Measuring Psychosocial Adversity as a Neurodevelopmental Risk Factor: Development & Validation of the Childhood Psychosocial Adversity Questionnaire–Bangladesh (CPAQ-B)

Table of Contents

Main Tables & Figures

Tables:

Figures:

Appended Tables & Figures

Tables:

Figures:

Glossary of abbreviations

ACE	Adverse Childhood Experience
BEAN Study	Bangladesh Early Adversity Neuroimaging Study
Crypto Study	Burden of Cryptosporidiosis Study
CPAQ-B	Child Psychosocial Adversity Questionnaire-Bangladesh
CTT	Classical test theory
DNA	Deoxyribonucleic acid
EEG	Electroencephalogram
EFA	Exploratory factor analysis
fNIRS	Functional Near-Infrared Spectroscopy
HOME Inventory	Home Observation Measurement of the Environment Inventory
HPA	Hypothalamic-pituitary-adrenal
ICC	Intraclass correlation coefficient
ICDDR,B	International Centre for Diarrhoeal Disease Research, Bangladesh
IQ	Intelligence Quotient
IRT	Item response theory
MRI	Magnetic Resonance Imaging
MSPSS	Multidimensional Scale of Perceived Social Support
N	Number
PFC	Prefrontal cortex
PROVIDE Study	Performance of Rotavirus & Oral Polio Vaccines in Developing Countries Study
SES	Socioeconomic status
U.S.	United States
UN	United Nations
WHO	World Health Organization
WPPSI	Wechsler Preschool and Primary Scale of Intelligence

Abstract

Background: Research examining effects of psychosocial adversities on child neurocognitive development has occurred largely in high-income countries. A key barrier to global extension of this work has been a lack of culturally appropriate measurement tools. This thesis reports on the development, validation, and initial application of the Childhood Psychosocial Adversity Questionnaire–Bangladesh (CPAQ-B), the first context-tailored tool to assess early childhood psychosocial adversity comprehensively as a neurodevelopmental risk factor in a low-resource setting.

Participants: 355 mother-child dyads from an urban slum of Dhaka, Bangladesh took part in qualitative work (N=45), item pretesting (N=25), questionnaire piloting (N=53), and administration of the CPAQ-B (N=232) at child ages 18, 24, 48, and/or 60 months. Community-based clinicians (N=35) also participated in qualitative sessions.

Methods: A conceptual model of early psychosocial adversity was developed to inform item selection. Initial items underwent expert review, pretesting, piloting, and item reduction. The CPAQ-B was then administered for psychometric analysis and validation including assessment of internal consistency, test-retest and inter-rater reliability, and criterion and predictive validity via correlation with parallel measures and child IQ scores. Finally, as an initial use of the tool, multivariate ordinary least squares regression modeling was used to assess whether 48-month CPAQ-B scores predict 60-month child intelligent quotient (IQ) scores when controlling for concurrent risks.

Results: 180 initial items were generated and piloted, and reduced to 64 items in the final scale. Exploratory factor analysis supported 11 subscales assessing for child maltreatment, various maternal and family stressors, and neighborhood-level exposures. Evidence supported strong subscale internal consistency and full-scale test-retest and inter-rater reliability. Subscale and full-scale scores correlated significantly with comparator measures. 48-month CPAQ-B scores significantly predicted 60-month WPPSI-IV IQ in multivariate regression models controlling for socioeconomic variables and malnutrition.

Conclusions: The CPAQ-B represents a novel research tool measuring childhood psychosocial adversities, with good initial validity evidence for use among low-SES children ages 18-60 months in Dhaka, Bangladesh. Future work may adapt the instrument for use in other countries in the region, while methods may inform similar efforts to develop context-tailored assessments across other socio-cultural settings.

I. Background

Lack of contextually appropriate measurement tools assessing psychosocial exposures may promote systematic underestimation of the importance of such variables in global health research. In high-income countries, tools assessing exposure to childhood abuse, neglect, family violence, caregiver psychopathology, and other "adverse childhood experiences" have generated evidence on links between early adversity and poorer health and development outcomes across the life course.[1-10] Considering neurodevelopment, specifically, studies link early psychosocial stress to poorer outcomes across domains including stress regulation, executive functioning, social cognition, language, and overall cognitive performance.[11-13] This evidence base is now enabling "biomarker-driven" research to inform improved early childhood development (ECD) interventions based on understanding of specific physiological pathways disrupted by excessive early psychosocial stress.[14,15]

A similar evidence base on the early developmental effects of childhood psychosocial adversities has not emerged in low-resource country settings, in part due to lack of tools to assess childhood psychosocial stress comprehensively in many settings. Yet children across global settings are thought bear substantial burdens of psychosocial risk factors found to impact child development in other settings. A 2016 systematic review estimated that at least 44% of children in developed countries and 59% in developing countries had been victims of physical, emotional, or sexual violence (excluding corporal punishment) in the preceding year.[16] The most recent United Nations (UN) World Report on Violence against Children, published in 2006, estimates that 133–275 million children annually regularly witness violence between primary caregivers, and at least 150 million girls and 73 million boys are victims of forced sexual activity each year.[17] Caregiver poor mental health is globally prevalent, with depression representing the

leading cause of disease-related disability globally per World Health Organization (WHO) estimates.[18] Despite growing focus on ECD in the global child health arena,[19,20] limited evidence on the nature and scale of developmental impacts linked to psychosocial risk factors hinders progress on interventions to improve child outcomes.[21]

This thesis reports on development and validation of the first context-tailored tool supporting comprehensive, prospective assessment of childhood psychosocial adversities in a non-Western country context, the Child Psychosocial Adversity Questionnaire–Bangladesh (CPAQ-B). I describe CPAQ-B and provide evidence suggesting that it is a valid tool to assess psychosocial risk factors for poorer neurodevelopmental outcomes among low-socioeconomic status (SES) children ages 18-60 months in Bangladesh. As a preliminary example of its usefulness, the instrument is then applied to generate evidence that psychosocial risk factors predict future child intelligent quotient (IQ) when controlling for other covariates.

II. Methods

Study

The CPAQ-B was developed within the Bangladesh Early Adversity Neuroimaging (BEAN) Study, a longitudinal investigation using neuroimaging and behavioral measures to characterize early child neurodevelopmental risk factors and outcomes in Dhaka, Bangladesh.[22] The study focuses on low-SES children in the urban slum of Mirpur who were randomly recruited from two preexisting longitudinal studies of mother-infant dyads, the Performance of Rotavirus and Oral Polio Vaccines in Developing Countries (PROVIDE) Study,[23] and the Burden of Cryptosporidiosis (Crypto) Study.[24] The BEAN Study assesses brain structure and function using electroencephalography (EEG), functional near-infrared spectroscopy (fNIRS), and structural and functional magnetic resonance imaging (MRI). It represents a collaboration

between Boston Children's Hospital, University College London, University of Virginia, and the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B) with funding from the Bill and Melinda Gates Foundation.

Sample

Original recruitment for the PROVIDE and Crypto Studies involved door-to-door visiting of all pregnant women in Mirpur. Parents have had regular home visits beginning either in pregnancy month 7 or at birth, and have visited the study over time for study activities and free child health services. Participants in the first phase of CPAQ-B development (qualitative sessions) were identified by stratified random sampling, first separating PROVIDE/Crypto by income tertile and ethnic group and then selecting randomly within groups. ICDDR,B study field workers taking part in staff focus groups were purposively selected based on depth of experience working with the study community. During phases of CPAQ-B piloting and validation, participants were recruited by inviting all subjects at routine BEAN, PROVIDE, and/or Crypto Study visits conducted when enrolled children were 18, 24, 48, and/or 60 months old. Based on PROVIDE/Crypto recruitment in pregnancy, all parents in the BEAN Study and CPAQ-B development are biological mothers. Adjunctive data informing CPAQ-B item reduction was drawn from a middle-SES sample used as a control group in the BEAN study, which was recruited by random door-to-door invitation in higher-SES wards adjacent to Mirpur.

Ethical approval

All activities were conducted in secure ICDDR,B study clinic rooms following informed consent with ethical approval from Boston Children's Hospital and ICDDR,B.

Instrument design & validation

CPAQ-B development proceeded through three stages, described below. Table 1 specifies dimensions of instrument validity assessed, *a priori* hypotheses, and evidence sources.

Table 1. Dimensions of validity assessed with hypotheses & data sources

Validation proposition: Use of the CPAQ-B as a research measure assessing early psychosocial adversity as a child developmental risk factor among low-SES, urban Bangladeshi children ages 18-60 months

Validity dimension	*A priori* hypotheses	Source of data
Construct validity Validity of the proposed use	a) Cumulative early psychosocial stress, including child and caregiver experiences, shapes developmental outcomes. b) The scientific construct of psychosocial adversity has locally specific and embedded manifestations in Mirpur.	- Literature review - Qualitative data - Expert review of conceptual model
Content validity Extent to which content captures construct	a) Item content captures major themes in the conceptual model without extraneous content. b) Factor analysis will support a final subscale structure that corresponds to the conceptual model.	- Comparison to conceptual model - Expert review - Pretesting with parents for comprehension & relevance - Factor analysis of CPAQ-B data
Internal consistency Content cohesion	After item reduction, most or all subscales will show good internal consistency with Cronbach's $\alpha \geq 0.70$; No subscale will have Cronbach's $\alpha < 0.60$.	- Classical test theory analysis of CPAQ-B data
Test-retest & inter-rater reliability Stability of scores over time/raters	a) Subscale and full-scale scores will show reasonably test-retest reliability over 2 weeks (accounting for potential true change), with adequate average ICC≥0.7. b) Inter-rater reliability will be lower based on necessary layering of variance due to rater and occasion (explained below), with adequate average ICC≥0.6.	- Retest administrations after 2-week interval with same interviewer (test-retest reliability) or different interviewer (inter-rater reliability minimum estimate)
Convergent validity Correlation with similar measures	Subscales scores will correlate with similar instruments, though likely only moderately given that comparator measures generally measure non-identical constructs and have not been validated in the study context.	- CPAQ-B and "comparator instruments" as described
Predictive validity Association with outcomes	Full-scale and subscale scores significantly predict future child cognitive performance, both in bivariate analyses and when controlling for other risks.	- 48-month CPAQ-B and 60-month WPPSI-IV scores
Incremental validity Extent of novel value	a) The CPAQ-B will explore new domains of psychosocial risk not captured by other instruments in Bangladesh while taking less time to administer than multiple instruments used previously. b) It will show better internal consistency and predictive validity than other related measures.	- Relative internal consistencies - Administration times - CPAQ-B subscale content vs. content of other available measures

CPAQ-B=Global Child Adversity Questionnaire-Bangladesh; SES–Socioeconomic status; CTT=Classical test theory; ICC=Intraclass correlation coefficient; WPPSI-IV=Wechsler Preschool and Primary Scale of Intelligence, 4th Ed.

