

**Validity of Phone-Based Assessment of Early
Childhood Development: Evidence from the
ECDI2030 Implementation in Nepal.**

Kenji Kitamura,^a Filipa De Castro,^b Perman Gochyyev,^c Dipu Shakya,^d
Claudia Cappa,^e Shyam Prasad Acharya,^f Devi Ram Acharya,^g and
Nicole Petrowski^e

^a Harvard Graduate School of Education, Cambridge, USA

^b Save the Children US, Washington DC, USA

^c University of California, Berkeley, USA

^d Education Section, UNICEF Nepal, Lalitpur, Nepal

^e Data and Analytics Section, Division of Data, Analytics, Planning
and Monitoring, UNICEF, New York, NY, United States

^f Education Review Office, Ministry of Education, Science and
Technology, Bhaktapur, Nepal

^g Ministry of Education, Science and Technology, Kathmandu, Nepal

Correspondence: Kenji Kitamura
Harvard Graduate School of Education
14 Appian Way, Cambridge, MA 02138, USA.
E-mail: kkitamura@g.harvard.edu

PHONE-BASED ASSESSMENT OF EARLY CHILDHOOD DEVELOPMENT

Abstract

The Early Childhood Development Index 2030 (ECDI2030) was created as a caregiver-reported measure to monitor progress toward Sustainable Development Goal target 4.2 and inform national policies. Using the ECDI2030 through phone-based administration for remote data collection is expected to contribute to efficient evidence generation, which is especially critical when traditional face-to-face data collection is not feasible. However, the existing evidence for the ECDI2030 and other caregiver-reported ECD measures are exclusively drawn from studies on assessments through face-to-face interviews. To address this, we present data from a validation trial in Nepal, comparing randomly divided computer-assisted personal interviews (CAPI) and computer-assisted telephone interviews (CATI) groups. The results of our analysis support the concurrent validity of phone-based ECDI2030 while suggesting the need to address challenges in the CATI administration. The current study contributes to the literature on ECD measurement and provides insights into the practical implications of different modes of ECD data collection.

Keywords: early child development, population-level measurement, validation, low- and middle-income countries, phone-based surveys, remote assessments

Validity of Phone-Based Assessment of Early Childhood**Development: Evidence from the ECDI2030 Implementation in
Nepal.**

Early childhood development (ECD) is a critical stage with far-reaching implications. Early interventions to enhance developmental outcomes are crucial not only for children to have better academic achievement and long-term health and economic outcomes (Campbell et al., 2014; Reynolds et al., 2011) but also for societies to prosper, as evidenced by high economic returns (Nores et al., 2005). Despite such evidence highlighting the significance of ECD, many children are at risk of not fulfilling their developmental potential (Black et al., 2017).

Recognizing the importance of the early childhood period and its problematic status globally, the United Nations Sustainable Development Goals (SDGs) included a target focused on ECD. Target 4.2 of the SDGs aims to ensure all girls and boys have access to quality early childhood development, care, and pre-primary education so that they are ready for primary education. Indicator 4.2.1, associated with this target, explicitly focuses on ECD outcomes as follows: *the proportion of children aged 24-59 months who are developmentally on track in health, learning, and psychosocial well-being* (United Nations, 2020). The Early Childhood Development Index 2030 (ECDI2030) was developed by UNICEF in collaboration with experts, partner agencies, and national statistical authorities to

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

monitor the progress toward SDG Target 4.2 through indicator 4.2.1 (UNICEF, 2021).

The ECDI2030 captures the achievement of key developmental milestones by children between the ages of 24 and 59 months based on reports from mothers (or from the children's primary caregivers if the mothers are not alive or not living in the same households) about their children's behaviors, skills, and knowledge (UNICEF, 2021).

The data are used to identify children who are developmentally on track in health, learning, and psychosocial well-being, and the resulting indicator can be used as the country's status on SDG indicator 4.2.1. Data collected through the ECDI2030 are also informative for country-level policy formulation and can shed light on the associations between governments' actions and children's development status at the population level. Furthermore, the disaggregation of results by relevant sociodemographic variables and associational analysis can help one understand disparities in the country and identify characteristics of children with greater needs.

The ECDI2030 was originally validated for use in household surveys to be administrated by interviewers in person (UNICEF, 2020). Testing and validation work followed a mixed methods approach to assess various aspects the ECDI2030: psychometric properties, suitability for implementation at scale in national household surveys, and relevance to capture the constructs of interest across different contexts (Cappa et al., 2021). Through iterative cognitive interviews across countries, the quality of

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

ECDI2030 items was tested and improved, focusing on respondents' comprehension and interpretation, which is crucial for ensuring cross-cultural relevance, reducing response bias, and minimizing measurement errors (Cappa et al., 2021). The analysis of each round of the testing resulted in the revision or removal of items found to have problems in the response process (Cappa et al., 2021). The iterative process of the testing, analysis, and revision led to the set of 20 questions considered to mitigate possible response bias (UNICEF, 2023).

Rao et al. (2021) presented evidence of concurrent validity of a caregiver report ECD measure that consists of a subset of the final version of the ECDI2030's 20 items, by showing its significant and moderate positive association with a direct ECD assessment score. The criterion validity of the caregiver report ECD measure was also provided by demonstrating its significant associations with factors expected to be related to child development, such as child age, maternal education, household wealth, child height, and weight (Rao et al., 2021).

Phone-based Early Childhood Development Assessment

In the context of the COVID-19 pandemic, data collection efforts across the world had to be adjusted to replace in-person interviews with remote ways of implementing surveys, including phone-based administration (Berman, 2020). While the existing literature supports the validity of ECDI2030 (Cappa et al., 2021; Rao et al., 2021) and

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

other caregiver-reported ECD measures, such as the Caregiver-Reported Early Development Index (CREDI; Waldman et al. 2021), the evidence regarding these measures is solely derived from studies using face-to-face caregiver interviews.

Angrist et al. (2020) explain that while the history of phone-based behavioral and learning assessments is limited, the literature presents evidence on the validity of phone-based assessments of other related constructs. Several studies focused on the phone-based assessment of cognitive function among elderly adult populations and found that phone-based measures can be comparable to gold standard measures administered face-to-face (Rankin et al., 2005; Rapp et al., 2012). Scholars also provided evidence for the validity of phone-based assessments of health outcomes among adults across low- and middle-income countries (LMICs; Greenleaf et al., 2017; Mahfoud et al., 2015). As such, the existing literature provides evidence that phone-based measures can collect valid and reliable data on certain constructs, such as cognitive function and health-related behaviors.

