. “Cabeza y pie” prueba (número correcto es más grande)
• 2. Lenguaje Expresivo , receptivo y comunicación		
24-36 meses	<ul style="list-style-type: none"> • Nombra partes del cuerpo más fácil² • Señala 3 partes del cuerpo al nombrárselas^{6 7} • Señala partes del cuerpo poco visibles (ej espalda)⁶ • Nombra animales , nombra objetos⁵ • Nombra 2 o 3 figuras de una lamina (casa, perro, niño, carro)⁶ • Reconoce al menos dos figuras ² • Reconoce 2 o 3 colores¹ • Tiene noción del color ² • Distingue colores y nombra alguno⁷ 	Receptiva: <ol style="list-style-type: none"> 1. Señala 3 partes del cuerpo fáciles 2. Reconoce 3 colores 3. señalar objetos familiares (casa, perro, niño, carro) Expresiva: <ol style="list-style-type: none"> 4. Dice su nombre (apellido no es necesario) 5. Nombra 1 color 6. Dice más de 20 palabras¹¹ (M)

	<ul style="list-style-type: none"> • Asocia dos palabras con significado. (M)² (ej asocia dos palabras como "quiero leche", "quiero dulce", "mamá leche".) • Combina verbo o sustantivo con allá, aquí, allí, ahí⁶ • Usa plurales; usa artículos; usa pronombres⁶ • Emplea adecuadamente los adjetivos femeninos o masculinos según el general de las personal u objetos⁶ • Usa oraciones completas^{4 6} (edades 3-4) • Dice su nombre (algunos completos, otros solo primer nombre)^{4 5 6} • Diferencia alto-bajo, grande-pequeno⁴ • Reconoce grande y chico, más y menos, largo-corto^{5 7 6} (edades 3-4) • Cumple dos órdenes que no se relacionan (ej tráeme el libro y cerrá la caja)⁶ • Ejecuta ordenes sencillas (1 orden)⁷ • Imita los sonidos de la vaca o perro⁷ • Dice para que sirven algunos objetos⁷ • ¿Imita actividades de la vida real, como por ejemplo alimentar a una muñeca, jugar a la casa, o imaginarse que es un personaje de televisión o de películas con usted u otra persona?⁸ 	<p>7. Dice palabras de más de 3 categorías (MacArthur II (M))</p> <p>8. Dice más de 1 forma de verbo¹¹ (M)</p> <p>9. Dice frases de más complejidad por lo menos en 3 ítems¹¹ (M)</p>
36-48 meses	<ul style="list-style-type: none"> • Nombra o reconoce 10 objetos usuales de lamina⁷ • Nombra partes del cuerpo más difícil¹ • Señala 3 partes del cuerpo (lengua, cuello, brazos, rodilla, dedo gordo)⁶ • Reconoce 3 o 4 colores^{1 7} • Reconoce y nombra 3 acciones de una lamina⁶ • Nombra 5 colores^{4 7} • Nombra y/o señala 3 colores⁵ • Nombra figuras geométricas (circulo, cuadrado, triangulo)⁵ • Sabe el nombre de tres personas cercanas.^{2 6} • Dice si es niño o niña^{2 5 7} (edad 2-3) • Conoce el nombre de sus padres (mama y papa)⁵ • Dice para que sirve 5 objetos^{4 7} • Conoce la utilidad de objetos^{5 7} • Diferencia adelante-atrás, arriba-abajo⁴ • Discrimina suave y duro⁶ 	<p>Receptiva</p> <p>10. Reconoce 5 figuras geométricas (o objetos usuales de lamina)</p> <p>11. Señala 3 partes del cuerpo más difícil</p> <p>12. Entiende sobre y abajo; adentro y afuera</p> <p>Expresiva</p> <p>13. Nombra 4 colores</p> <p>14. Dice para que sirve 2 objetos</p> <p>15. Dice 2 o más verbos de verbos¹¹ (M)</p> <p>16. Dice más de 5 pronombres¹¹ (M)</p> <p>17. Dice 10 o más frases con más complejidad¹¹ (M)</p> <p>18. Diferencia adelante-atrás, arriba-abajo</p>

	<ul style="list-style-type: none"> • Ejecuta ordenes relativas (“trae el grande”)⁷ • ¿Hace frases de cuatro o más palabras para preguntar algo que [él/ella] quiere hacer? (ejemplo: vamos a los columpios? (M)⁸ • ¿Juega en la casa a representar diferentes personajes: mamá, papá, hermano, tía, abuela, etc (M).⁸ • ¿Obedece una instrucción que contiene más de dos acciones? (Por ejemplo, si usted le pide que saque leche del mueble y la deje en la mesa, lo hace)(M)⁸ 	
48-59 meses	<ul style="list-style-type: none"> • Reconoce absurdos⁵ • Usa plurales⁵ • Usa el pretérito y el futuro de los verbos⁷ • Reconoce antes y después⁵ • Nombra órganos sensorial y su capacidad (ej: ojo, ver; nariz, oler)¹⁷ • Reconoce 7 colores¹ • Defina 4 palabras⁵ • Nombra características de objetos⁵ • Realiza una serie de 3 ordenes^{6 7} • Dice su nombre completo con apellido⁶ • Discrimina pesado y liviano⁵ (edad3)^{6 7} • ¿Es capaz de explicar por qué quiere algo o quiere hacer algo? (Por ejemplo, frente a la pregunta “por qué quieres jugo?”, es capaz de responder por ejemplo, “¿porque tengo sed”)⁸ • ¿Hace un relato breve de algún evento significativo como por ejemplo: “fui con la mamá a la feria y....”; “Jugamos a la pelota con Juanito y...”⁸ • Cuenta una historia o cuento que haya escuchado varias veces (ej. Caperucita roja, blanca nieves, cenicienta, etc.) aunque no lo haga perfecto?⁸ 	<p>Receptiva</p> <p>19. Reconoce 7 colores</p> <p>20. Realiza una serie de 3 ordenes</p> <p>Expresiva</p> <p>21. Nombra 3 órganos sensorial y su capacidad (ej: para que sirve su ojo? Para ver)</p> <p>22. Dicen todas las formas de verbos¹¹ (M)</p> <p>23. Dicen 20 pronombres¹¹ (M)</p> <p>24. Dicen 30 o más frases con la forma más compleja¹¹ (M)</p> <p>25. Explica por qué quiere algo o quiere hacer algo. Por ejemplo, frente a la pregunta “por qué quieres jugo?”, es capaz de responder por ejemplo, “¿porque tengo sed”</p> <p>26. Cuenta una historia que haya escuchado aunque no lo haga perfecto.</p>
<ul style="list-style-type: none"> • 3. Desarrollo socio emocional y problemas conductuales 		

24-36 meses	<ul style="list-style-type: none"> • Expresa cariño, enojo, celos. (M)² • Juega cerca de otros niños y comparte juguetes. (M)^{2 6} • Demuestra preferencia o rechazo por prendas de vestir o alimentos (M)⁶ • Manifiesta ternura, enojo, celos hacia las personas de su medio (M)⁶ • Tiene preferencia por un juguete determinado (niño contesta)⁶ • Pelea con sus amiguitos por la posesión de un juguete (M)⁷ • Se obstina en hacer las cosas el solo(M)⁷ • Canta algunos canciones sencillas (M)⁷ • Le encanta producir ruidos y sonidos con cualquier cosa (M)⁷ • Habitualmente sonríe, hace sonidos, pide jugar o hacer algo juntos cuando ve a sus personas favoritas?⁸ • ¿Habla con uno o más niños?⁸ • Usa frases cortas (de tres o más palabras) para decir lo que [ella/el] quiere, como por ejemplo “yo quiero eso” o “no quiero eso”?⁸ • ¿Ayuda en la casa haciendo cosas simples, como ayudar a guardar sus juguetes o traer algo cuando se lo piden⁸ • ¿Habitualmente mira, sonríe, hace sonidos, mueve sus brazos y pide jugar o hacer algo juntos cuando ve a sus personas favoritas?(M)⁸ • ¿Es capaz de explicar por qué quiere algo o quiere hacer algo? (Por ejemplo, frente a la pregunta “por qué quieres jugo?”, es capaz de responder por ejemplo, “¿porque tengo sed”)(M)⁸ • Indica si durante los últimos ____ meses, cada frase nunca jamás ha sido verdad, a veces ha sido verdad, o ha sido verdad con frecuencia para [NOMBRE DEL NIÑO].”⁹ • “¿Ha sido poco colaborador?” • “¿Tiene problemas conciliar el sueño?” • “¿Tiene problemas del habla?” • “¿Tiene rabietas o mal humor?” • “¿Se ha puesto nervioso o tenso?” • “¿Ha sido infeliz, triste o deprimido?” 	<ol style="list-style-type: none"> 1. El niño está jugando a que es animal o otra persona u otro rol con sus padres y hermanos (juego simbólico; M) 2. El niño está dispuesto a ayudar con quehaceres simples cuando la mamá se lo pide (disposición a colaborar) (M) 3. El niño sigue indicaciones (colabora) cuando se le pide (M) 4. El niño usualmente está contento (M) 5. El niño llora frecuentemente (sin motivo- M) 6. Frente a extraños el niño reacciona inicialmente con miedo (M) 7. Se integra a un grupo con facilidad (M) 8. “¿Su niño ha sido poco colaborador? (M) 9. “¿Su niño tiene problemas conciliar el sueño? (M) 10. “¿Su niño tiene problemas del habla?” (M) 11. Su niño mira, sonríe, hace sonidos, mueve sus brazos y pide jugar o hacer algo juntos cuando ve a sus personas favoritas?(M)
36-48 meses	<ul style="list-style-type: none"> • Ayuda en pequeñas tareas. (M)² • Le gusta llamar la atención. (M)² • Comprende que debe esperar su turno (observar si puede)⁶ 	<ol style="list-style-type: none"> 12. “¿Su niño tiene rabietas o mal genio?” (M) 13. “¿Su niño se ha puesto nervioso, encuerdado o tenso?” (M)

