Appendix 1. Study Protocol

Feasibility Study of the Masayang Pamilya Para sa Batang Pilipino (MaPa) Program for Families with Children and Adolescents

Lead Principal Investigator:

Prof Liane Peña Alampay, Ateneo de Manila University

Co-Principal Investigator:

Dr Jamie Lachman, University of Oxford

Co-Investigators:

Dr Rosanne Jocson, Ateneo de Manila University

Prof Catherine Ward, University of Cape Town

Prof Frances Gardner, University of Oxford

Mr Engels Del Rosario, Department of Social Welfare and Development

Dr Bernadette Madrid, Child Protection Network

Date: version 2.0 dated 31 March 2019

Project funding by: United Nations Children's Fund (UNICEF) Philippines

Table of Contents

Protocol Summary

Research Aim

The aim of this study is to test the feasibility of (1) an 8-session version of the *Masayang Pamilya* (MaPa) parenting program for families with children aged 2-9 (MaPa Kids) and (2) a 9-version culturally and contextually adapted MaPa parenting program for families with children aged 10-17 (MaPa Teens).

This study protocol will be registered on ClinicalTrials.gov.

Objectives

1. To test the feasibility of the MaPa Kids 8-session and MaPa Teens 9-session training modules for low-income families, by assessing:
    a. Implementation fidelity and quality;
    b. Recruitment, retention, and engagement of participants;
    c. Cultural and contextual relevance, acceptability, and satisfaction;
2. To evaluate the initial effects of the MaPa Kids 8-session and MaPa Teens 9-session training modules at immediate post-test on the primary outcome of child maltreatment, and the proximal outcomes of positive parenting; dysfunctional parenting; and attitudes toward corporal punishment.
3. To evaluate the initial effects of the MaPa Kids 8-session and MaPa Teens 9-session training modules at immediate post-test on the secondary outcomes of child neglect, parental monitoring, child behavior problems, parenting efficacy, parenting stress, self-efficacy in managing emotions, parental depressive symptoms, adolescent depressive symptoms, parent exposure to intimate partner violence, intimate partner coercion, family functioning, and community violence exposure.

Participants

Participants for the feasibility pilot for MaPa Kids and Teens will include 60 parents or primary caregivers, 30 children between the ages of 10 and 17, and 8 facilitators who deliver the program.

Data collection and analysis

Feasibility pilot process data for Mapa Kids and Teens will be collected through focus group discussions with 60 parents/primary caregivers, 30 children ages 10-17, and facilitators. A self-report questionnaire will also be administered to 60 parents/primary caregivers and 30 children who participate in the program.

The feasibility pilot outcome evaluation will be conducted using standardized, validated scales to assess the program's effect on the primary outcome of child maltreatment, and the secondary outcomes of positive parenting; dysfunctional parenting; and attitudes toward corporal punishment, child neglect, parental monitoring, child behavior problems, parenting efficacy, parenting stress, self-efficacy in managing emotions, parental depressive symptoms, adolescent depressive symptoms, parent exposure to intimate partner violence, intimate partner coercion, family functioning, and community violence exposure.

Full study protocol

A. Study team

A.1. Research team

Principal Investigator: Professor Liane Peña Alampay (Ateneo de Manila University)

Co-Principal Investigator: Dr Jamie Lachman (University of Oxford)

Co-Investigators: Dr Rosanne Jocson (Ateneo de Manila University); Prof Catherine Ward (University of Cape Town); Prof Frances Gardner (University of Oxford); Mr Engels Del Rosario (Department of Social Welfare and Development); Dr Bernadette Madrid (Child Protection Network)

A.2. Project staff

A Project Coordinator (Ms Jennel Reyes) will be contracted for the purpose of coordinating the administrative and logistical arrangements for the research and implementation aspects of the study.

A.3. Funder

United Nations Children's Fund (UNICEF) Philippines

B. Background

B.1. Literature review

The 2015 national baseline survey on violence against children (VAC) revealed that 80% of Filipino youth respondents had experienced violence in their childhood, with 60% of these cases occurring in the home. Mothers, fathers, and siblings were the most commonly reported perpetrators of harsh physical and psychological punishment [1]. A logical recourse to decrease child maltreatment in the country, therefore, is to implement intervention programs that improve parents' relationships with their children and their knowledge and skills in child behavior management.