However, this existing evidence pertains to adult respondents reporting about themselves, and it is uncertain whether such evidence can fully apply to phone-based caregiver report ECD measures. Unlike the assessments studied in the existing literature, caregivers report about their children as proxy respondents. Caregivers of young children are required to comprehend the intentions of questions and retrieve their memories of children's

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

behavior, skills, and knowledge. In such measurements, one cannot deny the possible problem that Jäckle, Roberts, and Lynn (2010) pointed out: administration modes can elicit different responses due to its effects on respondents' motivations to make the required efforts (e.g., retrieving the relevant information) and judgments about the appropriate responses (i.e., social desirability bias). Such effects can occur due to the differences in respondents' perception of privacy in the survey setting and the legitimacy of the survey, as well as rapport between the interviewer and respondent (Jäckle et al., 2010). Phone-based caregiver report ECD measures may introduce upward bias due to social desirability bias resulting from less rapport with interviewers, and they may result in an increase of the construct-irrelevant noise due to reduced perception of privacy and legitimacy by the respondents.

Furthermore, while the ECDI2030 was designed to mitigate measurement error by removing items found to cause confusion and misunderstanding (Cappa et al., 2021), it may be observed that at least some respondents may need clarification on one or more items. Therefore, in addition to considering the potential impact of different administration modes on response processes, it is also important to keep a close eye on the challenges faced by interviewers in addressing respondents' confusion when they lack access to non-verbal cues. The extent to which interviewers can effectively address respondents' confusion or misunderstanding under such circumstances is not well understood and warrants investigation.

The Current Study

This study aims to contribute to the literature on population-level measures of ECD outcomes by illustrating how the ECDI2030 performs under different administration modes. Specifically, we compared psychometric properties and administration difficulties of the ECDI2030 in face-to-face and phone-based interviews to investigate whether the phone-based administration can generate comparable data to the standard face-to-face administration of the ECDI2030. We analyzed the data from a pilot study project conducted by the Education Review Office (ERO) under the Ministry of Education, Science and Technology in Nepal in collaboration with UNICEF in March and April 2021.

We aim to address the following research questions through concurrent validity analyses: 1) Do the distributions of face-to-face ECDI scores and phone-based ECDI scores overlap?; 2) Are the differences in response means between face-to-face ECDI scores and phone-based ECDI scores significant?; 3) Are the percentage of children specified as developmentally on track significantly different between face-to-face ECDI group and phone-based ECDI group? Additionally, we conducted item analysis to compare the performance of each question and overall score reliability between face-to-face and phone-based ECDI2030, addressing the following research questions: 4) Are the properties of ECDI items significantly different between the face-to-face and phone-based groups in terms

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

of item difficulty and homogeneity?; 5) Does any item behave differently between the face-to-face and phone-based groups after adjusting for the latent trait (i.e., across respondents with the same value of the latent trait)?; 6) Is there a difference in the estimated reliability statistics between face-to-face and phone-based ECDI? Lastly, we conducted analyses of difficulties in administration for each administration mode. to investigate the last research question: 7) Does the frequency of administration difficulties that interviewers faced varied by the administration mode?

Method

Participants

The population of this study is children aged 24-59 months who attended pre-primary schools in the form of community school-based ECD centres in three municipalities (Chautara, Kakani, and Siddhalek) in Nepal. The three municipalities were purposefully selected so that the sample included both urban (i.e., from Chautara) and rural (i.e., from Kakani and Siddhalek) populations. We conducted cluster random sampling to select children representing the studied population. First, we selected ten community school-based ECD centres from each municipality based on data from the Integrated Education Management Information System. In this selection of ECD centres, we excluded those with five or fewer children, which account for approximately 5 % of the population, for the efficiency of the data collection. We contacted the selected 30

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

ECD centres in early March 2021 to obtain the list of eligible children. Among them, 27 ECD centres responded and shared the list. Based on the list, we randomly divided the children into two groups for computer-assisted personal interview (CAPI) and computer-assisted telephone interview (CATI) with age and sex stratification¹ within each centre.

Prior to the data collection, mothers or primary caregivers of the selected children were informed by teachers or principals about the study. Among the selected 215 children in the CAPI group and 209 children in the CATI group, 89% and 82% participated, respectively. The main reasons for non-participations were challenges in contacting, such as mothers being unavailable and difficulties in obtaining accurate contact numbers, rather than a refusal to participate. Among the 364 participants, the majority were children's mothers (96%), while other caregivers (i.e., fathers, grandmothers, and other household members) responded in 15 interviews when the mothers were not alive or not living in the same households. The average age of the children was 3.75 years old (44.95 months, SD = 9.03), and 41.76 percent of them were female. It is important to note that since our sample is drawn from children in community school-based ECD centers, which primarily target children aged 48-59 months, younger children are underrepresented in the sample, despite our efforts to stratify by age. Specifically, the

¹ We created six blocks (2 years girl, 2 years boy, 3 years girl, 3 years boy, 4 years girl, 4 years boy), within which children were randomly selected for either CAPI or CATI group.

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

proportions of 2-year-olds, 3-year-olds, and 4-year-olds are 18.68%, 36.54%, and 44.78%, respectively. Additional characteristics of the study sample are presented in the first column of Table 1.

[Table 1 near here]

We note that the sample does not fully represent the national population, although the purposive sampling aimed to capture sociodemographic diversity. Table 1 shows that the participants in this study had relatively higher socioeconomic status (SES) compared to the nationally representative survey sample, the Nepal Multiple Indicator Cluster Survey 2019 (MICS, 2019; CBS, 2020). For instance, our sample demonstrated higher levels of mothers' education and a larger number of children's books at home than the general population of Nepal.

Measures

The ECDI2030 consists of 20 questions covering the three domains of child development: *health, learning, and psychosocial well-being*, aligned with the SDG indicator 4.2.1. Most of the items use a binary 'YES/NO' response scale, inquiring whether the child can or cannot perform a specific task related to the corresponding developmental domain. In cases where the respondent is unsure of the child's ability, the response 'DON'T KNOW' was recorded for the item. Two items (ECD19 and ECD20) employ a non-binary response scale to capture the frequency or extent of specific behaviors. See Supplementary Table S2 for the complete list of ECDI2030

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

questions. Originally written in English, the questions and probes were translated into Nepali by a professional translator. The translation was thoroughly discussed and edited by ERO technical officers and then by a UNICEF national education specialist.