	<ul style="list-style-type: none"> • Copera en juegos de construcción ⁷ • Canta y baila a escuchar la música ⁷ • A veces se muestra terco y obstinado (M) ⁷ • Muestra un fuerte deseo de agradar (M) ⁷ • Juega en grupos de dos o tres cambiando continuamente de actividad ⁷ • Indica si durante los últimos ____ meses, cada frase nunca jamás ha sido verdad, a veces ha sido verdad, o ha sido verdad con frecuencia para [NOMBRE DEL NIÑO].” ⁹ • “¿Ha sido poco colaborador?” • “¿Tiene problemas conciliar el sueño?” • “¿Tiene problemas del habla?” • “¿Tiene rabietas o mal humor?” • “¿Se ha puesto nervioso o tenso?” • “¿Ha sido infeliz, triste o deprimido?” 	<p>14. “¿Su niño ha sido infeliz, triste o deprimido?”</p> <p>15. El niño le gusta jugar con otros niños, tiene amistades (M)</p> <p>16. El niño ayuda en pequeñas tareas cuando y limpiando la casa (M)</p> <p>17. El niño demuestra simpatía si una persona está herida (M)</p> <p>18. Su niño de vez en cuando canta y baila a escuchar música? (M)</p> <p>19. Su niño conozca el significado de festividades históricos (M)</p>
48-59 meses	<ul style="list-style-type: none"> • Le gustan los juegos de competencia. (M) ² • Hace pequeños mandados en la casa. (M) ² • Hace mandados fuera de la casa y visita a los amiguitos y a los familiares cercanos (M) ⁶ • Puede vestirse y desvestirse solo, o con poca ayuda (M) ⁶ • Tema la oscuridad, puede sentir temor irracional (M) ⁷ • ¿Dice lo que siente para explicar por qué está haciendo algo o quiere algo? (Por ejemplo poder responder “porque estoy contento/ triste/ enojado”(M)) ⁸ • ¿Sostiene conversaciones con amigos/as y adultos, intercambiando en una variedad de temas? (Por ejemplo: comidas, amigos, colegio, horario de acostarse(M)) ⁸ • ¿Imita actividades de la vida real, como por ejemplo alimentar a una muñeca, jugar a la casa, o imaginarse que es un personaje de televisión o de películas con usted u otra persona con uno o más niños?(M) ⁸ • ¿Anticipa una acción frente a un peligro, por ejemplo frente a un brasero o estufa dice ¡no quema! o frente a un enchufe ¡no duele!(M) ⁸ • ¿Anticipa algunas acciones frente a una situación de su vida cotidiana por ejemplo Ud. dice vamos a comprar y el va a buscar su chaqueta y se 	<p>20. El niño puede vestirse y desvestirse solo o con poca ayuda (M)</p> <p>21. El niño hace mandados fuera de la casa (M)</p> <p>22. El niño puede seguir las reglas de un juego (M)</p> <p>23. El niño se lleva bien con otros niños? (M)</p> <p>24. El niño se puede concentrar o prestar atención por mucho tiempo? (M)</p> <p>25. El niño se siente despreciable o inferior?” (M)</p> <p>26. El niño se ha puesto nervioso, encuerado o tenso? (M)</p> <p>27. El niño se haga demasiado joven por su edad?(M)</p> <p>28. ¿Su niño ha sido infeliz, triste o deprimido? (M)</p>

	<p>la pasa?(M)⁸</p> <ul style="list-style-type: none"> Indica si durante los últimos ____ meses, cada frase nunca jamás ha sido verdad, a veces ha sido verdad, o ha sido verdad con frecuencia para [NOMBRE DEL NIÑO].”⁹ Ver preguntas el seccion anterior 	
○ 4. Las destrezas emergentes académicas		
24-36 meses	<ul style="list-style-type: none"> Copia una línea vertical^{2 6 7} Copia una línea recta⁵ Copia con dificultad una línea horizontal^{6 7} Copia una circulo⁶ Pasa las páginas de un libro o una revista de una en una⁶ Hace garabatos con el lápiz⁷ 	<ol style="list-style-type: none"> mira dibujos en un libro brevemente. mira dibujos con mucho interés por un tiempo más largo. pasa las páginas de un libro El niño le gusta ver dibujos o libro(M) Cuenta hasta 5 (puede ser menos, hay que determinar rango relevante en piloto) Copia línea vertical Puede agarrar el lápiz con el punto abajo
36-48 meses	<ul style="list-style-type: none"> Copia un círculo.^{2 6 7} Rasga papel^{2 6} Dibuja figura humana rudimentaria^{4 5 7} Copia una cruz^{5 6 7} Le gusta mirar laminas, fotos e imagenes⁷ 	<ol style="list-style-type: none"> mira dibujos brevemente mira dibujos con mucho interés pasa las páginas de un libro El niño le gusta ver dibujos o libro(M) Cuenta hasta 10 (ver arriba) Copia línea vertical Pasa las páginas de un libro o una revista Dibuja una figura humana rudimentaria Reconoce 3 letras de los primeros 10 Puede escribir 2 letras de su nombre (B)(Bayley)
48-59 meses	<ul style="list-style-type: none"> Cuenta hasta 20 o más numeros³ Cuentan dedos de sus manos^{5 7} Copia un cuadrado^{2 5 7} (edad 3-4 y 4-5) Copia triangulo^{2 7} Comienza a recortar con tijeras² Nombrar por lo menos 10 letras³ Escribir su nombre³ Leer o haga como si estuviera leyendo³ Dibuja figura humana nivel 2^{4 5 6} (4 partes)⁷ (3 partes) 	<ol style="list-style-type: none"> Leer o hacer como si estuviera leyendo Cuenta hasta 20 Nombra por lo menos 10 letras Copia un cuadrado Copia un triangulo Reconoce su nombre Puede escribir por lo menos 2 letras de su nombre Dice nombres de las días de la semana

	<ul style="list-style-type: none"> • Dibuja imitando escalera^{4b} • Dibuja una (la letra) H⁶ • Utiliza las tijeras y trata de cortar papel siguiendo un trazo recto⁷ 	
--	--	--

1) Zil y Zev (2007)

2) Escala de Desarrollo de Guatemala

3) National Household Education Surveys

4) Escala Abreviada Nelson Ortiz (1999)

5) TEPSI (1985)

6) EDIN (1987)

7) Escala de Observacion de Desarrollo, Paraguay

8) Inventory of Early Development

9) Achenback Child Behavior Checklist

10) Bayley Scale of Infant Development

11) MacArthur II

M) Informe de madre

Annex 3
Final Item Mapping
Associated Factors (Phase III)

Survey of the Household

Sección	Información	Número de pregunta	Instrumento de origen de la pregunta
I. Composición de personas en la vivienda	Número de personas que viven en la casa	1	PRIDI
	Nombre de las miembros de la familia nuclear del niño/a que viven en la casa	2	MICS4
	Relación de parentesco con el niño/a	2	MICS4
	Sexo	2	MICS4
	Edad (en años completos)	2	MICS4
	Lengua / dialectos que habla	2	PRIDI
	Máximo nivel educativo alcanzado	2	PRIDI, MICS4
	Ha concluido dicho nivel educativo	2	PRIDI
II. Características de la vivienda	Material que predomina en los pisos de la vivienda	3	MICS4
	Material que predomina en el techo de la vivienda	4	MICS4
	Material que predomina en las paredes exteriores de la vivienda	5	MICS4
	Cuartos de la vivienda que se usan para dormir	6	MICS4
	Artefactos o servicios en la casa	7	PRIDI, MICS4
	Fuente de luz en la casa	8	PRIDI
	Principal fuente de agua	9	MICS4
	Clase de instalación sanitaria	10	MICS4
III. Materiales de estimulación de los niños	Número de libros para niños menores de 5 años o libros con dibujos que hay en casa	11	MICS4, PIRLS
	Número de libros para personas adultas que hay en la casa	12	PRIDI, PIRLS
IV. Salud de la madre / cuidador principal	Escala de depresión	13	Head Start FACES
V. Participación en programas sociales	Participación de la familia en programas sociales		
	Programa de nutrición	14	PRIDI
	Programa de salud	14	PRIDI
	Programa de lucha contra la pobreza	14	PRIDI
	Programa de transferencias condicionadas	14	PRIDI