There is extensive scientific evidence that parenting support programs are effective in reducing child maltreatment and associated risk factors such as corporal punishment and parent negative psychological health (e.g., [2, 3]). International development agencies including UNICEF and the World Health Organization (WHO) strongly advocate "parent and caregiver support" as among seven evidence-based strategies for ending VAC (i.e. INSPIRE strategies). Parenting for Lifelong Health (PLH) is one such initiative led by UNICEF HQ and WHO to support evidence-based parenting programs to reduce VAC in low and middle-income contexts. Likewise, the Philippine Plan of Action to End VAC 2017-2022 (PPAEVAC) has prioritized the development of "evidence-based parenting skills and positive discipline" as among its key result areas, the institutional lead of which is the DSWD. Thus, the department is a key partner in this proposed parenting program, which is also part of the Work Plan between DSWD and UNICEF.

To support these national and global initiatives to address VAC, the multisectoral collaboration Parenting for Lifelong Health (PLH)-Philippines[1] (of which ADMU is part) developed and tested a locally adapted parent support intervention based on a prototype program previously tested in South Africa. In 2016-2017, PLH-Philippines embarked on the cultural adaptation, feasibility study (N=30), and pilot randomized control trial[2] (RCT) (N=120) of the *Masayang Pamilya Para sa Batang Pilipino* (MaPa) program in NCR, which was implemented with families with children ages 2-6 and who were beneficiaries of the Department of Social Welfare and Development (DSWD) *Pantawid Pamilyang Pilipino Program* (4Ps). The results of this initial RCT were promising: at post-intervention, families who participated in MaPa reported 53% less physical

[1] PLH-Philippines is a collaboration between ADMU; Universities of Oxford, Bangor, Cape Town; Philippine Child Protection Network; Department of Social Welfare and Development (DSWD); UNICEF; UBS-Optimus Foundation

[2] Funded through a SSFA with ADMU

abuse, 44% less emotional abuse, and 40% less child neglect compared to control group families who underwent the 4Ps Family Development Sessions (FDS).

B.2. Research Aim

The current program aims to expand the development, implementation, and evaluation of the MaPa program as part of a systematic effort to further assess the acceptability, feasibility, and effectiveness of the interventions. The aim of the study is to test the feasibility of (1) an 8-session version of the Masayang Pamilya (MaPa) parenting program for families with children aged 2-9 (MaPa Kids) and (2) a 9-session culturally and contextually adapted MaPa parenting program for families with children aged 10-17 (MaPa Teens).

B.3. Objectives

1. To test the feasibility of the MaPa Kids 8-session and MaPa Teens training modules for low-income families, by assessing:
    a. Implementation fidelity and quality;
    b. Recruitment, retention, and engagement of participants;
    c. Cultural and contextual relevance, acceptability, and satisfaction;
2. To evaluate the preliminary effectiveness of the MaPa Kids 8-session and MaPa Teens 9-session training modules at immediate post-test on the primary outcome of child maltreatment, and the secondary outcomes of positive parenting, dysfunctional parenting, and attitudes toward corporal punishment.
3. To evaluate the preliminary effectiveness of the MaPa Kids 8-session and MaPa Teens training modules at immediate post-test on the secondary outcomes of child neglect, parental monitoring, child behavior problems, parenting efficacy, parenting stress, self-efficacy in managing emotions, parental depressive symptoms, adolescent depressive symptoms, parent exposure to intimate partner violence, intimate partner coercion, family functioning, and community violence exposure.

C. Pilot feasibility evaluation

The MaPa Kids 8-session and Teens 9-session modules will be piloted through a pre-post feasibility study for 60 families divided by child development age groups (i.e., 2-9 and 10-17). Piloting will include the following activities:

a) implementation of parenting intervention modules by local facilitators and coaches in communities identified in collaboration with DSWD;

b) monitoring of implementation feasibility including participation and engagement of families as well as quality of program delivery by facilitators;

c) pre-post assessment of the impact of the intervention on reducing violence against children, improving positive parenting, and addressing risk factors associated with violence against children.

C.1 Study design

The feasibility pilot of the program modules has a pre-post design with no control group, with the aims of assessing program implementation, cultural and contextual relevance, and study feasibility. Although there will be no comparison group and it is not designed to test effects, the pilot also has a provisional goal of reductions in child physical and emotional abuse at immediate post-intervention.