Phase 1: Conceptual model generation

A conceptual model was first developed to define dimensions of child and caregiver psychosocial stress relevant to developmental outcomes and salient in the study context. The conceptual model integrated scientific literature, qualitative work, and expert input.[25,26]

Ten semi-structured, one-hour interviews and four focus groups (2 hours long, 8-10 participants each) were conducted with mothers from Mirpur. Four focus groups of the same length and size were conducted with ICDDR,B community-based study staff. Sessions were conducted in Bangla using interview guides with prompts probing local norms and practices relating to caregiving generally and to child/caregiver experiences of stress. Sessions were audio-recorded, transcribed, translated, and analyzed using NVivo 11 analytic software (QSR International Pty Ltd, 2015). A qualitative content analysis integrated inductive generation of codes with deductive application of themes drawn from literature on psychosocial stress and child development. [27,28] Codes were generated by the lead author, and initial transcripts were double coded by the lead author and a Bangladeshi undergraduate research assistant at Boston Children's Hospital. Codes were refined until good inter-rater agreement was achieved ($\kappa>0.75$). Subsequently, roughly 15% of the transcripts were double coded and compared for discussion between the two coders. A conceptual model of early childhood adversity integrated the analysis with literature review.

Phase 2: Item generation, pretesting & reduction

Initial CPAQ-B items were generated to reflect conceptual model content, informed by qualitative work and by existing similar instruments.[25,29] Bangladeshi and U.S. academic experts and Bangladeshi study staff reviewed items for completeness and relevance. The final instrument draft was translated to Bengali and back-translated to English.

Initial items were pretested with mothers via cognitive interviewing, a think-aloud method in which participants respond to questionnaire items while articulating reasoning processes.[30-32] Pretesting led to further refinement of questions. Staff were trained in uniform administration methods including assessment of rating reliability.

After pretesting, items were piloted with study participants from Mirpur at child age 48 months (the only low-SES age cohort for which study visits coincided with piloting), and preliminary assessment of subscale coherence and item performance was performed within a classical test theory (CTT) framework assessing internal consistency (Cronbach's α), item-rest correlations, and distributions.[33] Items with inadequate distributions (e.g., mostly at ceiling or floor) were eliminated. Subsequently, items with low item-rest correlations were removed iteratively, targeting Cronbach's $\alpha > 0.70$ for each hypothesized subscale. Items could be retained despite low item-rest correlations to maintain theoretical content.[34]

To provide supplementary information, data were also pooled with BEAN middle-SES controls to achieve adequate sample size for Item Response Theory (IRT) modeling.[35] IRT estimates of item discrimination and information were used as adjunctive data to reflect further on CTT-based decisions.

Phase 3: Psychometric analysis

After item reduction, the CPAQ-B was administered to assess various dimensions of validity. All 48-month-olds received the longer, pilot CPAQ-B version, and data only from retained items were included for these participants in validation analyses to avoid losing representation of that age group in validation. Sensitivity analyses assessing impacts of this decision are discussed below while describing specific dimensions of validity assessed.

A. Content validity

The subscale structure of the questionnaire was assessed by exploratory factor analysis (EFA). Separate analyses were completed for child-focused questions (e.g., on caregiving, abuse, or neglect), for caregiver- or family-focused items (e.g. maternal depression, family conflict), and for community-focused items (e.g., crime, violence). Within logistical constraints, a minimum 5:1 item-to-subject ratio was targeted;[36] implications of this relatively low ratio are discussed. Promax rotation was utilized in final models if significant correlations were seen between subscales.

Final CPAQ-B scores were scaled to a 0-10 range, with contributions of subscales weighted subjectively based on prior literature to reflect evidence on the relative importance of various psychosocial exposures to child development (considered preferable to *de facto* subscale weighting by relative length). Assigned weights reflected literature indicating that severely stressful exposures in the immediate caregiving environment are particularly important in infancy and early childhood, as effects of more distal (e.g., community-level) exposures are mediated substantially by effects on caregivers.[37] Severe exposures in a child's immediate caregiving environment (e.g., neglect, abuse, domestic violence, maternal depression) most clearly linked to negative developmental effects were given a weight of 1.5, and more normative adverse exposures in the family environment or stressors in the community were weighted at 0.5. Implications of this subjective weighting process and alternative approaches are discussed.

B. Internal consistency

Subscale internal consistency and item properties were assessed by calculation of Cronbach's α, item means and standard deviations (SDs), and item-rest correlations within subscales.[33] Statistics were generated separately for each child age group (18, 24, 48, or 60

months) for the child-focused subscales (for instance relating to abuse and neglect) based on potential differential item functioning by age. In addition, internal consistency was calculated with and without the 48-month group as a sensitivity check to ensure that item selection based on contributions to internal consistency in this group at the pilot phase did not artificially inflate estimates.

C. Test-retest & inter-rater reliability

To assess test-retest reliability, a subset of participants completed the measure twice with the same interviewer after a two-week interval. Average intra-class correlation coefficients (ICCs) were calculated across administrations for subscale/full-scale scores. The two-week gap was selected to decrease bias from participants remembering items, but was expected to introduce some acceptable variability based on true changes in underlying constructs. Accounting for some expected true variation in latent traits over time, acceptable full-scale test-retest reliability was operationalized as ICC $\geq$0.85. Other participants completed the questionnaire twice over two weeks with different interviewers to assess inter-rater reliability. As rapport established in a one-on-one setting was expected to be the most substantial source of inter-interviewer variability, it was necessary to have different interviewers complete their assessments of each participant on separate occasions (versus, for instance, dual-rating of a single video-taped session). Average ICCs, therefore, were expected to underestimate inter-rater reliability, layering variance due to rater and occasion. Acceptable reliability for inter-rater assessments (as layered in a test-retest design) was operationalized as a lower average ICC $\geq$0.75 for full-scale scores.

D. Convergent validity

Pearson correlation coefficients were calculated comparing CPAQ-B subscale scores to scores on related instruments (described below). If no instrument was available with good content overlap, a measure was selected, where possible, that assessed a construct expected to correlate with the CPAQ-B subscale based on prior literature. Correlations in a direction consistent with predictions in the literature, while not evidence of true convergent validity, was considered at least to support the hypothesis that a CPAQ-B subscale measures what it intends.

E. Predictive validity

Correlation was further assessed between 48-month CPAQ-B subscale/full-scale scores and 60-month IQ as measured by the Wechsler Preschool and Primary Scale of Intelligence, 4th Edition (WPPSI-IV, described below). A negative relationship between all dimensions of psychosocial adversity and child cognition was expected. At this stage, simple correlation in the expected direction was taken as evidence that the CPAQ-B may be measuring the constructs it intends to measure. A one-year gap between CPAQ-B and WPPSI-IV administrations supports claims of a predictive relationship, and reflects a theoretical expectation that point-in-time IQ depends upon developmental processes over preceding years.

F. Incremental validity

Administration times and comprehensiveness of content was compared for the CPAQ-B and instruments previously used in Bangladesh to assess dimensions of psychosocial adversity. In addition, internal consistency (Cronbach's α) and WPPSI-IV correlations for comparator measures administered to this sample were compared to those for roughly parallel CPAQ-B subscales. Cronbach's α was standardized to a length of 10 items using the Spearman Brown Prophecy Formula given the tendency for Cronbach's α to increase with scale length.

Other measures

Psychosocial measures

At 24- and 60-month visits, participants completed the Bangla translation of the Edinburgh Postnatal Depression Scale (EPDS), a widely-used, 10-item measure created in Scotland in the 1980s to screen for postpartum depression based on past-week symptoms;[38] it it is now commonly used in both postpartum and non-postpartum samples.[39] Some validity data has been generated for the Bangla translation's use as a postpartum clinical screen in Bangladesh.[40] No published validation data from Bangladesh considers use in non-postpartum women or as a research tool assessing symptoms over a range of severities.

24-month-old BEAN participants were assessed with the Home Observation Measurement of the Environment (HOME) Inventory, a home visit-based observational measure estimating levels of cognitive stimulation and emotional responsiveness in the home. A version of this instrument has been adapted for Bangladesh, with some published validity evidence.[41] The BEAN study used the original version of the inventory, which differs somewhat from the Bangladesh adaptation, and for which no validation data on use in Bangladesh exists.