	Programa de apoyo a la actividad económica	14	PRIDI
	Participa o participó	14	PRIDI
	Nombre del programa en el que participa o participó	14	PRIDI
	Cantidad de años en los que ha participado	14	PRIDI

Survey of the Child

Sección	Información	Número de pregunta	Instrumento de origen de la pregunta
VI. Datos de la familia	Número de línea de la persona que responde el cuestionario	1	PRIDI
	La madre/cuidador principal trabaja fuera de la vivienda	2	PRIDI
	Nº de días de la última semana en que se dejó al niño/a solo o al cuidado de otro niño/a por más de una hora	3	MICS4
VII. Uso del tiempo	Cosas con las que el niño/a juega cuando está en casa	4	MICS4
	Frecuencia con que el niño/a juega con padre, madre, otros parientes, amigos, empleados de la vivienda	5	PRIDI
	Personas mayores de 15 años con las que el niño/a realizó actividades en los pasados 3 días	6	MICS4
VIII. Pautas de crianza	Reglas o rutinas de la casa sobre alimentación, hora de dormir, tareas en casa y momentos en que la familia come junta	7	Head Start FACES
	Formas de enseñar al niño/a sobre cómo comportarse	8	MICS4
	Percepción sobre el uso del castigo físico para la crianza o educación de un niño/a	9	MICS4
IX. Desarrollo y salud del niño/a	Desarrollo del niño/a	10	MICS4
	Calificación de la salud del niño/a	11	Head Start FACES
	Enfermedades o malestares crónicos	12	Niños del milenio, PRIDI
	Hábitos de higiene que practica el niño/a (lavado de dientes y manos)	13	Head Start FACES
X. Participación en programas de la primera infancia	Participación del niño/a en programas de la primera infancia		
	Programa educativo de 0 a 3 años	14	PRIDI
	Programa educativo formal de 3 a 5 años	14	PRIDI
	Programa educativo no formal de 3 a 5 años	14	PRIDI

	Programa de salud	14	PRIDI
	Programa de nutrición	14	PRIDI
	Participa o participó	14	PRIDI
	Nombre del programa	14	PRIDI
	Tipo de programa: público, privado o mixto	14	PRIDI
	Cantidad de años en los que ha participado	14	PRIDI
XI. Salud de la madre	Atención prenatal durante el embarazo con el niño/a	15	MICS4, Niños del Milenio
	Persona que dio la atención prenatal	16	MICS4
	Nº de veces en que asistió al médico o recibió atención prenatal	17	MICS4
	Nº de semanas de gestación de la madre	18	PRIDI
	Persona que atendió el nacimiento del niño/a	19	MICS4
	Lugar del nacimiento del niño/a	20	MICS4
	Peso y medida del niño/a al nacer (revisión de tarjeta de nacimiento)	21	MICS4, PRIDI
	Peso y medida del niño/a al nacer (reporte del entrevistado)	22	MICS4, PRIDI
	Alguna vez el niño/a ha sido amamantado	23	WHO y PRIDI
	Se sigue amamantando al niño/a	24	PRIDI
	Edad en la que se terminó de amamantar al niño/a	25	PRIDI
XII. Desarrollo socio-emocional	Desarrollo socio-emocional del niño/a	26	Varias fuentes y PRIDI

Annex 4
National Adaptations made in Paraguay, Costa Rica and Nicaragua (Peru = International Version)
Phase II²⁰

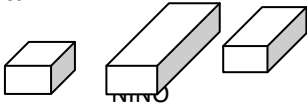
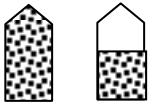
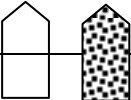
Form A					Form B				
	International/Peru	Costa Rica	Nicaragua	Paraguay		International/Peru	Costa Rica	Nicaragua	Paraguay
A1	pon			poné	B1	hacerle cariño			querés acariciarle
A2					B1	esta con hambre			tiene hambre
A3	sapo			vaca	B1	ponle			poné
A4	has			vos hacé	B1	de comer de nuevo		de comer otra vez	
A5	has hecho			hiciste	B2	pon			poné
A6	maderitas	tuquitos		maderas	B3	has hecho			hiciste
A6	ponlas			ponelas	B4				
A6	chiquita			chica	B5	cubos		cubos o bloques	
A7					B5	jueguitos			juegos
A8	dame			pasame	B6	pon			poné
A9	la pita	el cordón	el cordón	el cordón	B6	puedes			podés
A10	pon			poné	B6	agrupar		hacer grupos	
A10	chiquito			chico	B7	agrupar		hacer grupos	
A10	pavo real	gallina		pato	B8				
A10	puedes			podés	B9				


²⁰ FAN for Phase III incorporated changes made in Phase II; majority of national adaptations in Phase III were made to the indigenous language instruments.


A10	junto a		con	B10				
A11	un toque	un golpe	un golpe	B11	toque	Golpe	golpe	Golpe
A12				B11	tú			Vos
A13				B12				
A14				B13				
A15				B14				
A16	barriga		panza		escucha			
A17	sientate			B15	atentamente		oí con atención	
A18				B15	sientate			Sentate
	te haces una		te	B15	aplaude			Aplaudí
A19	herida		lastimás	B16	te haces una herida			te lastimás
A19	estas con sueño		tenés sueño	B16	estas con sueño			tenés sueño
A20	te voy		voy	B16	tienes hambre			tenés hambre
A20	dí		decí	B17	sembrando			Plantando
A21	mostrar		enseñar	B17	Columpiándose		Columpiándose-Meciéndose	Hamacándose
A21	cucharita		cuchara	B18	Columpiándose		Columpiándose-Meciéndose	Hamacándose
A22	sapo		vaca	B19	Columpiándose		Columpiándose-Meciéndose	Hamacándose
A23				B20				
A24				B21	pita	cordón	cuerda	Cordón
A25	pelota	Bola		B22	chocarlo		pegarle	
	arco	marco		B22	Ponte aquí			Ponete acá



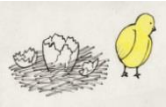


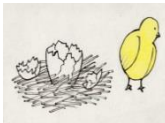
A26	al costado	a la par de	al lado		B23	pelota	Bola		
A26	cabecita			cabeza	B23	puedes cogerla		podes	
A27	pita	cordón		línea	B24	palita	Paleta	agarrarla	
A28	pita	cordón		línea	B24	tu continúa		vos seguís	
A28	pelotita			pelota	B25	pones		Poné	
A28	tirartela			pasartela	B26				
A28	cogerla			agarrarla	B27	vela	Candela		
A29	pita	cordón		línea	B27	media		calcetín	
A30	juntas	armas			B28	vela	Candela		
A31	juntas	armas			B29	al costado	a la par de	al lado	
A32	coge			agarra	B29	cabecita		Cabeza	
A32	haz			hacé	B30	junta a	a la par de		
					B30	jarra		pichel	
					B30	niña		Nena	
					B31				
					B32	barriga		barriga - panza	Panza
					B33				
					B34	mostrar		enseñar	Enseñar
					B34	dices			Decís
					B35	juntas	Armas		
					B37	haz			Hacé
					B38	enséñame			Mostrame

Annex 5
Eliminated Items in Forms A and B, Phase II to Phase III
Form A

	TITULO	CONSIGNA
A-3	DISCRIMINA EL OBJETO DIFERENTE <div> Materiales: <ul style="list-style-type: none"> - 4 sapos chicos - 1 sapo grande </div>	1. “¿Qué sapo es diferente?”
A-4	IMITA CONSTRUCCIÓN DE TORRE <div> Materiales: <ul style="list-style-type: none"> - 7 cubos de 2.5 x 2.5 cms </div>	1. “Mira lo que hago” 2. “Ahora has una torre como esta, lo más alta que puedas y me avisas cuando termines” Si FALLA se realiza un segundo intento diciendo: “Voy a desarmar la torre, para que vuelvas a hacer otra lo más alta que puedas”.
A-6	ORDENA DE MÁS GRANDE A MÁS CHICO <div> Materiales: <ul style="list-style-type: none"> - 3 bloques de madera: grande, mediano y chico </div>	1. “Acá tengo unas maderitas, ponlas en orden de la más grande a la más chiquita “ 
A-7	DIFERENCIA LA BOTELLA QUE PESA MENOS <div> Materiales: <ul style="list-style-type: none"> - Botella pesada (totalmente llena de arena) - Botella no tan pesada o de peso moderado (la mitad llena de arena) </div>	1. “Dame la botella que pesa MENOS”.  NIÑO
A-8	<div> Materiales: <ul style="list-style-type: none"> - Botella pesada (totalmente llena de arena) - Botella no tan pesada o de peso moderado (la mitad llena de arena) </div>	2. “Dame la botella que pesa MAS”. 