Based on the observed raw scores, the pre-defined age-wise cut-scores, which reflect the expected increase in the number of milestones achieved by children as they grow older, were used to generate a cross-nationally comparable indicator for the proportion of children who are developmentally on track. Specifically, a minimum of 7 items was established for children aged 24 to 29 months, 9 items for children aged 30 to 35 months, 11 items for children aged 36 to 41 months, 13 items for children aged 42 to 47 months, and 15 items for children aged 48 to 59 months (UNICEF, 2023).

Caregivers also reported on sociodemographic variables based on a background questionnaire. The background questionnaire included questions on the child's age (in months) and sex, ethnicity of the respondent, language spoken at home, both parents (mothers' and fathers') formal education levels and occupations, and the number of children's books at home.

At the end of each interview, interviewers reported difficulties in interview administration that they might face through post-interview questions. The questions focused on the following four aspects of difficulties: 1) difficulties in ensuring that the interview took place in a private and quiet place; 2) interruptions or

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

disruptions during the interview; 3) difficulties in ensuring the respondent answered the questions solely based on her perception and knowledge of the child, without directly checking with the child or consulting with other household members; and 4) difficulties to check whether respondents seemed to misinterpret or not understand the questions.

Procedure

In February 2021, The ERO selected nine interviewers who met pre-set criteria (e.g., bachelor's or higher degrees, a record of ECD-related coursework and/or relevant professional experience) and performed well in a mock phone-based interview using a subset of ECDI2030 items and a few background questions. In March 2021, a three-day in-person training was conducted. The first day focused on the review of the administration guideline and item-by-item instruction based on standard training materials for the ECDI2030 (UNICEF, 2020), followed by practice implementation of the ECDI2030 in the standard face-to-face administration with ERO staff. On the second day, the interviewers visited nearby schools to administer the ECDI2030 questionnaire through the standard face-to-face interview with two mothers of children attending the pre-primary classes. Then, they interviewed one or two mothers over the phone. The last day of the training focused on the reflection of challenges and solutions based on the experience of the field practices.

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

Upon completing the training, three teams, each consisting of three interviewers and one supervisor from the ERO, were grouped, and each of them visited one of the study sites for the data collection. The interviewers collected the data through either CAPI or CATI. The interviewers obtained consent from respondents before the interview based on a form that explained the purpose and procedure of the study and informed participants that their responses were voluntary and kept confidential.

Analysis Plan

The analytical scope of this study includes concurrent validity analyses, item analysis, instruments' reliability analysis, and administration difficulties analyses. The concurrent validity analyses focused on comparing ECDI2030 raw scores and criterion-based classification between the CAPI and CATI groups. The study design is an independent sample approach where participants differ across administration modes. Thus, the analytical validity depends on the assumption that the two groups are interchangeable in their characteristics and that the only difference is the assessment administration mode (Jäckle et al., 2010). To test this assumption, we conducted a covariate balance check based on standardized mean differences (Cohen's d) and variance ratios (r). The literature suggests that magnitudes of $|d| > .2$ (Abdia et al., 2017) and variance ratios outside of $[4/5, 5/4]$ (Rubin, 2001) suggest the concern of

imbalance. The second and third columns of Supplementary Table S1 reveal that a few variables are out of balance.

To address these background variable imbalances, we employed a propensity score (PS) based weighting approach. The PS is the conditional probability that children were in the CATI group rather than the CAPI group. The PS was estimated using logistic regression with all covariates presented in Supplementary Table S1. Based on the computed PS, we computed the inverse probability weights (IPW) where the probability of treatment is the conditional probability calculated based on the above model. The IPW was incorporated into the balance check. The fourth and fifth columns of Supplementary Table S1 show that this approach improves the balance, and all variables meet the balance criteria, indicating that the two groups can be considered interchangeable with IPW.

In the concurrent validity analysis, we weighted observations with the IPW. We began with a visual inspection of the ECDI2030 raw score distributions based on weighted kernel density estimations to address the first research question regarding the score distributions. Then, we conducted an independent t-test with the IPW to address the second research question concerning the raw score mean difference. Next, we compared percentages of children specified as developmentally on track based on the age-specific cut-off, using weighted chi-square tests, to address the third research question.

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

As for the item analyses, we started with classical test theory-based analyses, using two common indices: item difficulty index (p) (the percentage of respondents who answered 'YES') and item homogeneity index (adjusted item-test point biserial correlation: r_{pb}), and we compared these indices between the groups. As a scale-level measure related to item homogeneity, we calculated reliability in terms of internal consistency for each administration mode. We also examined if there are any differences in the likelihood of 'Don't Know (DK)' responses between the groups.

Additionally, we analyzed differential item function (DIF). First, we regressed item endorsement on the ECDI2030 score and administration mode through logistic regression with IPW in the first model. In the second model, we added the interaction between ECDI2030 score and administration mode as follows:

where significant β_1 and β_2 indicate that respondents have different probabilities of endorsing a given item between CAPI and CATI after controlling for overall ECDI2030 scores. We further conducted item response theory-based DIF analysis to investigate whether any item behaves differently across respondents with the same latent trait value (i.e., unobserved ECD outcomes) due to different administration modes, using Mantel-Haenszel (MH) tests (Holland & Thayer, 1988). Finally, we conducted analyses of difficulties in administration for each administration mode. We investigated

whether the frequencies of administration difficulties that interviewers faced varied by administration mode.

Results

Concurrent Validity Analysis Results

We began with a visual inspection of the ECDI2030 raw score distributions based on the weighted kernel density estimation. Figure 1 shows that the raw score distributions of two administration modes overlap across the score range. The result of the weighted independent t-test reveals that the difference in mean ECDI2030 scores between the CAPI group (14.36, se = 0.24) and the CATI group (14.27, se = 0.26) is very small and not statistically significant ($p = .80$). We also compared percentages of children identified as developmentally on track based on the age-specific cut-off through a weighted chi-square test. The result shows that the difference in the proportions of developmentally on-track children between the CAPI and CATI groups is very small (i.e., less than 1%) and not statistically significant ($\chi^2 = 0.05, p = .83$).

[Figure 1 near here]

Item Analysis Results

Table 2 presents the three indices for all the ECDI2030 items for the CAPI and CATI groups. Looking first at the results for item difficulties (the percentage of respondents who answered 'YES'), the fourth column shows no substantial differences in difficulties of any item by administration mode. The magnitudes of differences in item

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

difficulty for most items are within 0.03: that is, the proportions of answering “Yes” are different between the CAPI and CATI by three percent or less. Regarding the directions of differences, the CAPI group showed higher difficulty levels in 11 items, while the CATI group had higher difficulty levels in nine items, suggesting the absence of bias in item difficulty by the administration mode.