The Multidimensional Scale of Perceived Social Support (MSPSS) was administered to assess maternal social support at 24- and 60-month visits. This 12-item scale, developed among U.S. undergraduates in the 1980s, assesses perceptions of emotional, psychological, and logistical support.[42] No published validation data exists for use in Bangladesh.

For CPAQ-B items on intimate partner violence, it is noted that the WHO Multi-Country Study on Women's Health and Domestic Violence used its questionnaire in Bangladesh.[43] Given substantial participant burden involved in adding another measure of sensitive exposures, and the

lack of comprehensive published validation data for the WHO measure in Bangladesh, the decision was made to draw upon the WHO instrument in developing CPAQ-B items but not to co-administer it. Finally, SES indices were generated based upon household per capita income and parent education and occupation collected at child age 24 months in the PROVIDE study.

Developmental assessment

At 48 and 60 months, BEAN and PROVIDE participants completed the WPPSI-IV,[44] a cognitive performance measure validated for use in children roughly 30-84 months old in the U.S. and various developed-country settings. It assesses an array of cognitive domains including receptive and expressive vocabulary, picture memory, matrix reasoning, identification of differences and similarities. CPAQ-B scores were compared to WPPSI-IV raw scores versus familiar scaled t-scores (mean of 100, SD of 15), as substantial score distribution differences between our sample and the normative U.S. population used for scaling otherwise could bias analyses.

Preliminary application of the CPAQ-B

After validity assessments, an initial use case for the CPAQ-B as a measure of neurodevelopmental risk was tested. Multivariate ordinary least squares (OLS) linear regression models were fit, first predicting 60-month WPPSI-IV raw IQ score from 48-month CPAQ-B score alone (Model 1), then controlling for effects of stunted growth at 48 months, based on potential for confounding via shared association with low SES (Model 2), and finally, controlling for an index of socioeconomic status comprised of a weighted composite of maternal education, paternal education, paternal occupation, and log household income (Model 3). Stunting was selected as a marker of malnutrition shown in prior studies to correlate with both child cognition and poverty given potential for such non-psychosocial risk factors to bias estimates of IQ

associations with psychosocial risks (see review[21]). An additional model (Model 4) was fit predicting child IQ based on adversity score quartile ranks as a categorical variable to better visualize effect sizes at various exposure levels while controlling for all covariates included in Model 3. Models tested for hypothesized linear relationships between included predictors and 60-month raw IQ scores against a null hypothesis of no such relationship. OLS assumptions of linearity, residual normality and heteroscedasticity were assessed.

Follow-up with at-risk families

Ethical issues arose as participants shared experiences relating to child maltreatment, domestic violence, mental health needs, and other forms of psychosocial adversity. A follow-up plan was developed in collaboration with Bangladeshi Principle Investigators familiar with the study context and family violence research. The approach balances concerns relating to local cultural dynamics, relatively limited formal child protection services, and the aim to avoid unintended harm. First, items were defined for which affirmative response warranted automatic follow-up. Participants endorsing these items flagged as indicators of high risk were to be offered: (a) additional assessment by a clinical psychologist with expertise in family violence (optional except in cases of concern for child maltreatment); (b) referrals for legal and NGO services and clinical counseling; and (c) free counseling at the ICDDR,B clinic from a clinical psychologist. In extreme cases police involvement could be initiated for child protection (this has not occurred to date).

III. Results

Conceptual model

Based on literature review, *early childhood adversity* was conceptualized as a negative experience in early life observed in prior literature to increase risk of poorer development or

health outcomes, while *psychosocial* adversity is defined specifically as adversity related to a child's social environment (e.g., involving relationships with caregivers, family, and others) in interaction with psychological processes.[45] A "cumulative risk" approach, assessing sources of psychosocial stress from all sources, was selected. This approach is supported empirically by evidence of dose-response relationships between cumulative early adversity and later outcomes across many studies,[4,5,8] and mechanistically by the "allostatic load" paradigm identifying pathogenic effects of high all-cause stress.[46] Child exposures including abuse,[47] psychosocial deprivation or neglect,[12,48] and hostile or unresponsive caregiving[49,50] had been linked to adverse neurodevelopmental effects. Evidence also suggests that caregiver stress substantially influences stress levels in infants and young children, such that substantial caregiver stressors are generally included in measures of early life psychosocial adversities.[37] A model of how psychosocial stress can shape developmental and health outcomes is shown in the Appendix (Figure A2).

Although relevant, psychosocial protective factors are considered here to be distinct from adversities and are not included in the CPAQ-B. Protective factors may include some exposures that exist on a spectrum with adverse exposures (for instance, unusually responsive caregiving may represent a protective factor, where extreme caregiver unresponsiveness may be conceptualized as an adverse exposure); in these cases, it is noteworthy that psychometric analyses focused on identifying items that discriminate at the adverse end of the exposure spectrum. Meanwhile, other protective factors, such as access to early childhood enrichment activities, are not conceptualized here along a spectrum whereby their absence constitutes adversity. Capturing these exposures would likely require development of an additional measure. In full, while important for understanding a child's psychosocial environment holistically, protective factors are not considered given the already ambitious scope of the CPAQ-B.

Qualitative work subsequently served to define specific manifestations of child and caregiver stress in the study context. Ninety-four participants took part in qualitative sessions, including 62 mothers (10 in-depth interviews, 4 focus groups) and 32 staff (4 focus groups). Table 2 shows nine domains of psychosocial adversity identified in these sessions, along with selected observations and quotes. It was noted that participants readily shared their own use of corporal punishment and experiences of domestic violence, including in group settings, supporting a hypothesis that it may be possible, from a cultural and interpersonal perspective, to ask about sensitive topics and expect relatively open discussion.

Table 2. Conceptual model of adversity domains with selected qualitative findings

Domains of adversity	Examples of themes emerging in qualitative work	Quote illustrating one or more key themes
Poor caregiving quality & child neglect	Many hesitated to identify "poor" caregiving given economic constraints faced by parents. Failure to meet basic needs (cleanliness, safety, supervision) was criticized, but in a nuanced manner.	*"As an example [of poor caregiving], there is a mother in our house who has to go to work, leaving her 3-month-old baby with her sister-in-law. So, the baby does not get enough care. For instance, the baby needs milk but doesn't get it…She needs her mother's affections and lap but does not get that."* -Parent
Child discipline/abuse	Corporal punishment was described as widespread and generally acceptable if mild. Using slang in front of children was considered wrong based on religious dictums against cursing, as was more severe physical punishment causing injury.	R:*" I saw a mother throw a chair she had been sitting on at her child. By this, the child can get hurt...If my child does wrong, I can lock her in a room, can beat her slightly with a stick…This type of discipline is okay."* [I:*"But that which injures a child is not ok?"*] R:*"That is not ok. My child can be hurt very badly because of me, no?"* -Parent
Marital & family conflict	Marital disputes represented a significant source of maternal stress. Criticism of wife by in-laws and extended family was seen as a common source of distress and conflict.	*"Often, a mother-in-law treats her daughter-in-law very badly…asking her to do this and that, every household chore, but still she's not satisfied…[so] the husband starts beating his wife. He accuses her by saying, 'why did you talk badly with my mother?'"* -Parent
Domestic violence	Verbal and physical abuse of mothers was described as very common, sometimes causing serious injuries.	*"Husbands beat their wives with whatever they have in front of them. My child's father beats me in that way…If he gets a floor cleaner or stick, he beats me with that."* -Parent
Maternal social support	A number of participants described feeling isolated, devalued, and unseen.	*"[In families] it's like we do not have any value. No one listens to us and our words are not valued too."* – Clinician, member of community
Drugs, alcohol, & gambling	Drug, alcohol, and gambling addictions were substantial problems in some neighborhoods and rare in others.	*"Crime is happening because of drug trading. We have seen that there is a home in our slum, and from there fencidil and marijuana are being sold…maybe someone doesn't give the correct price, then a brawl starts."*-Clinician
Household economic stressors	Participants described household poverty as a major stressor and determinant of the quality of caregiving that parents are able to provide to children.	"*While parents are in the house there is no electricity. So they are busy arranging straw for fuel and cooking. So, we usually found them too busy [to play with children]…Those who are able to ensure home, food, clothes, education and medical treatment, are the good parents."*-Clinician
Maternal stress & depression	Participants readily launched into discussion of mental health problems in the community, though they did not generally use clinical terms or concepts.	*"At times I feel very sad, very low; I cannot sleep at night, I cannot eat...I don't even like my husband. I don't even like my children, I feel like going away, I feel very bad, my body becomes weak…"* [I: *"So what do you call this?"*] *"To us, it is called sorrow. It is our sadness, misery."*-Parent
Community adversities	Participants described stressors in the community, including theft, gang activity, drug trading, unsafe roads, political violence, and more rarely, cases of physical and sexual assault.	*"What will happen to him when he goes out on the street? Will he be hit by a car? Will he get into trouble? Will he be mugged by someone? Has anyone snatched his valuables, mobile phone, money? Where has he been? Where is he? Did he take food? Will he come home?"* -Parent
Illnesses/deaths/other traumas	Participants felt children were impacted by acute traumas like witnessing violence during political rallies, witnessing fires, or experiencing the death of relatives	*"[During political strikes,] if anyone does something against the political parties, they beat them up... They might slash someone's hand, someone's leg, thrash someone's head with bamboo stick...The kids get terrified, and they don't want to go to the school the next day."*-Parent

Item generation

180 initial items were generated from the conceptual model. Existing instruments consulted appear in Table A1 (Appendix).