Setting

The feasibility pilot will be conducted a low-income community in Quezon City, one of the 16 cities in the National Capital Region (NCR) in the Philippines. The primary criterion for site selection is that the 4Ps is operational in the community and includes a sizable number of recipient families with at least one child aged 2-9 or 10-17 years old. The recommendations and readiness of the city DSWD and 4Ps personnel to coordinate with the researchers is also a factor in site selection.

The outcome evaluation and self-report questionnaires with the 60 parents/primary caregivers and 30 children ages 10 to 17 will be conducted either in the homes of participants or a local community center, in a room that is relatively quiet and free from outside disturbance. Participants will also be invited to take part in focus group discussions (FGDs) at a meeting facility in the community where the modules are delivered (60 parents/primary caregivers, 30 children ages 10 to 17, 8 facilitators, and 2 coaches). The focus group discussions with the expert advisory group will take place at a meeting facility in Ateneo de Manila University.

Recruitment

This pilot will use targeted/purposive sampling to recruit participating families ($N = 60$ parents/caregivers and 30 child respondents for the Teens module). This study will rely on referrals from the 4Ps local municipal staff for recruitment of possible study participants. The research team will work closely with the 4Ps staff to identify potential eligible parents who are recipients of the conditional cash transfer program. 4Ps staff will provide the research team with a list of 50 participants with children ages 2 to 9 years and 50 participants with children ages 10 to 17 years, of which 30 will be randomly selected from each age group for initial recruitment in order to avoid potential issues regarding perceptions of fairness in the recruitment process. If less than 30 of the selected participants provide consent and are eligible from each age group, the remaining adults will be randomly selected from the original list until the full target sample size is reached.

If participating families have more than one child between the ages 2-17, the Computer Assisted Self-Interviewing tablets will randomly select one of the children for the parent to report on during the study. Research assistants from Ateneo de Manila University's Bulatao Center will be employed to contact referred participants, conduct informed consent procedures, and administer baseline and post-intervention assessments. All staff will have extensive experience working with vulnerable families. They will be thoroughly trained and supervised by the co-principal investigators and co-investigators on recruitment, informed consent, and assessment procedures.

All participating parents/primary caregivers and child participants and all 8 facilitators and 2 coaches will be invited to participate in the focus group discussion for the process evaluation at program completion.

Three focus group discussions will be convened with 5 Filipino experts in the field of parenting interventions in order to guide program adaptation. This study will use targeted sampling to recruit experts.

Eligibility criteria

Participants must have provided written, informed consent prior to the occurrence of any study procedures.

Inclusion criteria for participating parents or caregivers (n = 60):

1. Age 18 or older;
2. Primary caregiver responsible for the care of a child between the ages 2-9 (Kids) or 10-17 (Teens);
3. Spend at least four nights a week in the same household as the child in the previous month;
4. Recipient of the 4Ps conditional cash transfer program;
5. Provision of consent to participate in the full study;
6. Provision of consent for their child to participate in the full study (for MaPa Teens only).

Exclusion criteria for adult parents:

1. Any adult who has already participated in the Parent Effectiveness Service;
2. Any adult exhibiting severe mental health problems or acute mental disabilities;
3. Any adult that has been referred to child protection services due to child abuse.

Inclusion criteria for child respondents (n = 30):

1. Age 10 to 17 years at initial assessment
2. Lives in the house at least 4 nights per week
3. Must have an adult primary caregiver who lives in the household, who provides consent, and who participates in the study
4. Provides assent to participate in the full study.

Exclusion criteria for child respondents:

1. Any child who is either experiencing severe mental health problems, has acute developmental disabilities,
2. If the child participant has been referred to social services during baseline data collection due to reported or observed indications of significant harm.

The study will also assess program fidelity and quality of delivery by program facilitators ($N = 8$) and coaches ($N = 2$).

These facilitators and coaches will have the following inclusion criteria:

1. Age 18 or older;
2. Prior participation in a 5-day facilitator training workshop;
3. Agreement to implement the entire program;
4. Provision of consent to participate in the full study.

The participants of the expert advisory group have the following inclusion criteria ($N = 5$):

1. An academic expert, government official, NGO service provider, or practitioner who is involved in researching, advising, and/or providing services to families with children;
2. Provision of consent to participate in the study

Interventions

This study will assess the feasibility of two evidence-based interventions from the Parenting for Lifelong Health program suite: Parenting for Lifelong Health for Young Children and Parenting for Lifelong Health for Adolescents [4]. These programs are being adapted and tested for the Filipino context as the *Masayang Pamilya Para Sa Batang Pilipino* Parenting Program for families with children (MaPa Kids) and adolescents (MaPa Teens) [5].