[Table 2 near here]

Similarly, the adjusted item-test correlations (item-rest correlations) are comparable between the CATI and CAPI groups. While some items showed relatively large differences in adjusted item-test correlations (i.e., ECD5, ECD8, ECD9, ECD13, and ECD16), the directions of the differences are different across items, suggesting no systematic pattern. Table 2 also presents the proportions of DK responses for each item by administration mode. The last column shows no statistically significant differences in the proportion of DK responses by administration mode in any items. Furthermore, the average number of DK responses among the 20 ECDI2030 items was not statistically significant ($p = .26$).

The examination of scale reliability in terms of internal consistency further supports this finding on the items’ homogeneity: Cronbach’s alphas for the CAPI-based ECDI2030 and the CATI-based ECDI2030 were estimated at .76 and .75, respectively. These values indicate that the scores meet levels generally considered acceptable ($\geq .70$; Taber, 2018), demonstrating good internal consistency of the scale across different administration modes.

[Table 3 near here]

Regarding the DIF analysis, Table 3 represents the results of the analyses with only the main effect of administration mode (Model 1; for uniform DIF analysis) and the ones with the interaction terms (Model 2; for nonuniform DIF analysis). While none of the 20 items have uniform DIF, the second model finds that the coefficient for the interaction term is significant at a 95% confidence level for ECD13 (“child gives 3 objects”). This significant interaction term indicates that the holistic ECD outcome measured by ECDI2030 scores is more strongly correlated with the probability of endorsing ECD13 when the assessment is conducted through CATI rather than CAPI: among the relatively lower-development group, ECD13 is more difficult in CATI than in CAPI, whereas the same item is easier in CATI than in CAPI among the higher-development group. However, this significant nonuniform DIF may be spurious, given the number of statistical tests we performed. The risk of a false positive (i.e., Type I error) is $1 - 0.95^{20}$, approximately 64%. Moving to the DIF analysis based on the MH tests, Table 4 shows the absence of DIF across items.

[Table 4 near here]

Administration Difficulties Analysis Results

In investigating the administration difficulties in each mode of the administration, we first looked at the proportion of interviews in which the interviewers faced any of the four kinds of difficulties by mode (see Table 5). The most prevalent issue was interruptions or

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

disruptions during the interview (*FQ2. Disruptions* in Table 5): The interviewers experienced interruptions or disruptions in more than one-fifth of their interviews. This is followed by the difficulties in ensuring that the interview took place in a private and quiet place (*FQ1. Interview place*) and by difficulties in checking whether respondents seemed to misinterpret or not understand the questions (*FQ4. Misinterpretation*). The problem of *Interview place* and *Disruptions* are significantly related ($r = .60, p < .001$), suggesting that disruptions are likely to occur when interviews cannot be conducted in a private and quiet place.

[Table 5 near here]

The results of the weighted chi-square tests reveal that while the likelihoods of occurrence of the three most prevalent administration difficulties were not different by the administration mode, the interviewers were marginally more likely to experience difficulties in ensuring responses without consultation with others or direct check with the child (*FQ3. Consultation/Check*) in CATI than in CAPI ($p = .07$). Such difficulties in ensuring responses without consultation or direct check were more common when interviewers had difficulties in ensuring a quiet and private place for the interview, but such a relationship between these two difficulties was observed only in CATI but not in CAPI: while the correlation between *Interview place* and *Consultation/Check* is -0.01 ($p = .89$) for CAPI, it is much stronger for CATI ($r = 0.38, p < .001$).

Discussion

The phone-based administration of the ECDI2030 can greatly contribute to the population-level data generation helpful for evidence-based policy discussions. Given the safety concerns associated with traditional face-to-face data collection during the pandemic, there has been a growing demand for remote survey implementation (Berman, 2020). Moreover, as the data show that the number of mobile subscriptions in LMICs was 104 per 100 people in 2021 (World Bank, n.d.), the access to mobile phones has been expanding globally (96% of households have a phone in Nepal in 2019; CBS, 2019). The special needs under the pandemic and the expanding phone access present the demand and opportunity to collect important data on children safely and efficiently. However, the existing literature provides limited knowledge about the validity of ECD measurement through phone-based surveys. To address this gap, the current study analyzed multiple aspects of validity and reliability of the phone-based ECDI2030 by comparing it to the standard face-to-face administered ECDI2030 data in Nepal.

To examine the mode effect (i.e., the effects of different administration modes on caregivers' response), we employed an independent group approach by randomly assigning participants to one of the administration modes. The results show that the ECDI2030 data from the two administration modes are comparable in terms of overall score distributions, mean scores, and the proportions of of developmentally on-track children. These results

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

suggest that CATI-based ECDI2030 scores can capture the same pattern of responses on child development outcomes as standard CAPI-based ECDI2030 scores. Additionally, the item analyses revealed that the phone-based administration ensured that the ECDI2030 items were comparable to the data collected through the standard administration mode. This evidence reinforces the validity of ECDI2030 data irrespective of mode of administration.

Regarding the concern in the literature (Jäckle et al., 2010) that the administration modes can influence respondents' motivations to make the required efforts to respond (e.g., retrieving the relevant information), our study found indirect evidence for the absence of such an effect in the phone-based ECDI2030. If respondents are less motivated in one administration mode than the other, it is likely that they indiscriminately use one response category (e.g., keep saying "Yes"), which affects item homogeneity and internal consistency, or choose the "DK" response category. However, our study found no substantial differences in the item homogeneity index, internal consistency, and proportions of DK responses by administration mode.

Of note, the questionnaire in this study was concise (i.e., consisting of the 20 ECDI2030 questions and several sociodemographic questions), and it is likely shorter than survey questionnaires in which practitioners are interested in embedding the ECDI2030. The possible effects of administration mode on responses due to different amounts of effort needed to respond or

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

the respondent's motivation may vary depending on the length and complexity of the interview. The extent to which the evidence in the current study can be generalized to longer surveys is unknown. Thus, one needs to be cautious about the administration mode effect, especially when the ECDI2030 is integrated into a long survey. It is also recommended that those interested in conducting phone-based assessments follow the recommendation in the literature to keep the interview short to ensure the data quality (Angrist et al., 2020).