Pretesting & piloting

Items were pretested with 25 mothers. Particular adjustment was made to the specific anchoring and translation of the 5-point Likert scale by which participants gauged exposure frequency (most items using *Never, Rarely, Sometimes, Often,* and *Always*) to improve intelligibility given limited familiarity with similar tests in the sample. A single item was removed that asked about exposure to sexual violence within marriage as most women found the item confusing and sometimes offensive. Other adjustments to items were non-substantive consisting of adjusting wording to improve comprehension. After adjustment, all 180 items were adequately comprehensible and acceptable, and were piloted with 53 low-SES mothers at child age 48 months. Use of item-rest correlations and IRT item information functions to reduce items are illustrated in the Appendix (Figures A3 & A4). IRT models included data from 210 middle-SES controls at child ages 6 and/or 36 months (with only 36-month-olds included in child-focused subscales due to inappropriateness of items to infants). In full, 87 items were removed, resulting in a 93-item version of the CPAQ-B administered to a larger sample for analysis of psychometric properties and other dimensions of validity (details below).

Psychometric analysis of the CPAQ-B

The sample for psychometric analysis of the shortened CPAQ-B included 285 participants from Mirpur (N=71 18-months-old, N=106 24-months-old, N=53 48-months-old, N=55 60-months-old). Of these, 53 (all 48-month-olds) had completed the longer, pilot version, as noted. Demographic and socioeconomic data was available for participants from PROVIDE

(N=108); of note, BEAN and Crypto Study participants are considered representative samples from the same underlying Mirpur population. Among PROVIDE participants, mothers had a mean of 4.3 years of formal education, with 33% of the sample having had no formal schooling. At child age 2 years, 95% lived below the Bangladeshi national poverty line (U.S.$2/day), with a mean of $1.10 per household member daily. Most mothers (89%) identified as housewives, while 7% worked making handicrafts in-home and 4% worked outside the home.

Factor analysis & subscale structure

Initial exploratory factor analysis (EFA) on child-focused items indicated 11 factors. Five factors with eigenvalues greater than one were retained in a subsequent EFA (constrained to 5 factors). Seven items were dropped at this stage due to inadequate variance in the final sample (nearly all at floor) or low factor loadings. A final promax-rotated model (Table A2, Appendix) had four factors with eigenvalues 1.84-4.07 as defined by 21 indicators (subject-to-item ratio 13.57:1) with factor loadings 0.44-0.91 with no significant cross-loading onto other factors (defined as factor loading >0.30). Factors were identified as *Harsh discipline & abuse*, *Lack of supervision*, *Material neglect*, and *Low caregiver warmth*. On theoretical grounds, *Lack of supervision* and *Material neglect* were considered sub-constructs within a *Neglect* subscale.

Initial EFA on caregiver-focused items indicated 21 factors, of which 6 had eigenvalues greater than 1. An additional EFA was run constrained to 6 factors, and twelve items were dropped due to inadequate variance, low factor loadings, or loading on multiple factors. A final 6-factor constrained EFA resulted in a model with factors having eigenvalues 7.05-13.35. The 37 indicators (subject-to-item ratio 7.70:1) had factor loadings 0.40-0.97 and no significant cross-loading (Table A3, Appendix). Factors were identified as *Maternal depression*, *Domestic violence*, *Maternal social isolation*, *Family Conflict, Economic stress*, and *Marital breakdown.*

Marital breakdown was distinguished from the *Family Conflict* subscale in that the former included items about spousal separations, divorce, abandonment, and infidelity, sometimes implying a loss of regular contact, whereas the latter referred to day-to-day conflicts between family members, including spouses. Finally, EFA on items about the community environment showed a one-factor structure (second factor with eigenvalue 0.25). The single factor, identified as *Community stressors*, had eigenvalue 2.07 with 5 indicators and factor loadings 0.46-0.68 (Table A4, Appendix).

Internal consistency

Internal consistency (N=285) was adequate for all subscales (Cronbach's $\alpha > 0.65$), and was >0.70 for all but *Neglect* (Table 3). Slightly lower internal consistency (a measure of unidimensionality) for *Neglect* is consistent with EFA results suggesting two distinct item clusters. Cronbach's α was adequate (>0.60) across child-focused subscales and for all age groups when calculated separately, with the exception of the *Neglect* subscale in the 18-month age group (Cronbach's $\alpha=0.55$); implications are discussed below.

Estimating internal consistency while excluding 48-month-old participants for sensitivity check showed that all were within 0.03 Cronbach's α reliability units of original estimates with no impact on judgments about the extent or adequacy of consistency.

Retest reliability

Full-scale CPAQ-B scores had good test-retest reliability (average ICC=0.88, 95% confidence interval 0.77-0.94) in 39 retests (N=10 at 18, 24, and 60 months, and N=9 at 48 months). Minimum inter-rater reliability (conflated with test-retest reliability, as discussed), was also adequate (average ICC=0.79, 95% confidence interval 0.60-0.89) in 40 retests (N=10 at 18,

24, 48, and 60 months). Test-retest and inter-rater ICCs showed wide confidence intervals at the subscale level (Table A5, Appendix).

Convergent validity

Two CPAQ-B subscales had good construct overlap with comparator instruments: *Maternal depression* with the Edinburgh Postnatal Depression Scale (EPDS), and *Social isolation* with the Multidimensional Scale of Perceived Social Support (MSPSS) (with inverse score directionality). Symptom recall periods differed substantially between the CPAQ-B (past 6 months) and EPDS (past 7 days). CPAQ-B subscales related to child maltreatment and care overlapped partially (again, with inverse directionality) with HOME Inventory subscales, although the HOME Inventory did not attempt to assess for maltreatment or caregiver warmth comprehensively. For other CPAQ-B subscales, non-parallel but potentially correlating measures were identified. *Family conflict* and *Marital breakdown* were both expected to correlate negatively with the MSPSS when administered at the same age point, while *Domestic violence* was expected to correlate positively with concurrent EPDS scores based on prior literature in Bangladesh.[51] Economic stress at 48 months was expected to correlate positively with 24-month SES (concurrent data unavailable, as 24-month CPAQ-B sample was not from PROVIDE). No comparator instruments were available for the *Community stressors* or *Drugs, alcohol & gambling* subscales. Correlations with all comparator instruments were weak-to-moderate in magnitude and statistically significant apart from the *Low caregiver warmth* subscale and the HOME Inventory Avoidance of Punishment & Restriction subscale (N=106, r=-0.11, p=0.27) (Table 3).

Predictive validity

Full-scale CPAQ-B score at age 48-month correlated negatively with 60-month WPPSI-IV raw IQ ($r(51)=-0.61$, $N=52$, $p<0.001$). Correlations with 60-month IQ were also negative and significant for all 48-month CPAQ-B subscales apart from *Marital breakdown* ($r(51)=-0.25$, $p=0.08$), for which the analysis may have been underpowered (Table 3).

Table 3. Classical test theory statistics, convergent validity & predictive validity

		Classical Test Theory Statistics (N=285)			Correlation with Comparator Instruments			Correlation with WPPSI IQ[b] (N=52)
Subscale	**Items**	**Cronbach's α**	**Item-rest correlations**	**Mean (SD) item score, 0-4 scale[a]**	**Comparator instrument**	**N**	**Correlation coef. (p)**	**Correlation coef. (p)**
Child maltreatment								
Harsh discipline & abuse	9	0.85	0.27-0.83	0.80 (0.75)	HOME-A[c]	106	-0.33***	-0.28*
Neglect	6	0.67	0.30-0.65	0.53 (0.61)	HOME-O[c]	106	-0.33***	-0.38**
Low caregiver warmth	5	0.88	0.54-0.83	1.81 (1.08)	HOME-A[c]	106	-0.11	-0.43**
Maternal & family stress								
Depression	11	0.96	0.74-0.91	1.01 (0.91)	EPDS	149	0.42***	-0.42**
Physical domestic violence	9	0.91	0.64-0.84	0.26 (0.51)	EPDS	149	0.31***	-0.47***
Family conflict	5	0.81	0.56-0.65	1.27 (0.81)	EPDS	149	0.31***	-0.38**
Economic stressors	4	0.86	0.62-0.79	0.77 (0.87)	SES[d]	52	-0.34**	-0.42**
Social isolation	5	0.93	0.68-0.87	0.94 (1.07)	MSPSS	149	-0.26**	-0.30*
Marital breakdown	4	0.80	0.60-0.69	0.30 (0.64)	EPDS	149	-0.42***	-0.25
Environmental adversities								
Community stressors	6	0.77	0.45-0.60	0.50 (0.70)	SES[c]	52	-0.31*	-0.37**
Full scale	**64**			0.67 (0.50)				**-0.61*****

*p<0.05 **p<0.01 ***p<0.001

[b]CPAQ-B administered at child age 48 mo. vs. WPPSI-IV measured at child age 60 mo. (N=52) [c]Child age 24 mo. [d] 24-mo. SES vs. 48-mo. CPAQ-B subscale scores (concurrent data unavailable); HOME-A=HOME Inventory Avoidance of Punishment & Restriction subscale; HOME-O=HOME Inventory Organization of the Environment subscale; EPDS =Edinburgh Postnatal Depression Scale; MSPSS=Multidimensional Scale for Perceived Social Support

Incremental validity

Total administration time for the CPAQ-B is roughly 30 minutes. By comparison, total administration time for overlapping measures (EPDS, MSPSS, and HOME Inventory) was roughly 1 hour (additional administration time needed for existing WHO Multi-Country Domestic Violence Survey is not estimated here). The CPAQ-B also captures six additional dimensions of psychosocial stress not covered in the prior instruments.