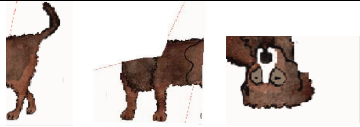
		<div></div> <div>NIÑO</div>														
A-9	<div>ENSARTA CUENTAS</div> <div><div>Materiales:</div><div><div>-</div><div>Cordón delgado de 28cm.</div></div><div><div>-</div><div>5 cuentas de aproximadamente 2cm de largo, con un agujero de 5mm en el centro</div></div></div>	<div>1. “Mira como paso la pita” mientras coge una cuenta y la ensarta en el cordón.</div> <div>2. “Ahora tú continúa metiendo todas”</div>														
A-11	<div>JUEGO DE TOQUES</div> <div><div>Materiales:</div><div><div>-</div><div>2 lápices</div></div><div><div>-</div><div>1 Plato</div></div></div>	<div>1. EJEMPLO 1: Vamos a jugar un juego: cuando yo dé un toque (el evaluador golpea una vez) yo quiero que tu des dos toques, así (si el niño no lo hace, el evaluador lo ayuda). Vamos a intentar hacerlo. Cuando yo de un toque (el evaluador da un golpe), tu das... (el niño, sin o con la ayuda del evaluador debe dar dos toques sobre el plato) Si el niño no lo hace o se equivoca, repetir el primer ejemplo solo una vez más. EJEMPLO 2: “Ahora, si yo doy dos toques (el evaluador da dos toques) yo quiero que tu des un toque, así (el evaluador da un toque). Vamos a intentar hacerlo. Cuando yo de dos toques (el evaluador da dos toques), tu das... (si el niño no da un toque, el evaluador lo ayuda)” Si el niño no lo hace o se equivoca, repetir el ejemplo 2 solo una vez más.</div> <div>2. “Ahora vamos a jugar el juego”.<table><tr><th>Secuencia de Toques del evaluador</th><th>Respuesta del niño</th></tr><tr><td>1</td><td></td></tr><tr><td>2</td><td></td></tr><tr><td>2</td><td></td></tr><tr><td>1</td><td></td></tr><tr><td>1</td><td></td></tr><tr><td>2</td><td></td></tr></table></div>	Secuencia de Toques del evaluador	Respuesta del niño	1		2		2		1		1		2	
Secuencia de Toques del evaluador	Respuesta del niño															
1																
2																
2																
1																
1																
2																
A-17	SEÑALA PARTES DEL CUERPO	<div>1. “Ahora quiero que me enseñes cuál es tu dedo...”</div> <div>2. “¿Dónde están...?”<div><div>-</div><div>tus cejas</div></div><div><div>-</div><div>tu rodilla</div></div><div><div>-</div><div>tu barriga</div></div><div><div>-</div><div>tu codo</div></div></div>														

A-18	SIGUE TRES INSTRUCCIONES Materiales: - Títere de perro	“Escucha atentamente: por favor tráeme el perro (señalándolo), siéntate (indicando la silla o el piso) y aplaude dos veces”							
A-21	NOMBRA OBJETOS COTIDIANOS Materiales: - Lápiz - Cucharita - Peine - Media - Gallina - Sapo	“Ahora yo te voy a mostrar algo y tú me vas a decir qué cosa es” <table><tr><td>Lista de palabras</td></tr><tr><td>Lápiz</td></tr><tr><td>Cucharita</td></tr><tr><td>Peine</td></tr><tr><td>Media</td></tr><tr><td>Gallina</td></tr><tr><td>Sapo</td></tr></table>	Lista de palabras	Lápiz	Cucharita	Peine	Media	Gallina	Sapo
Lista de palabras									
Lápiz									
Cucharita									
Peine									
Media									
Gallina									
Sapo									
A-22	SEÑALA OBJETOS COTIDIANOS Materiales: - Lápiz - Cucharita - Peine - Media - Gallina - Sapo	“Te voy a preguntar dónde está cada una de estas cosas y tú me la señalas, ¿listo? ¿Donde esta (nombre del objeto)?” <div></div> <ul style="list-style-type: none">- El sapo- El peine- La cucharita- El lápiz- La media- La gallina							
A-24	SEÑALA ACCIONES Materiales: - Lamina con 6 imágenes de personas realizando acciones.	<div>NOTA: En este ítem solo se pregunta por las imágenes que no nombró correctamente en el anterior. Las imágenes que si nombro correctamente en el ítem anterior se califican como correctas en este ítem.</div> Ahora quiero que señales la persona que está... <ul style="list-style-type: none">...Peinando...Lavando...Acariciando...Construyendo...Barriendo....Ordeñando							
A-26	ORDENA SECUENCIA TEMPORAL: 3 LAMINAS	<u>NACIMIENTO DEL POLLITO:</u> El evaluador pregunta: ¿Qué está pasando acá?							

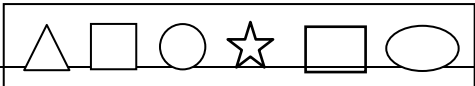
	<p>Materiales:</p> <ul style="list-style-type: none"> - Secuencia de láminas de NACIMIENTO DEL POLLITO 	 <p>Si el niño dice una respuesta correcta: “La cabecita de un pollito saliendo del huevo”. El evaluador refuerza su respuesta felicitándolo.</p> <p>Si el niño da una respuesta diferente o se queda callado, el examinador dice la respuesta correcta.</p> <p>El evaluador pregunta: ¿Qué está pasando acá?</p>  <p>Si el niño dice una respuesta correcta: “Hay un huevo”. El evaluador refuerza su respuesta felicitándolo.</p> <p>Si el niño da una respuesta diferente o se queda callado, el examinador dice la respuesta correcta.</p> <p>El evaluador pregunta: ¿Qué está pasando acá?</p>  <p>Si el niño dice una respuesta correcta: “Un pollito al costado del cascarón”. El evaluador refuerza su respuesta felicitándolo.</p> <p>Si el niño da una respuesta diferente o se queda callado, el examinador dice la respuesta correcta.</p> <p>A continuación colocan las 3 láminas en línea de la siguiente manera y pregunta:</p> <div style="display: flex; justify-content: center; align-items: center;">    </div> <p style="text-align: center;">NIÑO</p> <p>¿Qué pasa primero? _____</p> <p>¿Qué pasa después? _____</p> <p>¿Y al final? _____</p>
A-31	<p>COMPLETA ROMPECABEZAS: 3 PIEZAS</p>	<p><u>Perro</u></p> <p>Colocar las piezas en el siguiente orden:</p>

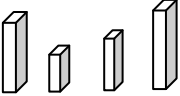
Materiales:

- Rompecabezas de 3 piezas de un perro


		 <p>NIÑO</p> <ol style="list-style-type: none"> 1. “Si juntas estas 3 piezas se puede ver un perro, mira como lo hago”. (Luego de hacer el ejemplo, volver a colocar las piezas en la posición inicial) “Ahora tu júntalas para ver el perro”
--	--	--


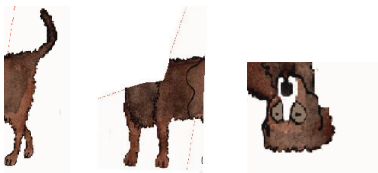

Form B

	TITULO	CONSIGNA
B-3	IMITA CONSTRUCCIÓN DE UN PUENTE: 3 BLOQUES <p>Materiales:</p> <ul style="list-style-type: none"> - 3 bloques rectangulares del mismo tamaño peso y color: 8.5cmsx4cmsx2cms 	<ol style="list-style-type: none"> 1. (Muestra del evaluador según esquema). “Mira lo que hago” 2. “Ahora tú intenta hacer un puente igual al mío y me avisas cuando termines” <p>Si FALLA realizar un segundo intento diciendo: “Veo que tú has hecho algo diferente a lo que yo hice. Ahora haz uno igualito al mío”.</p> <p>Si FALLA el segundo intento, pasar al ítem B-5</p>
B-7	CLASIFICACIÓN: POR 3 DIMENSIONES <p>Materiales:</p> <ul style="list-style-type: none"> - 1 cubo grande blanco, - 1 cubo mediano rojo, - 1 cubo chiquito negro; - 1 cilindro grande rojo, - 1 cilindro mediano negro - 1 cilindro chiquito blanco; - 1 bola grande negra - 1 bola mediana blanca - 1 bola chiquita roja 	<p>ESTE ITEM SOLO SE APLICA SI EN EL ITEM B-6 EL NIÑO OBTIENE UN PUNTAJE DE 2 PUNTOS</p> <ol style="list-style-type: none"> 1. “Acá hay más cosas, pon todos los que van juntos para hacer 3 grupos. Puedes usar cada uno de estos papeles para cada grupo”. (Señalar las tres hojas). 2. “Estos mismos juguetes se pueden agrupar de diferente manera, trata de encontrar otra manera de agruparlos para poner algunos en esta hoja, otros en esta y otros en la otra”. 3. ¿Ahora puedes intentar agruparlos de otra manera diferente?”
B-9	SEÑALA FIGURAS GEOMÉTRICAS <p>Materiales: Figuras grandes azules en forma de:</p> <ul style="list-style-type: none"> - cuadrado - círculo - triángulo - estrella - rectángulo 	<p>“Señala con tu dedo.....”</p> 