The literature also suggests that administration modes may affect respondents' judgments about the appropriate responses (i.e., social desirability bias) (Jäckle et al., 2010). Our study found no systematic differences in responses by mode at scale and item levels. However, the analysis of administration difficulty shows that respondents were more likely to consult other household members or directly check with the child in the phone-based administration than in the standard face-to-face administration. Such consultation and direct checks were especially prevalent when phone-based interviews were not conducted in private and quiet places. This suggests the greater challenge in standardizing the interview setting in the phone-based administration. Also, the interviewers reported that respondents, mostly mothers of the children, were more likely to seek help from other household members when they felt nervous or confused during the phone-based ECDI2030 administration. This finding implies the particular importance of rapport-building in phone-based administration. The importance of rapport building and

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

interview techniques to prevent consultations and direct checks

during interviews needs to be emphasized in training for

interviewers.

Limitations

We note limitations in this study. The current study's sample does not fully represent the general population of Nepal. Not only does this limitation in the sample have implications for the findings' external validity, but it also has to do with the variability of the ECDI2030 data. As discussed above, younger children are under-represented in the sample despite our use of the age-wise stratification: in fact, the mean of the youngest age group of children (2 years old) are closer to the upper range (35 months) than to the lower range (24 months) (i.e., mean is 30.84). Furthermore, according to the Nepal MICS 2019 (CBS, 2020), children in the Bagmati province, where our sample was drawn, showed better development outcomes than those in other provinces, and children who attended ECD centres showed better development outcomes than those who did not. Therefore, the current sample can be considered a relatively homogeneous and high-development group. Due to such limitations of the sample, some psychometric properties, such as item difficulty, homogeneity, and internal consistency, might be affected. Nevertheless, it should be emphasized that such evidence does not imply a validity problem with the assessment tool itself. In the future, conducting further testing of CATI

administration with samples that include children from lower resource settings would be beneficial.

Also, given the limited representativeness of the study sample, we caution against inferring the population-level development status of children in Nepal based on the data presented in this study. The focus of this study was to investigate whether the different administration modes of the ECDI2030 cause measurement errors. Due to this study aim and considering practical constraints caused by difficulties in the mobility under the pandemic, we did not take special measures to ensure national representativeness of the sample. However, it is important to note that phone-based surveys require special considerations in sample development, as they may introduce biases by primarily targeting wealthier populations who are more accessible via phone calls. Those interested in employing phone-based ECDI2030 to establish nationally representative evidence are advised to consult the existing literature on sampling methods for phone-based surveys (Lavrakas et al., 2017; Lepkowski et al., 2008).

Conclusion

The current study contributes to the literature related to ECD measures and mode comparison research by adding new knowledge on the feasibility of obtaining valid and reliable data on ECD based on caregiver reports collected through phone-based surveys. Specifically, our findings demonstrate that the phone-based

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

administration of the ECDI2030 can generate scores that exhibit a raw mean and proportion of developmentally on track children comparable to those generated by standard face-to-face administration. The analyses also indicate that the ECDI2030 from the phone-based and face-to-face administrations had comparable item properties, with no significant differential item functioning. These findings collectively provide evidence supporting the concurrent validity of the phone-based ECDI2030. The current study also provides informative lessons and suggestions regarding administration difficulties for policymakers and practitioners interested in collecting important data on children in a timely and safe manner, particularly when the traditional face-to-face administration is difficult or impractical.

References

- Abdia, Y., Kulasekera, K. B., Datta, S., Boakye, M., & Kong, M. (2017). Propensity scores based methods for estimating average treatment effect and average treatment effect among treated: A comparative study. *Biometrical Journal*, 59(5), 967-985. <https://doi.org/10.1002/bimj.201600094>
- Angrist, N., Bergman, P., Evans, D. K., Hares, S., Jukes, M. C., & Letsomo, T. (2020). Practical lessons for phone-based assessments of learning. *BMJ Global Health*, 5(7), e003030. doi: <http://dx.doi.org/10.1136/bmjgh-2020-003030>
- Berman, G. (2020). Ethical Considerations for Evidence Generation Involving Children on the COVID-19 Pandemic: Innocenti Discussion Papers, no. 2020-01. UNICEF Office of Research. <https://www.unicef->

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

irc.org/publications/1086-ethical-considerations-for-evidence-generation-involving-children-on-the-covid-19.html

Black, M. M., Walker, S. P., Fernald, L. C. H., Andersen, C. T., DiGirolamo, A. M., Lu, C., ... and Lancet Early Childhood Development Series Steering Committee. (2017). Early childhood development coming of age: Science through the life course. *Lancet*, 389(10064), 77-90. doi: [10.1016/S0140-6736\(16\)31389-7](https://doi.org/10.1016/S0140-6736(16)31389-7)

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., & Pan, Y. (2014).

Early childhood investments substantially boost adult health.

Science, 343(6178), 1478-1485. doi: [10.1126/science.1248429](https://doi.org/10.1126/science.1248429)

Cappa, C., Petrowski, N., De Castro, E. F., Geisen, E., LeBaron, P., Allen-Leigh, B., ... &

Scanlon, P. J. (2021). Identifying and Minimizing Errors in the Measurement of Early Childhood Development: Lessons Learned from the Cognitive Testing of the ECDI2030. *International journal of environmental research and public health*, 18(22), 12181. doi: [10.3390/ijerph182212181](https://doi.org/10.3390/ijerph182212181)

Central Bureau of Statistics (CBS). (2020). Nepal Multiple Indicator Cluster Survey 2019,

Survey Findings Report. Kathmandu, Nepal: Central Bureau of Statistics and UNICEF Nepal.

<https://www.unicef.org/nepal/reports/multiple-indicator-cluster-survey-final-report-2019>

Greenleaf, A. R., Gibson, D. G., Khattar, C., Labrique, A. B., & Pariyo, G. W. (2017). Building

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

the evidence base for remote data collection in low-and middle-income countries: comparing reliability and accuracy across survey modalities. *Journal of medical Internet research*, 19(5), e140.

doi: [10.2196/jmir.7331](https://doi.org/10.2196/jmir.7331)

Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel

procedure. In Wainer H & Braun HI (Eds.), *Test validity* (pp. 129–145). Hillsdale, NJ, US.

Jäckle, A., Roberts, C., & Lynn, P. (2010). Assessing the effect of data collection mode on

measurement. *International Statistical Review*, 78(1), 3-20. doi:

[10.1111/j.1751-5823.2010.00102.x](https://doi.org/10.1111/j.1751-5823.2010.00102.x)

Lavrakas, P., Benson, G., Blumberg, S., Buskirk, T., Cervantes, I. F., Christian, L., ... &

Shuttles, C. (2017). The future of US general population telephone survey research. AAPOR Report.