Considering the two instruments most parallel to CPAQ-B subscales, adjusted Cronbach's α was 0.81 for the EPDS versus 0.96 for the CPAQ-B *Maternal depression* subscale, and 0.84 for the MSPSS versus 0.98 for the CPAQ-B *Social isolation* subscale. Unfortunately, 48-month data was not available for either the EPDS or the MSPSS to assess correlations with 60-month IQ. To generate some estimate of predictive utility, 36-month comparator instrument scores were compared to 48-month IQ, and correlations were near zero for both EPDS ($r(122)=0.004$, $p=0.97$) and the MSPSS ($r(120)=0.07$, $p=0.42$). When measured concurrently in 78 participants, 60-month CPAQ-B depression subscale score correlated significantly with 60-month raw IQ ($r(77)=-0.24$, $p=0.03$) while 60-month EPDS score did not ($r(77)=-0.12$, $p=0.28$). Neither the 60-month CPAQ-B social isolation score ($r(77)=-0.18$, $p=0.11$) or 60-month MPSS score ($r(77)=0.19$, $p=0.09$) correlated significantly with concurrent 60-month raw IQ.

Application of the CPAQ-B as a measure of developmental risk

Multivariate regression models predicted 60-month WPPSI-IV IQ in 52 participants from 48-month CPAQ-B scaled score alone (Model 1), then controlling for child stunting (Model 2), and finally controlling for stunting with a composite index of socioeconomic status (Model 3). The sample had mean 60-month raw IQ of 53.08 (SD 9.29, range 34-74), scaled CPAQ-B score of 2.26 (SD 1.23, range 0.00-5.38), SES composite score of 21.26 (SD 12.66, range 2.38-55.50),

and stunting prevalence of 23.08%. For one participant, a missing 60-month WPPSI-IV score was imputed to have 60-month percentile rank equivalent to that of the participant's 48-month WPPSI-IV score percentile rank. One observation missing SES data was deleted list-wise. Robust standard errors were used given graphical evidence of residual heteroscedasticity over stunting.

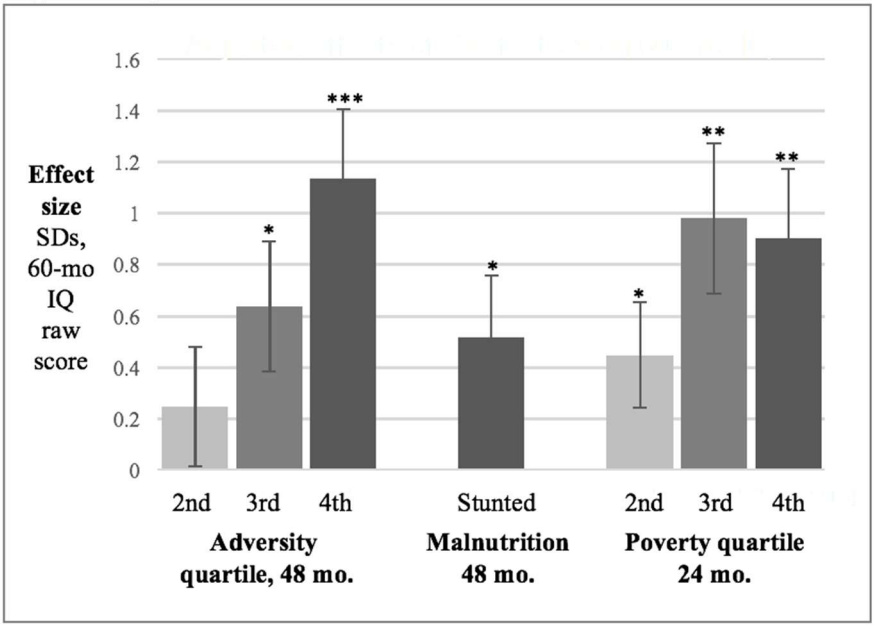
48-month psychosocial adversity was a statistically significant predictor of 60-month IQ in all models. Controlling for both SES and child stunting, a 1-point increment on the 0-10 adversity scale predicted a 3.88 point decrement in 60-month IQ ($t(48)=-4.89$, $p<0.001$). A child with an average level of 48-month psychosocial adversity (score 2.26) thus has a predicted 60-month IQ 8.77 points (0.94 SDs) lower than a child with adversity score of 0. Being stunted at 48 months predicted a 60-month IQ 6.39 points (0.69 SDs) lower ($t(48)=-2.93$, $p=0.005$). Socioeconomic status was also a significant predictor of 60-month IQ ($\beta=0.21$, $t(48)=2.89$, $p=0.006$). Covariates in Model 3 predicted 54.19% of 60-month IQ variance (Table 4).

Table 4. Multivariate regression coefficients for prediction of 60-month raw IQ (N=51)			
	Model 1	**Model 2**	**Model 3**
Adversity score, 48-mo.	-4.85*** (0.65)	-4.83*** (0.75)	-3.88*** (0.79)
Stunted (Yes/No), 48-mo.		-6.14* (2.38)	-6.388** (2.18)
SES composite score, 24-mo.			0.21** (0.07)
Constant	64.14*** (1.84)	65.51*** (2.12)	58.89*** (3.11)
R^2	0.40	0.47	0.54
df for model F statistic	1	2	3
df for parameter t statistic	50	49	48
F	55.58	21.44	17.75

Robust standard errors in parentheses; * p<0.05 **p<0.01 ***p<0.001

Model 4 assesses both SES and psychosocial adversity as categorical variables by quartile, with SES inverted as "poverty quartiles" (i.e. 4th quartile SES defined as 1st quartile poverty) for visualization of estimated effect sizes. Being in the top quartile of psychosocial adversity exposure (vs. bottom) predicted a 60-month IQ 1.13 standard deviations lower (t(44)=-3.88, p<0.001). Being in the lowest SES quartile (top quartile poverty) predicted a 60-month IQ 0.90 SDs lower (t(44)=-3.07, p=0.004), while being in the second SES quartile (second-to-top quartile poverty) predicted an IQ score 0.97 SDs lower (t(44)=-3.07, p=0.004) than someone in the highest SES quartile. On generalized linear hypothesis testing, there was no significant difference between estimated effect sizes for top quartiles of adversity versus poverty (F(1,44)=0.24, p=0.63). Being stunted predicted a 0.69 SD decrement in 60-month IQ (t(44)=-2.03, p=0.047) (Figure 1; See Table A6, Appendix for full model data).

Figure 1. Adjusted standardized associations of risk factors with 60-month IQ



Error bars: ±1SE *p<0.05 **p<0.01 ***p<0.001

IV. Discussion

Analyses presented here support the use of the CPAQ-B as a valid research measure assessing childhood psychosocial adversity as a neurodevelopmental risk factor among low-SES, urban Bangladeshi children ages 18-60 months. The scale appears to be culturally acceptable to participants and feasible. Qualitative work supports good construct and content validity with a high degree of tailoring to local context. Psychometric analyses show strong internal consistency, adequate test-retest and inter-rater reliability, and significant correlations both with existing similar measures and with future child IQ. Relatively lower internal consistency of the *Neglect* subscale at child age 18 months should not preclude use in this group, but does caution against direct score magnitude comparison across age groups based on potential differential item functioning.

Of note, this tailored measure of psychosocial risk has correlated more substantially with measures of child cognition, here WPPSI IQ, than non-tailored measures such as the MSPSS and

the EPDS. This observation highlights the value of context-tailored measures, and of local validation of tools adopted across settings. Lack of such efforts may bias the field towards underestimation of psychosocial variable effects. Initial application of the CPAQ-B as a measure of neurodevelopmental risk shows large effect sizes when predicting 60-month IQ based on 48-month adversity exposure. Considering standardized coefficients, for instance, being in the top quartile for psychosocial adversity was associated with a >1 SD difference in 60-month IQ, similar to the effect associated with being in the lowest quartile of SES. This is a very preliminary application, providing no evidence of causality. A particular issue is that measuring SES at 24 months may leave substantial residual confounding if 24-to-48-month change in SES accounts for adversity score variance. Still, the fact that stunting is also expected to correlate with SES allows some fruitful reflection upon relative influence of risk factors. Initial application of the CPAQ-B, regardless, supports a strong case that a meaningful variable has been captured warranting serious ongoing assessment of its relationships with neurodevelopmental outcomes.

Several limitations are important to consider. At this stage, sample sizes did not allow for confirmatory factor analysis using split-half samples to confirm EFA findings on subscale structure. At this stage, validity evidence is strongest for using overall instrument score as a marker of cumulative psychosocial risk; further validation efforts can strengthen evidence for use of subscale scores to assess specific risk sub-domains.

More broadly, a key challenge is the lack of "gold standard" measures against which to measure CPAQ-B performance. As noted, even where comparator instruments exist, there is often no published data on local validity. If scores differ versus the CPAQ-B, it remains difficult to know which to trust. Some divergence can be explained by differences in factors such as recall

period and imperfect construct overlap, though likely not all. Evidence that the CPAQ-B subscales may be better predictors of later child IQ might favor the context-tailored measure.

A further issue is the difficulty of determining whether psychosocial adversities have truly been captured comprehensively. While grounded in qualitative work, CPAQ-B subscales may leave important experiences unexplored given the complexity of social experience. Importantly, the CPAQ-B in its present format does not capture fathers' experiences; adaptation of the scale for use in male parents and non-parental caregivers remains a key priority.

Finally, the context-tailored nature of the CPAQ-B represents a strength but also a limitation. Additional work would be needed to adapt and validation the instrument for use in rural areas and other low resource settings outside of Bangladesh. More broadly, however, the methodology and basic conceptual model generated here could be used to achieve rapid adaptation of the CPAQ-B to other settings. Potential measurement gains achieved with development of the CPAQ-B (for instance, stronger evidence for ability to predict IQ) may highlight the value of investing in local tailoring and validation.