		<p style="text-align: center;">NIÑO</p> <p>.....el cuadrado el triangulo la estrella el ovalo el rectángulo el circulo</p>																					
B-10	<p>DIFERENCIA 4 OBJETOS POR ALTURA</p> <p>Materiales:</p> <ul style="list-style-type: none"> - 4 bloques de igual forma y color pero de diferente altura 	 <p style="text-align: center;">NIÑO</p> <ol style="list-style-type: none"> 1. <i>“Señala cuál es el más alto.”</i> 2. <i>“Señala cuál es el más bajo.”</i> 																					
B-11	<p>JUEGO DE TOQUES</p> <p>Materiales:</p> <ul style="list-style-type: none"> - 2 lápices - Plato 	<p style="text-align: center;">EJEMPLOS DE PRÁCTICA</p> <p><u>EJEMPLO 1:</u> Vamos a jugar un juego: cuando yo dé un toque (el evaluador golpea una vez) yo quiero que tu des dos toques, así (si el niño no lo hace, el evaluador lo ayuda). Vamos a intentar hacerlo. Cuando yo de un toque (el evaluador da un golpe), tu das... (el niño, sin o con la ayuda del evaluador debe dar dos toques sobre el plato) Si el niño no lo hace o se equivoca, repetir el primer ejemplo solo una vez más.</p> <p><u>EJEMPLO 2:</u> “Ahora, si yo doy dos toques (el evaluador da dos toques) yo quiero que tu des un toque, así (el evaluador da un toque). Vamos a intentar hacerlo. Cuando yo de dos toques (el evaluador da dos toques), tu das... (si el niño no da un toque, el evaluador lo ayuda)” Si el niño no lo hace o se equivoca, repetir el ejemplo 2 solo una vez más.</p> <p style="text-align: center;">EVALUACIÓN</p> <p>3. “Ahora vamos a jugar el juego”.</p> <table border="1"> <thead> <tr> <th>No de intento</th><th>Toques del evaluador</th><th>Respuesta</th></tr> </thead> <tbody> <tr><td>1</td><td>1</td><td></td></tr> <tr><td>2</td><td>2</td><td></td></tr> <tr><td>3</td><td>2</td><td></td></tr> <tr><td>4</td><td>1</td><td></td></tr> <tr><td>5</td><td>1</td><td></td></tr> <tr><td>6</td><td>2</td><td></td></tr> </tbody> </table>	No de intento	Toques del evaluador	Respuesta	1	1		2	2		3	2		4	1		5	1		6	2	
No de intento	Toques del evaluador	Respuesta																					
1	1																						
2	2																						
3	2																						
4	1																						
5	1																						
6	2																						

B-12	DIFERENCIA CONCEPTOS ENCIMA/DEBAJO Materiales: <ul style="list-style-type: none"> - Lápiz - Plato - 2 bloques de madera iguales 	1. “Coloca este lápiz encima del plato” (esperar la respuesta del niño). 2. “Ahora coloca este lápiz debajo del plato.”
B-15	SIGUE TRES INSTRUCCIONES Materiales: <ul style="list-style-type: none"> - Títere de perro 	1. “Escucha atentamente y haz lo que te digo”: “Por favor tráeme el perro , siéntate (indicando la silla o el piso) y aplaude dos veces”
B-17	UTILIZA PLURAL Materiales: <ul style="list-style-type: none"> - 2 lápices 	Colocar 1 lápiz: ¿Que es esto? _____ Colocar 2 lapices ¿Qué son estos? _____
B-19	SEÑALA ACCIONES Materiales: <ul style="list-style-type: none"> - Lamina con 8 imágenes de personas realizando acciones 	NOTA Si en el ítem anterior no nombro alguna imagen, preguntar por estas en el orden indicado más abajo. Calificar como correctas las imágenes que nombró adecuadamente en el ítem anterior “Ahora quiero que señales la figura que está...” ...Sembrando ...Ordeñando ...AcariciandoConstruyendo ...Lavando ...Tejiendo ...Columpiándose ...Haciendo trenzas

B-22	<div>TIRA LA PELOTA HACIA UN OBJETO</div> <div>Materiales:<ul style="list-style-type: none">- Una pita de 2 metros- Pelota de trapo- Plato de metal</div>	<div>“Ponte aquí en el extremo de la línea y tira la pelota hacia el plato y trata de chocarlo”</div>																		
B-26	<div>HOJEA LIBRO</div> <div>Materiales:<ul style="list-style-type: none">- Libro con 10 paginas pequeño y con colores</div>	<div>“Quiero ver como lees este libro”</div> <div>Si el niño no muestra interés por el libro se le puede decir una vez: “Quiero ver como lees este libro”.</div>																		
B-27	<div>NOMBRA OBJETOS COTIDIANOS</div> <div>Materiales:<ul style="list-style-type: none">- Pastilla- cucharita- Peine- Vela- Fósforos- Anteojos- Gallina- Media</div>	<div>“Ahora vamos a jugar un juego. Yo te voy a mostrar algo y tú me vas a decir qué cosa es, ¿listo?”</div> <table><tr><th>Lista de palabras</th><th>Respuestas</th></tr><tr><td>Pastilla</td><td></td></tr><tr><td>cucharita</td><td></td></tr><tr><td>Peine</td><td></td></tr><tr><td>Vela</td><td></td></tr><tr><td>Fósforos</td><td></td></tr><tr><td>Anteojos</td><td></td></tr><tr><td>Gallina</td><td></td></tr><tr><td>Media</td><td></td></tr></table>	Lista de palabras	Respuestas	Pastilla		cucharita		Peine		Vela		Fósforos		Anteojos		Gallina		Media	
Lista de palabras	Respuestas																			
Pastilla																				
cucharita																				
Peine																				
Vela																				
Fósforos																				
Anteojos																				
Gallina																				
Media																				
B-28	<div>SEÑALA OBJETOS COTIDIANOS</div> <div>Materiales:<ul style="list-style-type: none">- Pastilla- Cucharita</div>	<div>Voy a acomodar las cosas, espera un ratito sin tocarlas</div> <div></div> <div>Te voy a preguntar dónde está cada una de estas cosas y tú me la señalas, ¿listo?</div>																		

B-30	<p>ORDENA SECUENCIA TEMPORAL: 3 LÁMINAS DIFÍCILES</p> <p>Materiales:</p> <ul style="list-style-type: none"> - Secuencia de cartas de historias: “niña tomando agua” 	<p><u>Niña tomando agua</u></p> <p>Lámina 1: ¿Qué está pasando acá?</p>  <p>Respuesta válida: niña sirviéndose agua</p>
B-31	NOMBRA PARTES DEL CUERPO	<p>1. “Ahora vamos a jugar un juego. ¿Cómo se llama esto? (señalando su DEDO INDICE)</p> <p>2. “Ahora enseñame cómo se llama esto... (el examinador señala en su propia cara)</p> <ul style="list-style-type: none"> - tus cejas - tu rodilla - tu barriga - tu codo
B-32	SEÑALA PARTES DEL CUERPO	<p>3. “Ahora quiero que me enseñes cuál es tu dedo índice...”</p> <ul style="list-style-type: none"> - tus cejas - tu rodilla - tu barriga - tu codo
B-35	<p>COMPLETA ROMPECABEZAS: 3 PIEZAS</p> <p>Materiales:</p> <ul style="list-style-type: none"> - Rompecabezas de 3 piezas de un perro 	<p><u>Perro</u></p> <p>Colocar las piezas en el orden indicado:</p>  <p>NIÑO</p>
B-36	<p>COMPLETA ROMPECABEZAS: 4 PIEZAS</p> <p>Materiales:</p> <ul style="list-style-type: none"> - Rompecabezas de 4 piezas de un niño 	<p><u>Niño</u></p> <p>Colocar las piezas en el orden indicado:</p>  <p>NIÑO</p> <p>1. “Si juntas estas 4 piezas se puede ver un niño, junta estas cuatro</p>