The MaPa Kids program is delivered to groups of parents ($N = 15$ per group) and includes the following content: 1) spending one-on-one time with children; 2) describing actions and feelings for cognitive development and socio-emotional awareness; 3) using praise and rewards to encourage positive behavior; 4) establishing limits through effective instruction giving and consistent household rules; 5) nonviolent discipline such as ignoring negative attention seeking behavior, and consequences for noncompliance, rule-breaking, and aggressive behavior; 6) problem solving with children; 7) mindfulness based stress reduction for caregivers and adolescents; and 8) reflection and moving on. The MaPa Teens program is delivered to groups of parents (N = 15 per group) and groups of teens (N = 15 per group) and includes the following content: 1) one-on-one time with parents and teens; 2) positive reinforcement of positive behavior; 3) managing anger and stress; 4) establishing rules and routines; 5) family budgeting; 6) accepting responsibility for actions; 7) resolving family conflicts; 8) keeping safe in the community and resolving conflicts; and 9) reflection and moving on.

In 2016-2017, PLH-Philippines embarked on the cultural adaptation, feasibility study (N=30), and pilot randomized control trial (RCT) (N=120) of MaPa Kids in NCR, which was implemented with families with children ages 2-6 and who were beneficiaries of the Department of Social Welfare and Development (DSWD) *Pantawid Pamilyang Pilipino Program* (4Ps). The results of this initial RCT were promising: at post-intervention, families who participated in MaPa reported 53% less physical abuse, 44% less emotional abuse,

and 40% less child neglect compared to control group families who underwent the 4Ps Family Development Sessions (FDS). Based on feedback from the DSWD regarding cost constraints, MaPa Kids will be adapted from 12 to 8 sessions in this study [6].

The MaPa Teens program has been adapted from PLH Teens, one of the few low-cost parent-adolescent interventions that has been rigorously tested in a LMIC, as highlighted in the WHO's *INSPIRE: Seven Strategies to End Violence Against Children* [7]. It has been implemented in 15 countries by international and local NGOs as part of multiple USAID-funded projects. Originally a 14-session parenting program delivered by community facilitators to groups of parents and their teens (girls and boys), the MaPa Teens program has been reduced to 8 parenting sessions plus 1 session on family budgeting to be in alignment with MaPa Kids. Grounded in social learning theory, it uses non-didactic, participatory methods to reduce adolescent exposure to violence in the home and community by improving positive parenting and parent-child communication, while reducing interpersonal conflict, harsh discipline, parenting stress, and adolescent conduct problems, risky behavior, and mental ill-health. Families are also assisted in developing child safety plans, responding to abuse, family budgeting, and accessing medical and social services [8].

Incentives

Adult participants will be offered a gift card (PHP 500 or approximately £7 per participant) after baseline and again after post-intervention assessment. Child participants will receive a token worth PHP 200 (approximately £3 per participant) after baseline and again after post-intervention assessment. Refreshments (i.e., juice/soda, tea, sugar, and bread) will be provided during data collection.

All participants will receive a certificate of completion at the end of the program, as well as a certificate for participating in the study at post-test. To date, we have consistently found a high level of willingness among communities to participate in research studies on parenting programs.

Participants of focus group discussion will be offered refreshments (i.e., juice/soda, tea, sugar, and bread) during the sessions.

C.2 Parenting experts working group

a) <u>Overview</u>

Three focus group discussions will be convened with five experts in the field of parenting interventions in order to guide program adaptation. Drawing from the results of the formative evaluation, the Parenting Experts Working Group will determine whether and to what extent the MaPa programs should be adapted, considering its contextual and cultural relevance and feasibility. The Working Group will advise the research team on the content of the program, including whether any components may need to be altered or excluded, as well as if additional components or aspects should be incorporated. The Working Group will also provide recommendations on adaptations to the structure of the program and methods of delivery, with a focus on ensuring a balance between fidelity and fit to the targeted population in the Philippines. On-going email communication and consultations with particular Working Group members on a one-on-one basis will assist in guiding the program developers in designing the curriculum for training facilitators and supervisors, as well as in the development of the facilitator manual and parent guidebook. Records (e.g., minutes of discussions) generated by this Expert Working Group will also form data for the study, as they will be used to inform data analysis and recommendations [9].