Additional future priorities include generation of further data to assess retest reliability at the subscale level given large ICC confidence intervals. Ultimately, future work will make use of the CPAQ-B to consider how psychosocial risks shape child development more broadly, across a variety of low-resource settings. In relating CPAQ-B to neuroimaging findings, outcomes of particular interest include measures of EEG power and coherence, volumetric parameters on structural MRI, and analysis of differences in response to social and nonsocial stimuli on fNIRS.

V. Conclusion

The CPAQ-B is a novel research tool measuring childhood psychosocial adversities among low-SES children ages 18-60 months in Dhaka, Bangladesh, with strong initial validity

evidence. Moving forward, the instrument may be employed to examine relationships between early psychosocial stress and altered neurodevelopment as assessed by IQ testing, MRI, EEG, fNIRS, and other methods. Experience with the CPAQ-B affirms that measuring early psychosocial risks is difficult but important in light of potential links to child adverse outcomes. It highlights the importance of seeking good fit between context and measure, and of assessing instrument validity across diverse populations. The CPAQ-B offers an important step in beginning to assess difficult-to-measure child risks, and stands to make a significant contribution if used thoughtfully.

APPENDED TABLES

Table A1. Prior instruments consulted during item generation

General
Adverse Childhood Experiences Questionnaire[2]
WHO Adverse Childhood Experiences International Questionnaire[19]

Child abuse
Conflict Tactics Scale Parent-Child version[20]
International Society for Prevention of Child Abuse & Neglect Child Abuse Screening Tool, Parent Version[35]
Multiple Index Cluster Survey Household Questionnaire[37]

Child neglect
Home Observation Measurement of the Environment (HOME) Inventory[33,34]
Multidimensional Neglectful Behavior Scale[36]
Multiple Index Cluster Survey Household Questionnaire[37]

Caregiver depression, social isolation, and stress
Center for Epidemiological Studies Depression Scale
Edinburgh Postpartum Depression Scale[32]
Multidimensional Scale of Perceived Social Support[29]

Domestic violence
Demographic & Health Surveys Domestic Violence Module[31]
Psychological Maltreatment of Women Inventory[38]
WHO Multi-Country Domestic Violence Surveys[39]

Table A2. Factor loadings from exploratory factor analysis on child-focused items

Variable	Item	**Factor1** Abuse	**Factor2** Low warmth	**Factor3** Neglect A-Supervision	**Factor4** Neglect B-Basic care
q20	Threatened to beat child	0.8612			
q13	Beat child lightly	0.8542			
q18	Criticized child with names	0.6865			
q21	Frightened child	0.6739			
q19	Cursed child with slang	0.6213			
q17	Shook child	0.6016			
q14	Beat child harshly	0.5272			
q15	Beat child with object	0.4648			
q16	Beat child so cut, swollen, bruised	0.3678			
q9	Called sweet names(-)		0.8953		
q11	Showed physical affection(-)		0.827		
q10	Praised child(-)		0.8111		
q8	Talking for enjoyment(-)		0.7356		
q12	Took on lap or bosom to quiet(-)		0.5889		
q6	In house >1hr without supervision			0.909	
q5	Left in care of other qld >1hr			0.9017	
q7	Left free outside			0.4916	
q1	Stayed dirty for a whole day				0.7631
q2	Dirty clothes or naked				0.7485
q3	No one gave food or milk				0.4434

Note: Factor loadings <0.30 not shown

APPENDED TABLES

Table A3. Factor loadings from exploratory factor analysis on caregiver-focused items

Item	Question (paraphrase)	Factor1 maternal depression	Factor2 social isolation	Factor3 domestic violence	Factor4 economic stress	Factor5 family conflict	Factor6 marital breakdown
q55	Don't like anything	0.9652					
q54	No mood to talk/socialize	0.9491					
q57	Unable to work	0.9323					
q53	Loss of appetite	0.9009					
q52	Trouble sleeping	0.8265					
q56	Crying a lot	0.7171					
q51	Physical sx	0.712					
q50	Restless/shaky	0.7007					
q49	Anxious/worried	0.6934					
q48	Persistent sadness/sorrow	0.6922					
q58	Unable to feel joy with child	0.6343					
q32	No one to give kindness/encouragement		0.9171				
q33	No one to give advice		0.9124				
q31	No one to cheer you up		0.901				
q30	All alone		0.7656				
q29	Home was hell or no peace		0.4235				
q43	Beaten after using drugs/alcohol			0.5614			
q41	Child saw/heard you beaten			0.6481			
q37	Pushed, shook, or threw you			0.7282			
q39	Threatened to beat you			0.7771			
q35	Beat you harshly			0.8215			
q36	Beat you with object			0.8371			
q34	Beat you a little			0.8436			
q40	Afraid for your life			0.8462			
q38	Kicked/dragged you, beat you up			0.8822			
q46	Child lacked nutritious food				0.9811		
q47	Could not buy child needed basics				0.8882		
q45	Food insecurity				0.781		
q44	Tension about money				0.5619		
q28	Quarreled w husband's fam					0.876	
q27	Husband's fam criticized					0.8425	
q25	Quarrels with husband					0.5256	
q42	Child saw/heard you yelled at					0.4687	
q24	Husband very critical					0.4019	
q23	Husband unfaithful						0.6766
q21	Husband not giving money						0.6545
q22	Separated due to problems						0.6522
q26	Husband denied you food						0.5625

Note: Factor loadings <0.30 not shown

Table A4. Factor loadings from exploratory factor analysis on community-focused items

Item	Question (paraphrase)	Factor1 Community stress
q62	Gangs or drug trade	0.6915
q59	Violence in community	0.6514
q60	Crime in community	0.6502
q63	Eve teasing	0.6018
q61	Trash & poor sanitation	0.4568
q64	Class discrimination	0.4159

APPENDED TABLES

Table A5. Subscale test-retest and inter-rater reliability				
	Test-Retest Reliability		Inter-rater reliability	
	ICC	95% CI	ICC	95% CI
Child maltreatment				
Harsh discipline & abuse	0.87	0.75-0.93	0.78	0.58-0.88
Neglect	0.75	0.53-0.87	0.67	0.37-0.82
Low caregiver warmth	0.81	0.63-0.90	0.75	0.53-0.87
Maternal & family stress				
Depression	0.83	0.67-0.91	0.52	0.09-0.75
Physical domestic violence	0.84	0.69-0.91	0.70	0.44-0.84
Family conflict	0.67	0.37-0.83	0.67	0.37-0.82
Economic stressors	0.87	0.75-0.93	0.74	0.52-0.86
Social isolation	0.62	0.28-0.80	0.50	0.06-0.74
Marital breakdown	0.88	0.77-0.94	0.91	0.83-0.95
Environmental adversities				
Community stressors	0.70	0.43-0.84	0.47	0.00-0.72
Full scale	0.88	0.76-0.93	0.78	0.58-0.88

Table A6. Prediction of age-60-month child IQ by risk factors modeled categorically	
	Model 4
Adversity score, 48-mo.	
2nd Quartile	-0.25 (0.25)
3rd Quartile	-0.64* (0.27)
4th Quartile	-1.13*** (0.29)
Stunted (Yes/No), 48-mo.	-0.54* (0.26)
Poverty (inverse SES), 24-mo.	
2nd Quartile	-0.45* (0.22)
3rd Quartile	-0.98** (0.32)
4th Quartile	-0.90** (0.29)
Constant	64.14*** (1.84)
R^2	0.40
df for model F statistic	1
df for parameter t statistic	50
F	55.58