		<i>piezas para ver el niño"</i>
B-38	<p>RECONOCE FIGURAS, LETRAS Y NUMEROS</p> <p>Materiales:</p> <p>- Cartilla con número, letras y figuras</p>	<ol style="list-style-type: none"> 1. <i>"Mira esto, enséñame donde hay una figura"</i> 2. <i>"Enséñame dónde hay un número"</i> 3. <i>"Enséñame dónde hay una letras"</i>

Annex 6
Changes in the Surveys, Manuals and Forms
Phase II to Phase III

1. Manual for Applying Surveys and Processes

Section	Change
1.2. Consideraciones éticas – Casos éticos	Se incluyó un párrafo para dar instrucciones sobre cómo proceder en casos en los que el niño/a evaluado presente una discapacidad no identificadas por sus padres o cuidadores.
2. Descripción de instrumentos	Se indicó qué instrumentos se encuentran unidos en un solo documento.
3.4. Uso de códigos	Se modificó el cuadro de códigos a partir de la nueva numeración de códigos: ahora NS=7 (antes 6), NC=8 (antes 7), NA=9 (antes 8). Los códigos para No administrado y el Error en la administración se han unido.
3.6. Cuaderno anecdótico	Se creó esta sección para indicar que los examinadores deben hacer un registro de sus observaciones sobre las aplicaciones.
4.1. Entrevista preliminar	Se modificó la explicación de selección del niño/a a evaluar, ahora por orden alfabético.
4.3. Registro de Visitas a Viviendas	Se especificó el uso de los códigos de las UPM.
4.5. TVIP	Se modificó la información sobre los códigos a usarse en los casos en los que no se obtiene respuesta del niño/a y en los que haya error de aplicación.
	Se añadió que indica que la prueba puede detenerse y retomarse en otro momento según el estado del niño/a.
4.6. Antropometría	Se detalló mayor información sobre el uso del tallímetro.
5. Pautas para los supervisores durante el trabajo de campo	Se creó esta sección para dar indicaciones sobre el rol de los supervisores en el trabajo de campo.

2. Training Manual

Section	Change
1. Perfil de examinadores	Se modificó el perfil para que incluya a personas de ambos sexos y se hizo énfasis en la necesidad de tener experiencia en trabajos de campos en zonas rurales.
2.2. Evaluación de participantes	Se añadió criterios de evaluación para evaluar el desempeño de los participantes durante los juegos de roles.
2.4. Materiales	Se actualizó la lista de materiales.
3. Horario de capacitación	Se añadió información sobre el cronograma para los días 4, 5, 6, 7, 8, 9 y 10 (se juntaron las actividades del día 10 y 11, por lo que ahora la capacitación tiene duración de 10 días.)
4.1. Día 1	Se modificaron los ejercicios de práctica del llenado de la Entrevista Preliminar, a partir de los cambios en el método de selección del niño/a.
Días 4, 5, 6, 7, 8, 9, y 10	Se incorporó información sobre las actividades a realizar en dichos días.
Anexos	Se modificó la tabla de práctica del llenado de la Entrevista Preliminar a partir

	de los cambios en dicho instrumento y se hizo algunas modificaciones en la Prueba escrita según el nuevo método de selección de niños/as.
--	---

3. Preliminary Interview

Question	Change
	Se colocó un nuevo encabezado, en el que se añadió el código de la UPM y el código del examinador.
5	5.A. se añadió espacios para colocar el código de la UPM.
	5.I. Se modificó el texto para indicar el Nuevo método de selección del niño/a (por orden alfabético).
	Se eliminó la columna en la que se colocaba el orden de los niños/as a partir de la modificación en la forma de selección del niño/a.
	Se eliminó la tabla de selección de niños/as

4. Household Registry

Question	Change
	Se colocó un nuevo encabezado, en el que se añadió el código de la UPM, nombre de la comunidad, número total de viviendas en la UPM y número de viviendas seleccionadas en la UPM.
	Se redujo el número de códigos de participación a 6.

5. Informed Consent Form

Question	Change
	Se colocó un nuevo encabezado con el código de la UPM

6. Family Survey

Question	Change
	Se colocó un nuevo encabezado con el código de la UPM.
2	Se añadió columna G: ¿Asiste actualmente a dicho nivel educativo?
	Columna F: se unió las opciones de maestría y doctorado, se hizo una categoría para postgrado incompleto y postgrado completo, se eliminó la palabra bachillerato para hacer alusión a la educación universitaria y se indicó que el llenado de años completados también debe hacerse para código 01 (educación inicial).
9	Se eliminó “agua embotellada/envasada” de la opción “Otros”.

11	Se cambió la opción “10 o más libros” a “11 o más libros” Decía 10 ó más, y se cruzaba con la opción anterior
12	Se cambió la opción “10 o más libros” a “11 o más libros”
14	Se añadió la indicación de que en caso la familia participe o haya participado en más de un tipo de programa, se anote aquel en el que participó por más tiempo.

7. Child Survey

Question	Change
2	Se modificó el formato para añadir la pregunta “¿Cuántas horas semanales trabaja?”
7	Se eliminó la opción “No tiene tareas para la casa”.
8	Opción h: se modificó el orden de la opción.
	Opción f: se eliminó “le dio una zurra”.
13	Se modificó el fraseo de oraciones y el formato de pregunta para que las respuestas se registren con frecuencias.
14	Se añadió la indicación de que en caso el niño/a participe o haya participado en más de un tipo de programa, se anote aquel en el que participó por más tiempo y se especificó que solo se registren los programas en los que únicamente participó el niño/a.
24	Se especificó que la pregunta hace se refiere a la ingesta de leche materna.
26	Opción n: se añadió “adulto conocido”.
	Tarjetas: se añadió tarjetas de frecuencia para la pregunta 13.

Annex 7
UPMs By Country

Costa Rica	# UPMs			Nicaragua	# UPMs			Paraguay, Sample 1	# UPMs			Peru*	# UPMs
Departament	Total	Urban	Rural	Department	Total	Urban	Rural	Department	Total	Urban	Rural	Region	Total
SAN JOSE	23	23	0	Boaco	2	1	1	Ascunción	18	18	0	Arequipa	20
ESCAZU	1	1	0	Carazo	3	1	2	Concepción	7	3	4	Ayacucho	7
DESAMPARADOS	5	3	2	Chinandega	4	2	2	San Pedro	12	3	9	Ancash	9
PURISCAL	3		3	Chontales	3	1	2	Cordillera	10	4	6	Cajamarca	19
TARRAZU	3	2	1	Estelí	3	1	2	Guaira	7	3	4	Callao	12
ASERRI	4	1	3	Granada	2	1	1	Caaguazu	15	6	9	Junin	18
MORA	2	0	2	Jinotega	3	1	2	Caazapa	6	2	4	Lambayeque	12
GOICOECHEA	3	3	0	León	5	2	3	Itapua	18	8	10	La Libertad	20
ALAJUELITA	4	4	0	Madriz	3	1	2	Misiones	4	2	2	Lima	118
VAZQUEZ DE CORONADO	2	2	0	Managua	6	4	2	Paraguari	9	3	6	Ica	10
ACOSTA	1	0	1	Masaya	4	2	2	Alto Parana	23	17	6	Iquitos	10
TIBAS	2	2	0	Matagalpa	4	1	3	Central	54	46	8	Huanuco	8
TURRUBARES	1	0	1	Nueva Segovia	3	1	2	Deembucu	3	2	1	Pasco	4
CURRIDABAT	1	1	0	Miskitu	6	2	4	Amambay	6	4	2	Piura	22
PEREZ ZELEDON	3	0	3	Rio San Juan	3	1	2	Canindeyu	5	2	3	Puno	5
LEON CORTES	1	0	1	Rivas	3	1	2	Presidente Hayes	3	1	2	San Martin	12
ALAJUELA	13	11	2									Ucayali	6
SAN RAMON	2	0	2	Total	57	23	34	Total	200	124	76		
NARANJO	3	1	2									Total	312
OROTINA	2	0	2					Paraguay, Sample 2	# Districts/UPMs			*No disaggregation of urban/rural available	
SAN CARLOS	3	0	3					Department	Total	Urban	Rural		
ALFARO RUIZ	1	0	1					Ascunción	8	8	0		
UPALA	3	1	2					Concepción	2	0	2		
CARTAGO	4	2	2					San Pedro	5	2	3		

[illegible]