Lepkowski, J. M., Tucker, N. C., Brick, J. M., de Leeuw, E. D., Japec, L., Lavrakas, P. J., Link,

M. W., & Sangster, R. L. (Eds.). (2008). *Advances in telephone survey methodology*.

John Wiley & Sons.

Mahfoud, Z., Ghandour, L., Ghandour, B., Mokdad, A. H., & Sibai, A. M. (2015). Cell phone

and face-to-face interview responses in population-based surveys: how do they compare?. *Field methods*, 27(1), 39-54. doi:

[10.1177/1525822X14540084](https://doi.org/10.1177/1525822X14540084)

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

Nores, M., Belfield, C. R., Barnett, W. S., & Schweinhart, L. (2005).

Updating the economic

impacts of the High/Scope Perry Preschool program. *Educational*

Evaluation and Policy Analysis, 27(3), 245-261. doi:

[10.3102/01623737027003245](https://doi.org/10.3102/01623737027003245)

Rankin, M. W., Clemons, T. E., & McBee, W. L. (2005). Correlation analysis of the in-clinic

and telephone batteries from the AREDS cognitive function ancillary

study. AREDS Report No. 15. *Ophthalmic Epidemiology*, 12(4), 271-

277. doi: [10.1080/09286580591003815](https://doi.org/10.1080/09286580591003815)

Rao, N., Chan, S. W., Su, Y., Richards, B., Cappa, C., De Castro, E. F., & Petrowski, N.

(2021). Measuring Being “Developmentally on Track”: Comparing

Direct

Assessment and Caregiver Report of Early Childhood Development in

Bangladesh, China, India and Myanmar. *Early Education and*

Development, 1-

23. doi: [10.1080/10409289.2021.1928446](https://doi.org/10.1080/10409289.2021.1928446)

Rapp, S. R., Legault, C., Espeland, M. A., Resnick, S. M., Hogan, P. E., Coker, L. H., ... &

CAT Study Group. (2012). Validation of a cognitive assessment

battery administered

over the telephone. *Journal of the American Geriatrics Society*, 60(9),

1616-1623. doi:

[10.1111/j.1532-5415.2012.04111.x](https://doi.org/10.1111/j.1532-5415.2012.04111.x)

Reynolds, A. J., Temple, J. A., Ou, S. R., Arteaga, I. A., & White, B. A.

(2011). School-based

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

early childhood education and age-28 well-being: Effects by timing, dosage, and subgroups. *Science*, 333(6040), 360-364. doi:

[10.1126/science.1203618](https://doi.org/10.1126/science.1203618)

Rubin, D. B. (2001). Using Propensity Scores to Help Design Observational Studies:

Application to the Tobacco Litigation. *Health Services and Outcomes Research*

Methodology, 2, 169-188. doi: [10.1023/A:1020363010465](https://doi.org/10.1023/A:1020363010465)

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research

Instruments in Science Education. *Research in Science Education*, 48(6), 1273-1296.

<https://doi.org/10.1007/s11165-016-9602-2>

UNICEF. (2020). *ECDI2030: Instructions for Interviewers*. UNICEF.

UNICEF. (2023). *The Early Childhood Development Index 2030: A new measure of early*

childhood development. UNICEF, New York, 2023.

UNICEF. (2021). *Early childhood development index 2030: A new tool to measure*

SDG indicator 4.2.1. Author. <https://data.unicef.org/resources/early-childhood-development-index-2030-ecdi2030/>

United Nations. (2020). Global indicator framework for the Sustainable Development Goals and

targets of the 2030 agenda for sustainable development. Author.

https://unstats.un.org/sdgs/indicators/Global%20Indicator%20Framework%20after%202021%20refinement_Eng.pdf

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN
NEPAL

Waldman, M., McCoy, D. C., Seiden, J., Cuartas, J., & Fink, G. (2021).

Validation of

motor, cognitive, language, and socio-emotional subscales using the
Caregiver Reported Early Development Instruments: An application
of multidimensional item factor analysis. *International Journal of
Behavioral Development*, 45(4), 368–377.

<https://doi.org/10.1177/01650254211005560>

World Bank. (n.d.) Mobile cellular subscriptions (per 100 people).

Retrieved from

<http://data.worldbank.org/indicator/IT.CEL.SETS.P2>

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN
NEPAL

Table 1. Sample and population characteristics

Variable Name	Sample (n=364)	National Population* (n=4196)
CHILD AGE		
% of 2 years old	18.68%	30.19%
% of 3 years old	36.54%	35.70%
% of 4 years old	44.78%	34.10%
CHILD SEX (Female)	41.76%	47.86%
HOME LANGUAGE		
NEPALESE	77.47%	44.49%
OTHER LANG	22.53%	55.51%
MOTHER'S EDUCATION		
NONE	21.33%	27.93%
BASIC (Gr 1-8)	37.67%	33.45%
LOWER BASIC (Gr 1-5)	20.78%	16.64%
UPPER BASIC (Gr 6-8)	16.90%	16.81%
SCONDARY (Gr 9-12)	35.46%	31.62%
LOWER SCONDARY (Gr 9-10)	20.78%	19.80%
UPPER SCONDARY (Gr 11-12)	14.68%	11.82%
HIGHER	5.54%	6.97%
FATHER'S EDUCATION		
NONE	9.30%	13.36%
BASIC (Gr 1-8)	49.01%	41.35%
LOWER BASIC (Gr 1-5)	25.92%	20.40%
UPPER BASIC (Gr 6-8)	23.10%	20.95%
SCONDARY (Gr 9-12)	36.34%	34.39%
LOWER SCONDARY (Gr 9-10)	22.25%	21.89%
UPPER SCONDARY (Gr 11-12)	14.08%	12.50%
HIGHER	5.35%	10.89%
BOOKS (Standard Deviation)	1.22 (2.06)	.46 (.03)

Note: *The national population is children aged 24-59 months regardless of their attendance in ECD centres. The data is based on the MICS 2019.