APPENDED FIGURES

Figure A1. Phases of instrument design methodology

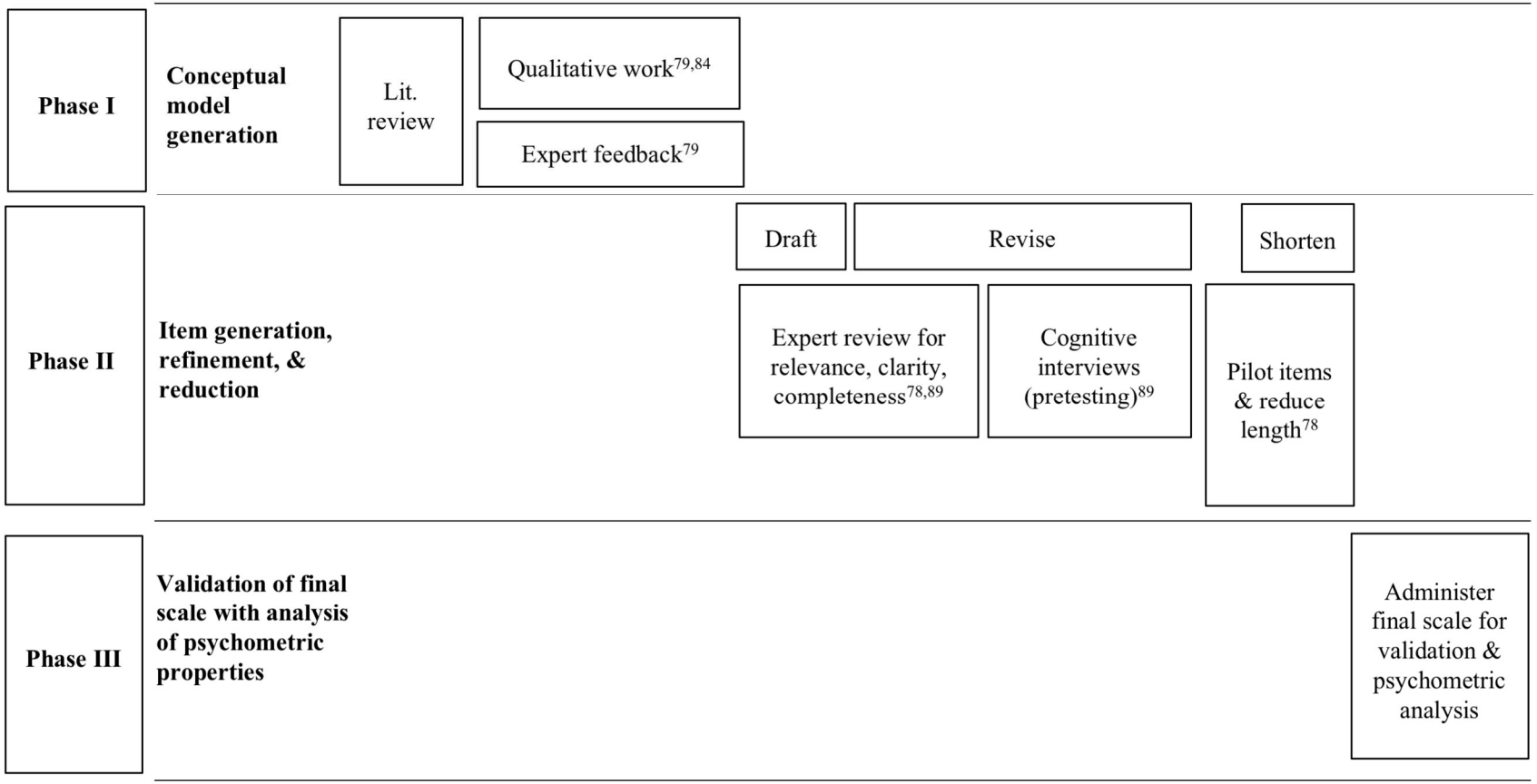
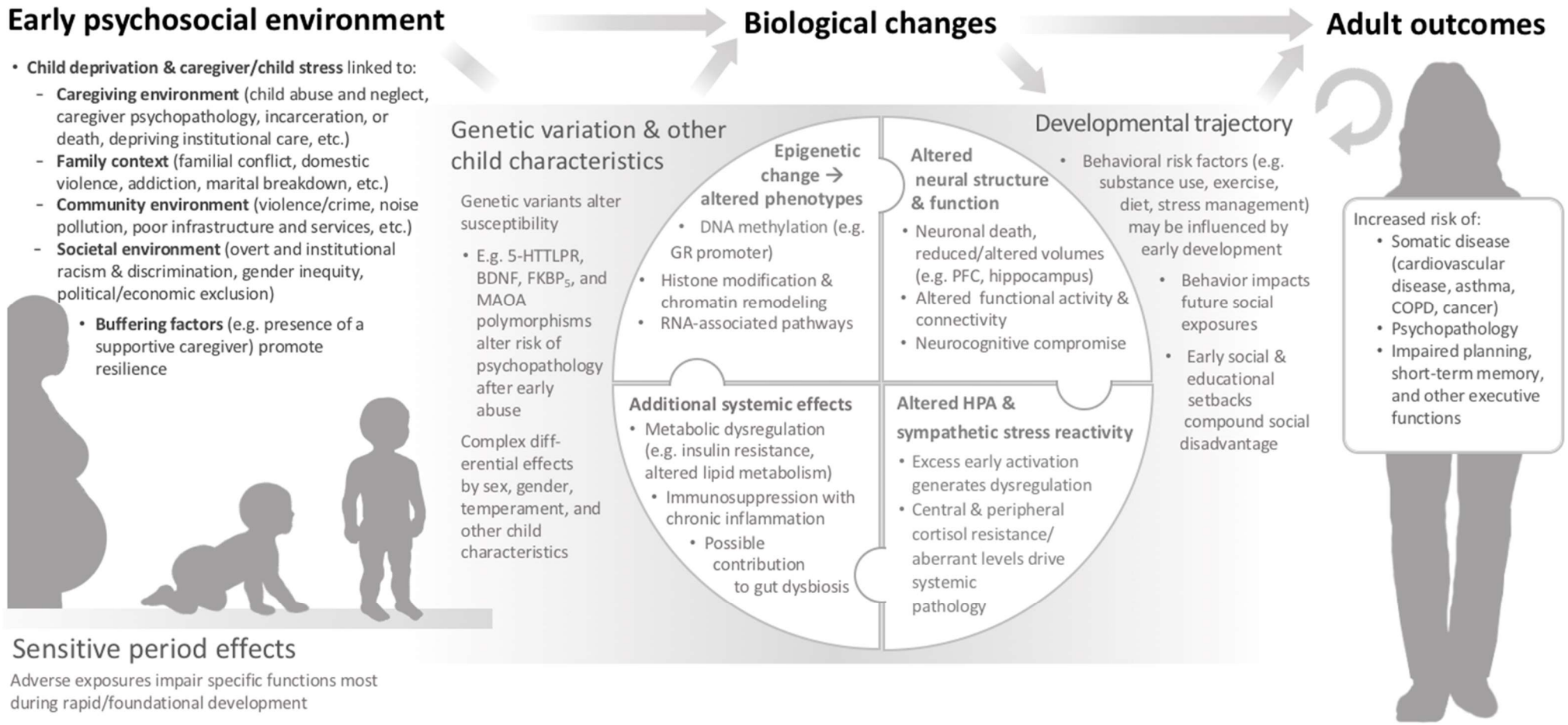


Figure A2. Model of the biological embedding of early psychosocial adversity



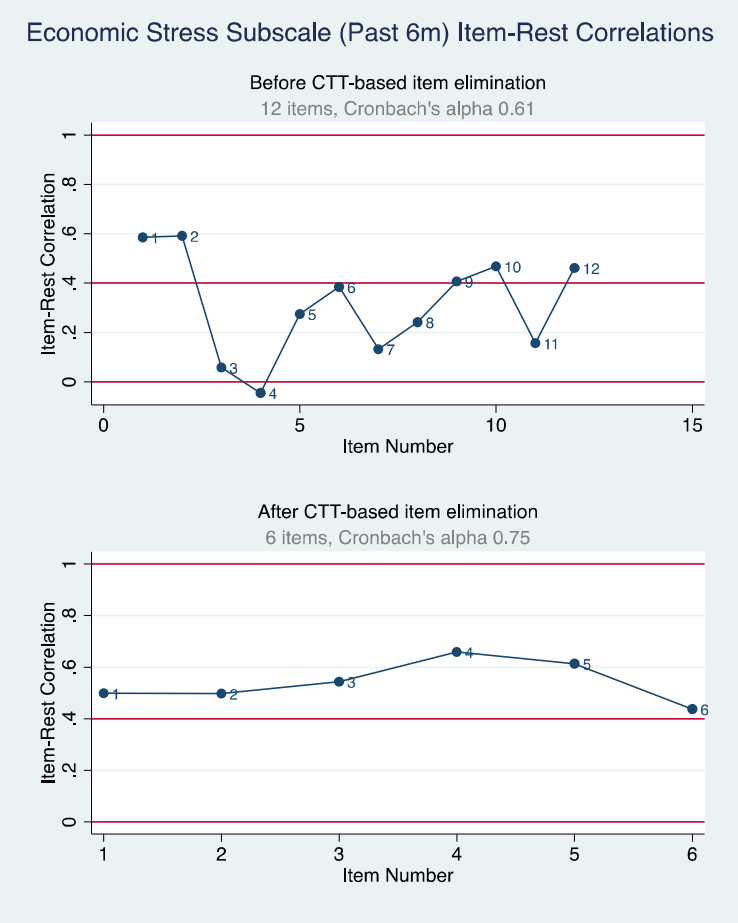
Legend—GR: glucocorticoid receptor; PFC: prefrontal cortex; COPD: chronic obstructive pulmonary disease

Appears in Berens et al. (2017)[52]

APPENDED FIGURES

Figure A3. CTT-based scale shortening—Item-rest correlations for economic stress items

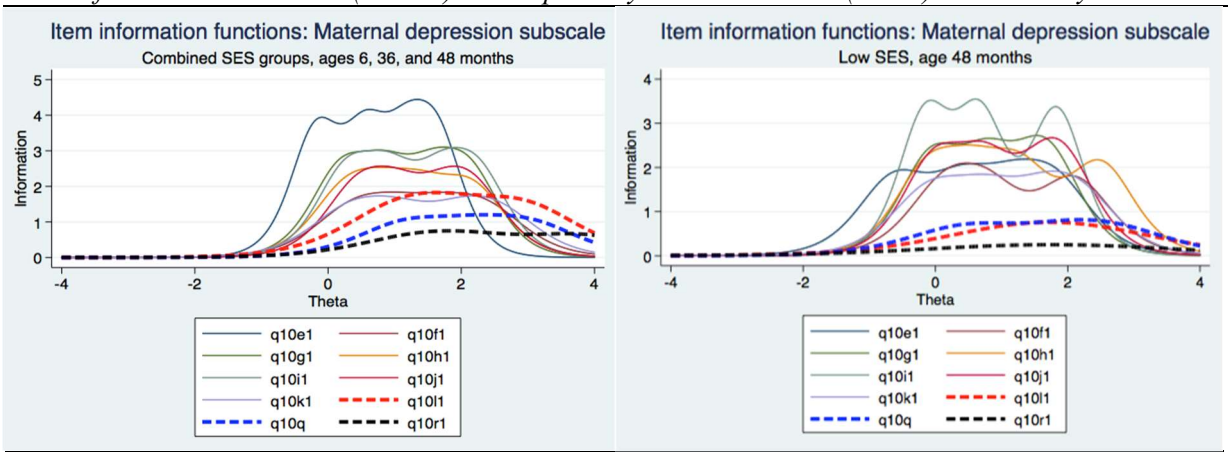
Before & after CTT-based item elimination