Bibliography

- Achenbach, T. M. (1991). Manual for the Child Behavior Checklist/4-18 and 1991 Profile. Burlington, VT: University of Vermont, Department of Psychiatry.
- Almond, D. y Currie, J. (2010). Human capital development before age five. National Bureau of Economic Research Working Paper 15827. Cambridge: NBER.
- American Educational Research Association, American Psychological Association y National Council on Measurement in Education, AERA (1999). Standards for Education and Psychological Testing, Washington D.C. AERA, APA y NCME.
- Ayoub, C., O'Connor, E., Rappolt-Schlichtmann, G., Vallotton, C., Raikes, H. y Chazan-Cohen, F. (2009). Cognitive skill performance among young children living in poverty: Risk, change, and the promotive effects of Early Head Start. *Early Childhood Research Quarterly*. 24:3, pp. 289-305.
- Bakermans-Kranenburg, M. J., IJzendoorn, M. H. V., and Kroonenberg, P. M. (2004). Differences in attachment security between African-American and white children: ethnicity or socio-economic status? *Infant Behavior and Development*, 27(3), 417-433.
- Backhoff, E., Andrade, E., Sánchez, E. and Peon, Z. (2008). *El Aprendizaje en Tercero de Preescolar en México. Lenguaje y Comunicación. Pensamiento Matemático*. México: Instituto Nacional de Evaluación Educativa (INEE).
- Banco Interamericano de Desarrollo (2009). Indicadores y datos regionales sobre el desarrollo infantil temprano y el aprestamiento. Plan de operaciones.
- Bates, E., Bretherton, I., y Snyder, L. (1988). From first words to grammar: Individual differences and dissociable mechanisms. New York: Cambridge University Press.
- Behrman, R. J., Glewwe, P. y Miguel, E. (2007). *Methodologies to Evaluate Early Childhood Development Programs*. Washington D.C.: The World Bank.
- Bernier, A., Carlson, S. M., y Whipple, N. (2010). From External Regulation to Self-Regulation: Early Parenting Precursors of Young Children's Executive Functioning. *Child Development*. 81:1, pp. 326-339.
- Berk, Laura (1999) Desarrollo del niño y del adolescente. Prentice Hall, Madrid.
- Bornstein, M. H. y Haynes, O. M. (1998). Vocabulary Competence in Early Childhood: Measurement, Latent Construct, and Predictive Validity. *Child Development*, 69:3, pp. 654-671.
- Cabell, S. Q., Justice, L. M., Zucker, T. A., y Kilday, C. R. (2009). Validity of Teacher Report for Assessing

the Emergent Literacy Skills of At-Risk Preschoolers. *Language, Speech, and Hearing Services in Schools*, 40, pp. 109-115.

Cameron, C. E., Brock, L. L., Murrah, W. M., Bell, L. H., Worzalla, S. L., Grissmer, D., y Morrison F. (2012). Fine Motor Skills and Executive Function Both Contribute to Kindergarten Achievement. *Child Development*. 83:4, pp. 1229-1244.

Carlson, S. M., Moses, L. J. y Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Child Psychology*. 87, pp. 299-319.

Carter, A. S., Briggs-Gowan, M. J., Jones, S. M., y Little, T. D. (2003). The Infant–Toddler Social and Emotional Assessment (ITSEA): Factor Structure, Reliability, and Validity. *Journal of Abnormal Child Psychology*. 13:4, pp. 495-514.

Castro, V., Laguna, J. R., and Vijil, J. (2010). Informe de Resultados: ELI 2009. Caribe. Manuaga: Centro de Investigación y Acción Educativa Social.

Center on the Developing Child Harvard University, National Forum on Early Childhood Program Evaluation, y National Scientific Council on the Developing Child (2007). A Science Based Framework for Early Childhood Policy Using Evidence to Improve Outcomes in Learning, Behavior and Health for Vulnerable Children. On line: <http://www.ncdhhs.gov/dma/casemgmt/CMEBDPolicy.pdf>.

Chaddock, Laura, Charles H. Hillman, Sarah M. Buck, and Neal J. Cohen (2011a). “Aerobic Fitness and Executive Control of Relational Memory in Preadolescent Children.” On the web at: <http://kch.illinois.edu/Research/Labs/neurocognitive-kinesiology/default.htm> .

Chaddock, Laura, Matthew B. Pontifex, Charles H. Hillman, and Arthur F. Kramer (2011b). “A Review of the Relation of Aerobic Fitness and Physical Activity to Brain Structure and Function in Children.” *Journal of International Neuropsychological Society* 16, 1-11.

Cooper-Kahn, J. y Dietzel, L. (2008a). *Late, Lost and Unprepared*. Bethesda, MD: Woodbine House.

Cooper-Kahn, J. y Dietzel, L. (2008b). What is Executive Functioning? En línea: <http://www.idonline.org/article/29122/?theme=print>.

Denham, S. y R. Weissberg. (2004). Social-Emotional Learning in Early Childhood: What We Know and Where to Go From Here.” en Chesebrough, et al., (eds.). *A Blueprint for the Promotion of Prosocial Behavior in Early Childhood* (pp. 13-50. New York: Kluwer Academic/Plenum Publishers.

Diamond, Adele and Kathleen Lee (2011). “Interventions Shown to Aid Executive Function Development in Children 4 to 12 Years Old.” *Science* 33. 19 August 2011. www.sciencemag.org.

Du Toit, M. (Ed.) (2003). “IRT from SSI: BILOG-MG, MULTILOG, PARSCALE, TESTFACT”: Scientific Software International.

- Engle, P., Black, M. M., Behrman, J. R., Cabral de Mello, M., Gertler, P. J., Kapiriri, L., Martorell, R., Young, M. E. and the International Child Development Steering Group (2007). Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *Lancet*, 369, January 20, pp. 229-242.
- Fernald, A., Marchman, V.A., and Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science* 16:2: 234-248.
- Fernald, L.C., Kariger, P., Hidrobo M. and Gertler, P. J. (2012). "Socioeconomic gradients in child development in very young children: evidence from India, Indonesia, Peru, and Senegal." *Proceedings of the National Academy of Science* 109:2: 17273-80. On line: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3477389/?report=reader>
- Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., Pethick, S. y Reilly, J. S. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego, CA: Singular Publishing Group.
- Fernald, L. C. H., Kariger, P., Engle, P. and Raikes, A. (2009). *Examining Early Child Development in Low-Income Countries: A Toolkit for the Assessment of Children in the First Five Years of Life*. Washington D.C.: The World Bank. On line: http://siteresources.worldbank.org/INTCY/Resources/395766-1187899515414/Examining_ECD_Toolkit_FULL.pdf.
- Ferrer, G. (2006). *Sistemas de Evaluación de los Aprendizajes en América Latina Balance y Desafíos*. Santiago de Chile: PREAL.
- Fischer, K. W. (2008). Dynamic cycles of cognitive and brain development: Measuring growth in mind, brain, and education. En A. M. Battro, K. W. Fischer y P. Léna (Eds.). *The educated brain* (pp. 127-150). Cambridge U.K.: Cambridge University Press.
- Fischer, K.W. y Bidell, T.R. (2006). Dynamic development of action, thought, and emotion. In W. Damon & R.M. Lerner (Eds), *Theoretical models of human development. Handbook of child psychology* (6th ed, Vol 1, pp. 313-399). New York: Wiley.
- Gelman, S. A., Waxman, S. R, y Kleinberg, F. (2008). The role of representational status and item complexity in parent-child conversations about pictures and objects. *Cognitive Development* 23, pp. 313-323
- Gerber, R. J., Wilks, T y Erdie-Lalena, C. (2010). *Developmental Milestones : Motor Development*. *Pediatrics in Review*, 31:7, pp. 267 -277.
- Goel, V. y Dolan, R. (2000). Anatomical Segregation of Component Processes in an Inductive Inference Task. *Journal of Cognitive Neuroscience*, 12:1, pp. 1-10.
- Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., Strupp, B. and el International Child Development Steering Group (2007). Developmental potential in the first 5 years for children in developing countries. *Lancet*, 369:6, pp. 60-70.