This study will use targeted sampling to recruit Filipino experts in the field of parenting interventions in order to guide program adaptation. These experts will be men and women who have provided written informed consent. It is anticipated that a total of five persons will be appropriate for this Working Group, in order to promote close collaboration and active engagement by all members. A limited Working Group size will also help to ensure that all focus group discussions can be held with all members present, which is necessary given the time limitations for this project. Potential respondents will be recruited via telephone or e-mail.

b) <u>Data collection</u>

Each of the three focus group discussions will last approximately 2-3 hours and will occur in office meeting rooms. If more time is required, the lead facilitator will ask the Working Group if members are willing to

continue meeting for an additional 30 minutes; the time will only be extended if all members agree. Discussions will be audio recorded, transcribed verbatim, and translated into English when necessary. Co-Investigator Jocson will also record simultaneous written notes in English, with particular attention paid to key decisions made and issues that particularly stirred debate. In the event that the audio or video recording fails or is incomplete, these notes will serve as a critical backup.

c) Focus group protocol

The Parenting Experts Working Group discussions will be conducted using an open-ended approach to soliciting views and feedback from Working Group members, in a manner that is unbiased and non-judgmental. The lead facilitator (Co-Principal Investigator Alampay) and the co-facilitator (Co-Investigator Jocson), will conduct the Working Group discussions according to the major agenda items for each meeting (refer to Appendix 5), while simultaneously allowing additional items and tasks to emerge based on the importance placed on particular topics of discussion by Working Group members. The major agenda items may shift between Working Group discussions, depending on the flow of discussion and progress made in-between meetings (e.g. via e-mail).

d) Data analysis

The research team will use a thematic approach within an experiential framework in order to analyze the data from the focus group discussion transcripts. An initial coding framework will be developed from the transcripts by two independent coders using NVivo. If translations of focus group discussions are required, any questions arising due to translation or context will be clarified with the lead facilitator and/or professional translators. The codes will then be grouped into themes based on the respondents' feedback on PLH 2-9 core themes, program structure and schedule, program logistics, methods for program delivery, and implications of key findings from the feasibility pilot on further program adaptation. After the coders have agreed on the emergent, commonly expressed views and deviating opinions, transcripts will again be assessed. Findings will then be discussed within the research team with an emphasis on the validity and representativeness of individual themes. Finally, we will select data extracts that represent key themes as well as diverging viewpoints.

C.3 Outcome data collection

Ten data collectors who are fluent in Filipino (the local language) and who have prior experience working with vulnerable families in the Philippines will serve as research assistants. They will be extensively trained in ethics, informed consent, interviewing, and observational assessment techniques by the study investigators. All data collection points will have a detailed protocol, and will be scripted to ensure consistency across data collectors. Research assistants will be managed by co-PI Alampay and co-I Jocson. Quantitative self-report assessments and qualitative interview/focus group guides will be translated into Filipino by bilingual researchers and the translation checked by back-translation [10]

The proposed study will use an innovative, low-cost technological data collection tool using Computer-Assisted Self-Interviewing ('CASI') methods with e-tablet technology to administer consent forms and questionnaires. This method of data collection has been pilot-tested during previous feasibility RCT studies with high acceptability to respondents. Trained data collectors will explain the CASI procedures, read out questions, and assist participants to key in responses to their handsets. CASI supports multiple languages, and so can include English as well as local languages. Although ongoing studies report exceptionally high acceptability of CASI in the Philippines, if any participants are unable or uncomfortable with the use of tablets, a paper-and-pen interviewer-assisted questionnaire will also be available in English and Filipino.

In addition, the study will be using audio-CASI to administer sensitive items on the questionnaires regarding child maltreatment, adolescent risk behavior, and intimate partner violence. Audio-CASI techniques have been shown to be especially effective in increasing participant willingness to disclose highly stigmatized activities or experiences [11, 12].. In particular, audio-CASI may decrease a respondent's anxiety in answering questions face-to-face and thus increase disclosure of abuse [13]. Audio-CASI will also help to address low levels of literacy encountered in low-income populations. Participants will have the option to either read or use the audio function when responding to questions regarding child maltreatment and intimate partner violence. After a tutorial on how to use the audio function, Respondents will receive earphones during this part of the questionnaire while the interviewer completes paperwork in order to allow the respondent as much privacy as

possible. In ongoing studies in the Philippines, the team has found this approach to result in a significant increase in the disclosure of sensitive items, as well as a high level of acceptability. The present study will also assess its efficacy and acceptability during the process evaluation.