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

Table 2. Item difficulty, homogeneity statistics and analysis of DK responses

Item	Item difficulty (p)			Adjusted Item-Total Correlation (r_{pb})		Proportion of DK response		
	CAPI	CATI	diff	CAPI	CATI	CAPI	CATI	value
ECD1	0.80	0.78	0.01	0.28	0.24	1.19%	1.06%	.01
ECD2	0.83	0.84	-0.01	0.22	0.15	6.34%	6.57%	.07
ECD3	0.66	0.68	-0.02	0.36	0.35	1.68%	1.95%	.03
ECD4	0.42	0.41	0.01	0.44	0.38	8.78%	4.62%	2.03
ECD5	0.97	0.98	0.00	0.30	0.13	0.00%	0.61%	1.02
ECD6	0.94	0.92	0.02	0.31	0.31	0.00%	0.00%	NA
ECD7	0.70	0.67	0.04	0.41	0.39	4.83%	3.42%	.36
ECD8	0.89	0.83	0.06	0.33	0.46	0.86%	0.44%	.32
ECD9	0.95	0.95	0.01	0.34	0.24	0.00%	0.35%	1.01
ECD10	0.50	0.54	-0.04	0.55	0.52	4.96%	2.94%	1.00
ECD11	0.22	0.22	0.00	0.44	0.47	2.55%	1.13%	.80
ECD12	0.46	0.49	-0.03	0.56	0.49	4.92%	2.54%	1.09
ECD13	0.67	0.64	0.02	0.26	0.46	17.23%	12.62%	1.36
ECD14	0.42	0.42	0.00	0.43	0.45	9.05%	5.97%	0.96
ECD15	0.50	0.50	0.00	0.27	0.19	11.19%	11.52%	.01
ECD16	0.96	0.96	0.00	0.16	0.29	1.40%	0.00%	1.01
ECD17	0.73	0.72	0.01	0.13	0.15	5.39%	5.83%	.03
ECD18	0.92	0.93	-0.01	0.16	0.23	0.47%	1.83%	1.61
ECD19	0.95	0.93	0.03	0.11	0.09	2.81%	5.19%	1.34
ECD20	0.87	0.87	-0.01	0.14	0.21	2.93%	5.51%	.09

Note: *** $p < .001$, ** $p < .01$, * $p < .05$. All analyses were conducted with IPW.

Table 3. Logistic regression table

Item	Model 1	Model 2
ECD1	-.06 (.28)	-.13 (.32)
ECD2	.14 (.30)	.06 (.34)
ECD3	.17 (.27)	.15 (.28)
ECD4	-.02 (.26)	.06 (.28)
ECD5	.16 (.67)	-1.60 (1.89)
ECD6	-.39 (.47)	-.58 (.93)
ECD7	-.17 (.28)	-.23 (.31)
ECD8	-.62 (.39)	-.17 (.65)
ECD9	-.06 (.57)	-1.18 (.96)

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

ECD10	.37 (.28)	.38 (.29)	-.38 (.45)
ECD11	-.02 (.35)	-.10 (.70)	.11 (.71)
ECD12	.28 (.28)	.37 (.31)	-.65 (.54)
ECD13	-.09 (.27)	.06 (.28)	.67* (.33)
ECD14	.07 (.27)	.08 (.30)	-.04 (.38)
ECD15	.00 (.24)	.02 (.24)	-.27 (.28)
ECD16	.13 (.65)	1.99 (1.14)	1.42 (.79)
ECD17	-.06 (.26)	-.05 (.26)	.04 (.26)
ECD18	.18 (.43)	.52 (.47)	.41 (.39)
ECD19	-.47 (.47)	-.55 (.52)	-.13 (.38)
ECD20	.09 (.34)	.21 (.43)	.23 (.39)

Note: *** $p < .001$, ** $p < .01$, * $p < .05$. All the analyses were conducted with the IPW.

Table 4. Mantel-Haenszel test statistics

Item	value	<i>p-value</i>	Odds ratio
ECD1	0.14	0.71	0.86
ECD2	0.11	0.74	1.15
ECD3	0.40	0.53	1.21
ECD4	0.02	0.88	0.93
ECD5	0.02	0.90	0.87
ECD6	0.06	0.81	0.80
ECD7	0.14	0.71	0.87
ECD8	3.83	0.05	0.44
ECD9	0.00	0.95	0.85
ECD10	2.29	0.13	1.63
ECD11	0.02	0.90	1.02
ECD12	0.49	0.48	1.26
ECD13	0.04	0.84	0.92
ECD14	0.00	0.99	1.03
ECD15	0.01	0.93	1.05
ECD16	0.01	0.91	0.89
ECD17	0.01	0.94	0.95
ECD18	0.02	0.88	0.97
ECD19	0.01	0.92	0.94
ECD20	0.00	0.97	1.07

Note: *** $p < .001$, ** $p < .01$, * $p < .05$.

Table 5. Chi-square tests of administration difficulties by administration mode

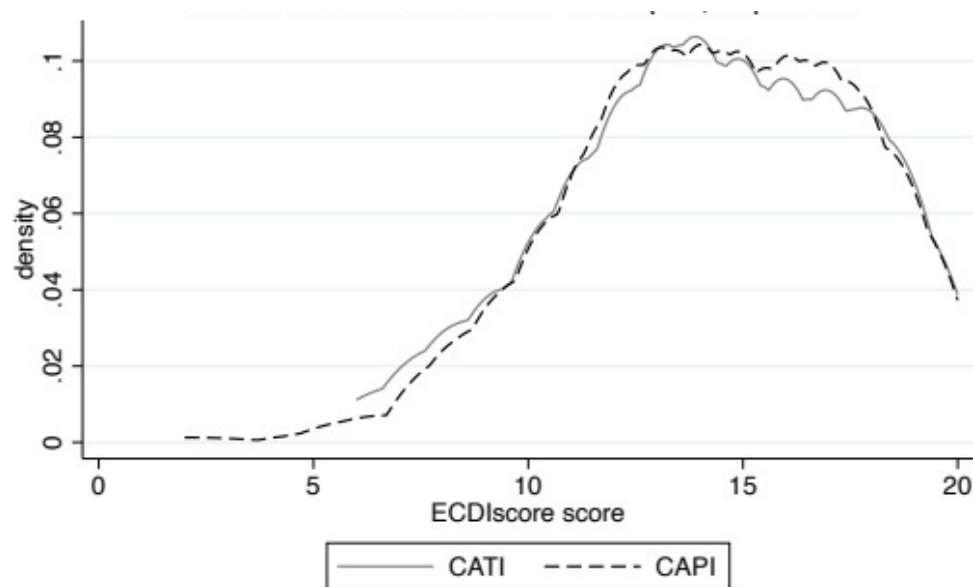
	CAPI n = 192	CATI n = 172	value	<i>p-value</i>
Difficulties				
FQ1. Interview place	15.3%	13.3%	.26	.61
FQ2. Disruptions	22.1%	22.7%	1.34	.26
FQ3. Consultation/Check	4.9%	12.3%	2.75	.07
FQ4. Misinterpretation	10.6%	10.3%	.04	.95

Note: *** $p < .001$, ** $p < .01$, * $p < .05$. The proportion reported here indicates the proportion of interviews in which interviewers experienced a given

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

difficulty. The three response categories that indicate difficulties in FQ2-FQ4 (2 "Only sometimes," 3 "Often," 4 "Always") were combined into one category for the calculation of the proportions (second and third columns). We used original categories in the chi-square tests. The analyses were conducted with IPW.