- Greenwood, C. R., Luze, G. J., Cline, G., Kuntz, S., y Leitschuh, C. (2002). Developing a General Outcome Measure of Growth in Movement for Infants and Toddlers. *Topics in Early Childhood Special Education*, 22:3, pp. 143-157
- Hack, M., Taylor, H. G., Drotar, D., Schulchter, M., Cartar, L., Wilson-Costello, D., Klein, N., Friedman, H., Mecuri-Minich, N., y Morrow, M. (2005). Poor Predictive Validity of the Bayley Scales of Infant Development for Cognitive Function of Extremely Low Birth Weight Children at School Age. *Pediatrics*, 116:2, pp. 333 -34.
- Hamadani, J., Tofail, F., Hilaly, A., Huda, S., Engle, P., and Grantham-McGregor, S. (2010). "Use of Family Care Indicators and Their Relationship with Child Development in Bangladesh." *Journal of Health Population and Nutrition* 28(1):23-33. On line: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2975843/>
- Heckman, J. (2000). Policies to Foster Human Capital. *Research in Economics* , 54 (1), 3-56.
- Heckman, J. and Masterov D. V. (2004). The productivity argument for investing in young children. Working Paper No. 5, Invest in Young Children Working Group, Committee for Economic Development.
- Heineman, K. R., Middelburg, K. J. y Hadders-Algra, M. (2010). Development of adaptive motor behaviour in typically developing infants. *Acta Pædiatrica* 99:4 pp. 618-624.
- Heineman, K. R. y Hadders-Algra, M. (2008). Evaluation of Neuromotor Function in Infancy—A Systematic Review of Available Methods. *Journal of Developmental and Behavioral Pediatrics*, 29:4, pp. 315-323.
- Hoff, E. (2009). *Language Development*. 4th Edition. Belmont, CA: Wadsworth, Carnegie Learning.
- Hohm, E., Jennen-Steinmetz, C., Schmidt, M. H., y Laucht, M. (2007). Language development at ten months. *European Child And Adolescent Psychiatry*, 16:3, pp. 149-156.
- Iverson, J. M. (2010) Developing language in a developing body: the relationship between motor development and language development. *Journal of Child Language*, 37(2), 229–261.
- Lerner, R. M. (2006). Developmental science, developmental systems, and contemporary theories of human development. En Lerner, R. M. (Editor), *Handbook of Child Psychology. Volume 1. Theoretical Models of Development*, (pp. 1-17). New Jersey: John Wiley & Sons, Inc.
- Mara, S. (2000). Estudio de Evaluación de Impacto de la Educación Inicial en el Uruguay (Mimeo). Montevideo: Administración Nacional de Educación Primaria. Proyecto MECAEP/ANEP/BIRF.
- McCall, R. B. y Carriger, M. S. (1993). A Meta-Analysis of Infant Habituation and Recognition Memory Performance as Predictors of Later IQ. *Child Development*, 61:4, pp. 57–79.

Ministerio de Salud, MINSAL (2007). Minuta Comunicacional. Resultados II Estudio de Calidad de Vida y Salud. 12 de julio.

Morales, J., Calvo, A. and Bialystok, E. (2013). "Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*. 114 (2): 187.

Mustard, J. F. (2005). Desarrollo del cerebro basado en la experiencia temprana y su efecto en la salud, el aprendizaje y la conducta. Organización de Estados Americanos. En línea: <http://www.oas.org/udse/dit2/relacionados/archivos/desarrollo-cerebral.aspx>.

Overton, W. F. (2006). Developmental science, developmental systems, and contemporary theories of human development. En Lerner, R. M. (Editor), *Handbook of Child Psychology*. Volume 1. Theoretical Models of Development, (pp. 18-88). New Jersey: John Wiley & Sons, Inc.

Paxson, C., and N. Schady. (2007). "Cognitive Development among Young Children in Ecuador: The Roles of Wealth, Health, and Parenting." *Journal of Human Resources* 42 (1): 49–84. On line: https://www.princeton.edu/rpds/papers/Paxson_Schady_Cognitive_Development_Among_Young_Children_in_Ecuador_JHR.pdf

Pollitt, E. y Bacallao, J. (2009). Informe Técnico. La utilización de hitos motores en estudios poblacionales como indicadores del desarrollo en niños menores de dos años. Lima: OPS.

Pollitt, E. y Caycho, T. (2010) Desarrollo motor como indicador del desarrollo infantil durante los primeros años de vida. *Revista de Psicología de la PUCP*, Vol. 28 (2) p. 382-406.

Pollitt, E. (2001). The developmental and probabilistic nature of the functional consequences of iron-deficiency anemia in children. *J. Nutr.*, 131(2), 669S-675S.

Ramey, C. T., Ramey, S. L. y Lanzi, R. G. (2006). Developmental Science, Developmental Systems, and Contemporary Theories of Human Development. En Renninger, K. A. y Sigel, I. E. (Editores), *Handbook of Child Psychology*. Volume 4. Theoretical Models of Development, (pp. 864-892). New Jersey: John Wiley y Sons, Inc.

Raver, C. 2002. Emotions Matter: Making the Case for the Role of Young Children's Emotional Development for Early School Readiness. SRCD Social Policy Report, 16:3.

Reese, H. W., Lee, L. J., Cohen, S. H. y Puckett, J. M. (2001). Effects of intellectual variables, age, and gender on divergent thinking in adulthood. *International Journal of Behavioral Development*, 25:6, pp. 491-500.

Rescorla, L. y Alley, A. (2001). Validation of the Language Development Survey (LDS): A Parent Tool for Identifying Language Delay in Toddlers. *Journal of Speech, Language and Hearing Research*, 44(2), 434-435.

Ruhm, C. y Waldfogel, J. (2011). Long-term effects of early childhood care and education. IZA, Discussion Paper No. 6149.

- Saarni, C. (1999). *The Development of Emotional Competence. The Guilford Series on Social and Emotional Development*. New York: Guilford Press, Guilford Publications.
- Schady, N. (2011). "Parents' Education, Mothers' Vocabulary, and Cognitive Development in Early Childhood: Longitudinal Evidence from Ecuador." *American Journal of Public Health* 101(12): 2299-2307. On line: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3222428/#!po=58.7500>
- Schady, N. (2006). Early Childhood Development in Latin America and the Caribbean. *Economia*, 6(2), 185-225.
- Serpell, R. and Msamenang, A. B. (2014). *Locally relevant and quality ECE Programmes: Implications of research on indigenous African child development and socialization*. Paris: UNESCO.
- Shonkoff, J. P., y Phillips, D. A. (Eds.). (2000). *From Neurons to Neighborhoods: The Science of Early Childhood. Development Committee on Integrating the Science of Early Childhood Development*. Washington D.C.: National Academy Press.
- Snow, C. E. y Van Hemel, S. B. (2008). *Early Childhood Assessment. Why, What, and How*. Washington D.C.: The National Academies Press.
- Son, S. H. y Meisels, S. (2006) The Relationship of Young Children's Motor Skills to Later Reading and Math Achievement. *Merrill-Palmer Quarterly*, 52 (4), 755-778.
- Soprano, A. M. (2010). *Cómo evaluar la atención y las funciones ejecutivas en niños y adolescentes*. Buenos Aires: Paidós.
- Tamis-LeMonda, C. S., Bornstein, M. H., y Baumwell, L. (2001). Maternal Responsiveness and Children's Achievement of Language Milestones. *Child Development*, 72:3, pp. 748-767.
- Thelen, E. (2000) Motor development as foundation and future of developmental psychology. *International Journal of Behavioral Development*, 24 (4), 385-397.
- Thompson, R. A. y Nelson, C. A. (2001). Developmental science and the media: early brain. *American Psychologist*, 56, pp. 5-15.
- UNESCO (2006). *EFA Global Monitoring Report 2007. Strong Foundations. Early Childhood Care and Education*. Paris: UNESCO.
- Vegas, E. and Santibáñez, L. (2010). *La Promesa del Desarrollo en la Primera Infancia en América Latina y el Caribe*. Washington D. C.: The World Bank and Mayol Ediciones.
- Verdisco, A. (2009). "The need for regional indicators," IDB Education Newsletter 1. On line: <http://www.iadb.org/en/topics/education/early-childhood-development-the-need-for-regional-indicators,6468.html>.
- Verdisco, A. (2010). "Without data, there is no action," IDB Education Newsletter 3. On line:

<http://archive.constantcontact.com/fs085/1103432343906/archive/1103623483957.html>.

Verdisco, A., Thompson, J. and Hepworth, K. (2013). "Proyecto Regional de Indicadores de Desarrollo Infantil: PRIDI. Procesos, resultados y retos encontrados hasta la fecha," *Espacio para la Infancia*. Julio. On line: <http://www.earlychildhoodmagazine.org/es/?s=verdisco>.

Verdisco, A., Cueto, S., Thompson, J. and Neuschmidt, O. (2014). Urgency and Possibility. Results of PRIDI. A First Initiative to Create Comparative Data on Child Development in Four Latin American Countries. Washington, DC: Inter-American Development Bank.

Walker S.P., Chang, S. M, Vera-Hernandez, C. M. and Grantham-McGregor, S. M. (2011). "Early Childhood Stimulation Benefits Adult Competence and Reduces Violent Behaviour", *Pediatrics* 127(5): 849-857. On line: <http://pediatrics.aappublications.org/content/127/5/849.full.pdf+html>

Walker, S., Wachs, T. D., Gardner, J. M., Lozoff, B., Wasserman, G. A., Pollitt, E., Carter, J. A. and the International Child Development Steering Group (2007). Child development: risk factors for adverse outcomes in developing countries. *Lancet*, 369, Enero 13, 2007, 145-157.

Wetherby, A. M., Allen, L., Cleary, J., Kublin, K., y Goldstein, H. (2002). Validity and Reliability of the Communication and Symbolic Behavior Scales Developmental Profile With Very Young Children. *Journal of Speech, Language, and Hearing Research*, 45, pp., 1202-1218.

World Health Organization (2007). AnthroPlus. Online: <http://www.who.int/growthref/tools/en/>.

Young, M. E., with Richardson, L (2007). (Eds.). *Early Childhood Development: From Measurement to Action. A Priority for Growth and Equity*. Washington, DC: World Bank.