CTT=Classical test theory

Figure A4. IRT-based scale shortening—Item information for maternal depression items

Model fit to mixed-SES data (N=263) then separately to low-SES data (N=53) as sensitivity check



IRT=Item response theory; SES=Socioeconomic status

REFERENCES

1. Barboza Solís C, Kelly-Irving M, Fantin R, et al. Adverse childhood experiences and physiological wear-and-tear in midlife: Findings from the 1958 British birth cohort. *Proceedings of the National Academy of Sciences.* 2015;112(7):E738-E746.
2. Thomas C, Hypponen E, Power C. Obesity and type 2 diabetes risk in midadult life: the role of childhood adversity. *Pediatrics.* 2008;121:e1240-1249.
3. Surtees PG, Wainwright NWJ, Willis-Owen SAG, Luben R, Day NE, Flint J. Social adversity, the serotonin transporter (5-HTTLPR) polymorphism and major depressive disorder. *Biological Psychiatry.* 2006;59:224-229.
4. Danese A, Moffitt TE, Harrington H, et al. Adverse childhood experiences and adult risk factors for age-related disease: depression, inflammation, and clustering of metabolic risk markers. *Archives of pediatrics & adolescent medicine.* 2009;163(12):1135-1143.
5. Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. *American journal of preventive medicine.* 1998;14(4):245-258.
6. Liu Y, Croft JB, Chapman DP, et al. Relationship between adverse childhood experiences and unemployment among adults from five U.S. states. *Soc Psychiatry Psychiatr Epidemiol.* 2013;48(3):357-369.
7. Giovanelli A, Reynolds AJ, Mondi CF, Ou SR. Adverse Childhood Experiences and Adult Well-Being in a Low-income, Urban Cohort. *Pediatrics.* 2016;137(4).
8. Flaherty EG, Thompson R, Dubowitz H, et al. Adverse childhood experiences and child health in early adolescence. *JAMA pediatrics.* 2013;167(7):622-629.
9. Richards M, Wadsworth MEJ. Long term effects of early adversity on cognitive function. *Archives of Disease in Childhood.* 2004;89(10):922-927.
10. Kelly-Irving M, Mabile L, Grosclaude P, Lang T, Delpierre C. The embodiment of adverse childhood experiences and cancer development: potential biological mechanisms and pathways across the life course. *Int J Public Health.* 2013;58(1):3-11.
11. Berens AE, Nelson CA. The science of early adversity: is there a role for large institutions in the care of vulnerable children? *Lancet (London, England).* 2015;386(9991):388-398.
12. De Bellis MD, Hooper SR, Spratt EG, Woolley DP. Neuropsychological Findings in Childhood Neglect and their Relationships to Pediatric PTSD. *Journal of the International Neuropsychological Society : JINS.* 2009;15(6):868-878.
13. Knudsen EI. Sensitive periods in the development of the brain and behavior. *Journal of cognitive neuroscience.* 2004;16:1412-1425.
14. Garner AS, Shonkoff JP, Siegel BS, et al. Early Childhood Adversity, Toxic Stress, and the Role of the Pediatrician: Translating Developmental Science Into Lifelong Health. *Pediatrics.* 2011;129(1):e224.
15. Berens AE, Jensen SK, Nelson CA. Biological embedding of childhood adversity: from physiological mechanisms to clinical implications. *BMC Medicine.* 2017;15:135-147.
16. Hillis S, Mercy J, Amobi A, Kress H. Global prevalence of past-year violence against children: a systematic review and minimum estimates. *Pediatrics.* 2016;137(3):1-13.
17. Pinheiro PS. *World Report on Violence Against Children.* Geneva: United Nations Secretary-General's Study On Violence Against Children;2006.

18. Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *The Lancet.*370(9590):851-858.
19. Black MM, Walker SP, Fernald LCH, et al. Early childhood development coming of age: science through the life course. *The Lancet.* 2016;389(10064):77-90.
20. Engle PL, Black MM, Behrman JR, et al. Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *The Lancet.* 2007;369(9557):229-242.
21. Jensen SKG, Berens AE, Nelson CA, 3rd. Effects of poverty on interacting biological systems underlying child development. *The Lancet Child & Adolescent Health.* 2017;1(3):225-239.
22. Storrs C. How Poverty Affects the Brain. *Nature Magazine*. London: Nature; 2017.
23. Kirkpatrick BD, Colgate ER, Mychaleckyj JC, et al. The "Performance of Rotavirus and Oral Polio Vaccines in Developing Countries" (PROVIDE) study: description of methods of an interventional study designed to explore complex biologic problems. *The American journal of tropical medicine and hygiene.* 2015;92(4):744-751.
24. Korpe PS, Haque R, Gilchrist C, et al. Natural History of Cryptosporidiosis in a Longitudinal Study of Slum-Dwelling Bangladeshi Children: Association with Severe Malnutrition. *PLoS neglected tropical diseases.* 2016;10(5):e0004564.
25. Gehlbach H, Brinkworth ME. Measure twice, cut down error: A process for enhancing the validity of survey scales. *Review of General Psychology.* 2011;15(4):380.
26. Rattray J, Jones MC. Essential elements of questionnaire design and development. *Journal of clinical nursing.* 2007;16(2):234-243.
27. Onwuegbuzie AJ, Dickinson WB, Leech NL, Zoran AG. A qualitative framework for collecting and analyzing data in focus group research. *International journal of qualitative methods.* 2009;8(3):1-21.
28. Hsieh H-F, Shannon SE. Three approaches to qualitative content analysis. *Qualitative health research.* 2005;15(9):1277-1288.
29. Holmbeck GN, Devine KA. An author's checklist for measure development and validation manuscripts. Oxford University Press; 2009.
30. Kane MT. Validating the Interpretations and Uses of Test Scores: Validating the Interpretations and Uses of Test Scores. *Journal of Educational Measurement.* 2013;50(1):1-73.
31. Schaeffer NC, Presser S. The Science of Asking Questions. *Annual Review of Sociology.* 2003;29(1):65-88.
32. Krosnick JA. Survey research. *Annual review of psychology.* 1999;50(1):537-567.
33. Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951;16(3):297-334.
34. DeVellis RF. *Scale development: Theory and applications.* Vol 26: Sage publications; 2016.
35. Yen W, Fitzpatrick A. Item response theory. In: Brennan R, ed. *Educational measurement*. 4th ed. Westport, CT: American Council on Educaiton, Praeger Publishers; 2006.
36. Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment.* 2005;10(7):1-9.

37. Doom JR, Gunnar MR. Stress in Infancy and Early Childhood: Effects on Development. *International Encyclopedia of the Social & Behavioral Sciences.* 2015;23:577-582.
38. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale. *The British journal of psychiatry.* 1987;150(6):782-786.
39. Cox JL, Chapman G, Murray D, Jones P. Validation of the Edinburgh Postnatal Depression Scale (EPDS) in non-postnatal women. *Journal of affective disorders.* 1996;39(3):185-189.
40. Gausia K, Fisher C, Algin S, Oosthuizen J. Validation of the Bangla version of the Edinburgh Postnatal Depression Scale for a Bangladeshi sample. *Journal of Reproductive and Infant Psychology.* 2007;25(4):308-315.
41. Aboud FE. Evaluation of an Early Childhood Parenting Programme in Rural Bangladesh. *Journal of Health, Population, and Nutrition.* 2007;25(1):3-13.
42. Zimet GD, Dahlem NW, Zimet SG, Farley GK. The Multidimensional Scale of Perceived Social Support. *Journal of Personality Assessment.* 1988;52(1):30-41.
43. Garcia-Moreno C, Jansen HA, Ellsberg M, Heise L, Watts CH. Prevalence of intimate partner violence: findings from the WHO multi-country study on women's health and domestic violence. *The lancet.* 2006;368(9543):1260-1269.
44. Pearson Clinical. *Wechsler Preschool and Primary Scale of Intelligence - Fourth Edition.* New York: Pearson North America; 2014.
45. Martikainen P, Bartley M, Lahelma E. Psychosocial determinants of health in social epidemiology. *International Journal of Epidemiology.* 2002;31(6):1091-1093.
46. McEwen BS, Gianaros PJ. Stress- and Allostasis-Induced Brain Plasticity. *Annual Review of Medicine.* 2011;62:431-445.
47. Hart H, Rubia K. Neuroimaging of child abuse: a critical review. *Frontiers in human neuroscience.* 2012;6:52.
48. Nelson CA, Fox NA, Zeanah. CH. *Romania's abandoned children: Deprivation, brain development, and the struggle for recovery.* Harvard University Press; 2014.
49. Smith KE, Landry SH, Swank PR. The Role of Early Maternal Responsiveness in Supporting School-Aged Cognitive Development for Children Who Vary in Birth Status. *Pediatrics.* 2006;117(5):1608-1617.
50. Yates TM. The Developmental Consequences of Child Emotional Abuse. *Journal of Emotional Abuse.* 2007;7(2):9-34.
51. Ellsberg M, Jansen HAFM, Heise L, Watts CH, Garcia-Moreno C. Intimate partner violence and women's physical and mental health in the WHO multi-country study on women's health and domestic violence: an observational study. *The Lancet.* 2008;371(9619):1165-1172.
52. Berens AE, Jensen SK, Nelson CA. The Biological Embedding of Early Adversity: Physiological Mechanisms & Implications for Clinical Practice *Under Review.* 2017.