Figure 1. Weighted kernel density of ECDI2030 scores by administration mode



VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

Supplementary files

Table S1. Balance check

Variable Name	Baseline		IPTW	
	d	r	d	r
Number of eligible children (%)				
Chautara	-0.08		-0.03	
Kakani	0.03		0.01	
Siddhalek	0.05		0.01	
AGE (in month)	0.07	1.09	0.01	1.16
% of 2 years old	0.00		0.06	
% of 3 years old	-0.13		-0.12	
% of 4 years old	0.13		0.07	
SEX (Female)	-0.04		0.00	
CAST/ETHNICITY				
BRAHMIN	0.07		0.00	
CHETTRI	0.01		0.00	
JANAJATI	-0.08		-0.02	
DALIT	-0.12		-0.01	
OTHERS	0.09		0.02	
DON'T KNOW	0.14		0.03	
HOME LANGUAGE				
NEPALESE	-0.06		0.01	
OTHER LANG	0.06		-0.01	
MOTHER'S EDUCATION				
NONE	-0.01		0.01	
BASIC (Gr 1-8)	-0.10		-0.02	
LOWER BASIC (Gr 1-5)	-0.12		-0.03	
UPPER BASIC (Gr 6-8)	0.01		0.01	
SCONDARY (Gr 9-12)	0.10		0.02	
LOWER SCONDARY (Gr 9-10)	0.04		0.00	
UPPER SCONDARY (Gr 11-12)	0.09		0.02	
HIGHER	0.03		0.00	
FATHER'S EDUCATION				
NONE	0.09		-0.01	
BASIC (Gr 1-8)	-0.14		-0.01	
LOWER BASIC (Gr 1-5)	-0.14		-0.02	
UPPER BASIC (Gr 6-8)	-0.02		0.00	
SCONDARY (Gr 9-12)	0.02		0.01	
LOWER SCONDARY (Gr 9-10)	0.07		0.01	
UPPER SCONDARY (Gr 11-12)	-0.05		0.00	
HIGHER	0.10		0.00	
MOTHER'S OCCUPATION				
AGRICULTURE	-0.25*		0.00	
LABOR	-0.12		-0.03	
BUSINESS	-0.06		-0.01	
EMPLOYEE	0.04		0.00	
NOT WORKING	0.24*		0.01	
OTHERS	0.15		0.02	
FATHER'S OCCUPATION				
AGRICULTURE	0.12		0.09	
LABOR	-0.10		0.00	
BUSINESS	-0.07		-0.01	
EMPLOYEE	0.04		0.00	
OREIGN	0.00		0.01	
NOT WORKING	0.17		-0.01	
OTHERS	-0.10		-0.10	
BOOKS	0.11	1.30*	0.00	1.06

Note: *indicates that the variable is a concern of imbalance suggested in the literature (Abdia et al., 2017; Rubin, 2001).

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

Table S2. Early Childhood Development Index 2030 questionnaire

ECD1. Can (name) walk on an uneven surface, for example, a bumpy or steep road, without falling?	YES 1 NO 2 DK 8
ECD2. Can (name) jump up with both feet leaving the ground?	YES 1 NO 2 DK 8
ECD3. Can (name) dress (him/herself), that is, put on pants and a shirt, without help?	YES 1 NO 2 DK 8
ECD4. Can (name) fasten and unfasten buttons without help?	YES 1 NO 2 DK 8
ECD5. Can (name) say 10 or more words, like 'mama' or 'ball'?	YES 1 NO 2 DK 8
ECD6. Can (name) speak using sentences of 3 or more words that go together, for example, "I want water" or "The house is big"?	YES 1 NO 2 DK 8
ECD7. Can (name) speak using sentences of 5 or more words that go together, for example, "The house is very big"?	YES 1 NO 2 DK 8
ECD8. Can (name) correctly use any of the words 'I,' 'you,' 'she,' or 'he,' for example, "I want water" or "He eats rice"?	YES 1 NO 2 DK 8
ECD9. If you show (he/she) an object (he/she) knows well, such as a cup or animal, can (he/she) consistently name it? By consistently we mean that (he/she) uses the same word to refer to the same object, even if the word used is not fully correct.	YES 1 NO 2 DK 8
ECD10. Can (name) recognize at least 5 letters of the alphabet?	YES 1 NO 2 DK 8

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

 8
ECD11. Can (name) write (his/her) name?	YES 1 NO 2 DK 8
ECD12. Can (name) recognize all numbers from 1 to 5?	YES 1 NO 2 DK 8
ECD13. If you ask (name) to give you 3 objects, such as 3 stones or 3 beans, does (he/she) give you the correct amount?	YES 1 NO 2 DK 8
ECD14. Can (name) count 10 objects, for example 10 fingers or 10 blocks, without mistakes?	YES 1 NO 2 DK 8
ECD15. Can (name) do an activity, such as colouring or playing with building blocks, without repeatedly asking for help or giving up too quickly?	YES 1 NO 2 DK 8
ECD16. Does (name) ask about familiar people other than parents when they are not there, for example, "Where is Grandma?"?	YES 1 NO 2 DK 8
ECD17. Does (name) offer to help someone who seems to need help?	YES 1 NO 2 DK 8
ECD18. Does (name) get along well with other children?	YES 1 NO 2 DK 8
ECD19. How often does (name) seem to be very sad or depressed? Would you say: daily, weekly, monthly, a few times a year, or never?	DAILY 1 WEEKLY 2 MONTHLY..... 3 A FEW TIMES A YEAR 4 NEVER 5 DK

VALIDATION OF ECDI2030 FOR PHONE-BASED ADMINISTRATION IN NEPAL

 8
ECD20. Compared with children of the same age, how much does (<i>name</i>) kick, bite, or hit other children or adults? Would you say: not at all, the same or less, more, or a lot more?	<p>NOT AT ALL 1</p> <p>THE SAME OR LESS 2</p> <p>MORE 3</p> <p>A LOT MORE 4</p> <p>DK 8</p>