Data on implementation fidelity will be collected by video-recording all program sessions. While only a subset of sessions is required for the assessment of quality of delivery, all sessions will be recorded as some videos may not be easy to code due to poor recording, or may not be recorded due to human error. Also, it may provide a greater level of consistency for parents and facilitators to have all sessions recorded so that participants are more comfortable with the video assessments. It will be explained to parents that the purpose of the video is to assess facilitator skills and ensure that they are getting the best possible service. Nonetheless, recordings for assessment will be selected from videos of sessions eight, nine, or ten in order to assure that the facilitator assessed is leading the activity of interest.

Transcripts of video recordings will be produced prior to assessment. Two assessors who are fluent in Filipino, and trained supervisors of the MaPa program will code impressions based on a 4-point Likert scale: a score of "0" is defined as "little or no evidence that the facilitator has exhibited competency;" a score of "1" is defined as "the facilitator has a general understanding of what is expected but does not meet expected level of proficiency;" a score of "2" is defined as "the facilitator exhibits the required skills and is able to perform the activity with a high level of proficiency;" and a score of "3" is defined as "the facilitator demonstrates exceptional ability to deliver required skills or activity." Scores for overall proficiency as well as for each category and subscale will be normalized by calculating a ratio of facilitator score to the total possible score. In order to determine inter-rater reliability, both raters will code the same videos until assessors achieve at least an inter-cluster correlation coefficient threshold of 0.70 as an acceptable level of inter-rater reliability [14].

Program adherence data will be collected using attendance registers administered by research assistants assigned to the process evaluation. Program participants will indicate whether they completed their home activity assignments using self-report checklists.

Qualitative data will involve focus group discussions with all participating parents/primary caregivers and program facilitators and coaches. Data will be collected by co-PI Alampay or trained research assistants, all of whom speak fluent Filipino, using an interview schedule with an open-ended approach to soliciting views and feedback from participants. Interviewers will guide the focus groups along the following steps while allowing the discussion to evolve as themes emerge. Each step will begin with an overarching question followed by more specific questioning and inquiry:

> Question 1: What are participants' overall experiences and observed change during program?
>
> Question 2: Are the program materials, delivery, and content culturally acceptable to participants and program facilitators and coaches?
>
> Question 3: What were any potential barriers to participation/implementation in the program and what are suggestions to overcome these barriers?

Focus groups will be captured on digital recorders with written notes as a backup. All equipment will be tested onsite prior to use. Focus groups will last approximately 120 minutes and occur in a quiet room a local community center. Data will be transcribed verbatim and will be translated into English and verbally back-translated into Filipino to verify accuracy.

C.4 Outcome measurements

Data collection during the study will take place at two assessment points: baseline and immediate post-intervention. See Tables 1 and 2 for summaries of feasibility study outcome and process evaluation measurements for parents/caregivers and adolescents. All measures will be translated into Filipino and back-translated to check the accuracy of the translation.

a) Adult questionnaire

Sociodemographic information

1. Basic caregiver and child demographic information

Basic caregiver and child demographic information will be asked using items from the UNICEF Multiple Indicators Cluster Survey (MICS) Household Survey [15]. The MICS was developed to monitor the situation of children and women on a global level and is based on Demographic and Health Surveys. It has been used widely throughout low- and middle-income countries (LMIC) including the Philippines. It assesses caregiver/child age, gender, marital status, employment status, education level, basic literacy, child's relationship to caregiver, presence of child's biological parents (including reasons for absence), and other household members' age, gender and relationship to caregiver. It also assesses other household socio-demographic characteristics including household structure, family employment, and whether or not the family receives any government grants

2. Household ladder

Subjective social status will be assessed using the MacArthur Subjective Scale of Social Status [16]. Parents will be presented with a drawing of a 10-rung ladder that represents where people stand in their communities. The top of the ladder represents the highest standing in the community and the bottom ladder represents the lowest standing in their community. They will be asked to place themselves on the rung that best represents where they see themselves relative to their community (1 = lowest standing to 10 = highest standing).

3. Household income

An objective measure of household income will be obtained by asking parents to indicate their family's monthly income using a 16-item scale with income ranges from 0 = no income to 15 = over 83,000 monthly [17].

4. Adult history of child maltreatment

Parental/primary caregiver history of experiencing child maltreatment will be measured through an adapted version of the International Society for the Prevention of Child Abuse and Neglect (ISPCAN) Child Abuse Screening Tools Retrospective version (ICAST-R) (4 items) [18]. This scale utilizes parental self-reports of maltreatment experiences during childhood (until age 18 years) to assess the history of ever experiencing physical abuse (e.g., "When you were growing up (before age 18), did your caregiver ever discipline or punish you physically by hitting, spanking, slapping, kicking, or shaking you?") (2 items) and emotional abuse (e.g., When you were growing up (before age 18) did any person ever discipline or punish you by insulting or criticizing you, to make you feel that you were bad, stupid or worthless?") (2 items). In this study, ever experiencing physical abuse or emotional abuse will be assessed on a frequency scale (0 = Never; 1 = Once or twice; 2 = 3-5 times; 3 = More than 5 times). There will also be a dichotomous score for overall indication of previous child abuse (0 = no abuse; 1 = previous abuse).

5. Basic necessities

Relative poverty will be assessed using the Basic Necessities Scale (8 items). Developed by the Center for South African Social Policy in the 'Indicators of Poverty and Social Exclusion Project,' the Basic Necessities Scale measures levels of economic deprivation by identifying basic household items that families are unable to afford [19]. These include food, toiletries, clothes, shoes, and school uniforms, equipment, and fees. Items are coded dichotomously for positive or negative responses and summed to create an overall score. Previous

studies in have used this scale in multiple settings with high internal reliability (α = .84) [20, 21]. This information will be collected from both children and caregivers.

6. Household hunger

Relative poverty will be based on household hunger assessed using the Hunger Scale Questionnaire [22]. This scale examines food shortage and hunger in the household. Parents respond positively or negatively regarding the occurrence of hunger in the household, whether it occurred during the past 30 days, and if so, whether it occurred more than 5 times in the past 30 days (e.g., "the household has run out of money to buy food"). The scale produces scores for single occurrence and intensity of hunger

7. Food consumption

This study will also assess food consumption via caregiver report on average meals consumed per day in the past week for both the caregiver and child participants based on items from the UNICEF MICS Household Survey (e.g., "how many meals did you consumer per day") [15].

8. *Caregiver alcohol use*

Parental dependency on alcohol will be assessed by assessing alcohol consumption during the past month (1 item). Dependency is based on 3 or more drinks per day for female participants and 5 or more per day for male participants [23]. Due to the sensitive nature of these items, additional items dealing with other activities to reduce stress have been included in this section to encourage accuracy (3 items; e.g., "In the past month, have you been for a walk or done some other exercise to help you relax?").

Primary outcome

9. Child maltreatment – physical and emotional abuse

Physical abuse (including abusive discipline) and emotional abuse will be measured using 21 items from an adapted and expanded version of the ISPCAN Child Abuse Screening Tool-Trial Parent version (ICAST-TP) [24]. The ICAST-T is an adaptation of the multi-national and consensus-based survey instrument ICAST-Parent version (ICAST-P) [25], and has been used successfully in low and middle-income countries, including recently in the Philippines. The ICAST-TP measures parental reports of the incidence of abuse perpetrated against their child over the past month using a frequency score on a scale of 0 to 7, or 8 or more times (e.g., "In the past 4 weeks, how often did you discipline [Child Nickname] by pushing, grabbing, or kicking him/her?"). This study will assess incidence of child maltreatment for physical abuse (13 items), emotional abuse (8 items), as well as an overall indication of previous child abuse (0 = no abuse; 1 = previous abuse). We will also assess frequency of overall abuse by summing all of the subscales as well as for each individual subscale.

Proximal outcomes

10. *Positive parenting – MaPa Kids Only*

Positive parenting behavior will be assessed using the Parenting of Young Children Scale (PARYC, 21 items) [26]. The PARYC measures the frequency of parent behavior over the previous month. Items are summed to create total frequency scores for positive parenting (7 items, e.g., "how often do you play with your child"), setting limits (7 items, e.g., "how often do you stick to your rules and not change your mind") and proactive parenting (7 items, e.g., "how often do you explain what you want your child to do in clear and simple ways"). It has been used with strong reliability in previous studies on the program adapted in this study [27].

11. Dysfunctional parenting – MaPa Kids Only

Dysfunctional parenting behavior will be assessed using the Parenting Scale (PS, 30 items) [28]. This scale examines parent attitudes and beliefs regarding discipline practice. Responses are based on a 7-point Likert scale in which parents are presented with a situation and then are asked to choose between two alternative responses to a situation (1 = most effective; 7 = most ineffective; i.e., situation: "When I say my child can't do something;" response, score = 1: "I stick to what I said;" or response score = 7: "I let my child do it anyway"). Items are summed to create an overall score as well as for three subscales: Laxness, Over-reactivity, and

Verbosity. The PS has been used widely to assess the effectiveness of parenting programs, including in low-resource settings such as Panama [29].

12. Parenting behavior – MaPa Teens Only

Parent-child interaction, will be measured using two subscales from the Alabama Parenting Questionnaire – Adult Report [30]: Positive Parenting (8 items, e.g., "you praise your child if s/he behaves well") and Parent Involvement (8 items, e.g., "you take your child to a special activity"). The APQ has been shown to have moderate to strong internal reliability for both parent and child reports (Cronbach's $\alpha = 0.50$ to 0.89). It has been used widely including in LMIC such as South Africa and Mexico [31-36]. Caregivers report on the frequency of parenting behavior based on a 4-point Likert scale (0 = never; 3 = Often, more than 5 times). Items are summed to create total frequency score (range 0 to 24) as well as for each subscale.

13. Attitudes toward punishment

Attitudes toward punishment will be assessed using one item from the UNICEF Multiple Indicator Cluster Survey (MICS) 5 Child Discipline module[15], five items from the ISPCAN Child Abuse Screening Tool-Intervention Efficacy and Attitudes sub-scales (ICAST-I) [25]. The MICS item asks the parent/primary caregiver: "In order to bring up, raise up, or educate a child properly, the child needs to be physically punished." Parents/primary caregivers will report whether they disagree or agree with the statement based on a 5-point Likert scale of 0 to 4 (0 = Disagree strongly; 4 = Agree strongly). The two ICAST-I Efficacy items asks how often parents/primary caregivers reacted to child misbehavior over the past month using a frequency score on a scale of 1 to 7, or 8 or more times (e.g., "In the past 4 weeks, how often did physical discipline seem like the only option for stopping [Child Nickname's] bad behavior?" Four of the ICAST-T Attitudes sub-scale items assess parent/primary caregiver refer to a scenario in which a child is "always getting into trouble," and asks about the level of effectiveness of various disciplinary responses based on a 5-point Likert scale of 1 to 5 (1 = Very ineffective; 5 = Very effective). The fifth item on this sub-scale then asks "Which of the above approaches do you feel is most effective for disciplining children?" and permits a single select response.

Secondary Outcomes

14. Child neglect

Child neglect will be assessed using an adapted version of the ICAST-T Caregiver (mentioned above) Neglect subscale. This subscale has 3 items for assessing medical, physical, and educational neglect, including "In the past month, how often was [Child Nickname] not taken care of when sick or injured, even when you or another caregiver were able to do so and could afford it?" and "In the past month, how often was [Child Nickname] not given a meal that he or she needed, even when you or another caregiver was able to afford it?"

15. Parental monitoring – MaPa Teens Only

Parental monitoring and supervision practices will be measured using an adapted parental monitoring scale (11 items) [37]. The scale measures parents' solicitation of information (5 items, i.e., "How much do you try to know…") concerning their child's activities and friendships (e.g., who your child spends time with) using a 3-point scale (0 = I don't try; 2 = I try a lot); rule-setting, or how often parents set rules or limits about the same items using a 4-point scale (0 = Never; 3 = Always). Four items on monitoring of online activities were added to the solicitation of information subscale (e.g., "How much do you try to know the websites your child visits?) and two items were added to the rule-setting subscale (e.g., "How much do you set rules or limits on how long your child uses his or her device?), derived from the Global Kids Online survey [38]. Mean scores on the two subscales are standardized and summed to create a total parental monitoring score, with higher scores indicating higher levels of monitoring.

16. Child behavior problems

Child behavior problems will be measured using the Child and Adolescent Behavior Inventory (CABI) [39]. The CABI assesses a wide range of internalizing and externalizing symptoms in children and adolescents and is relatively shorter than the Child Behavior Check List (CBCL), making it a practical and reliable tool for measuring behavior problems. Parents report on their child's behavior during the past month (0 = Not True, 1 = Somewhat or Sometimes True, 1 = Very True). The irritability subscale (4 items, e.g., "has frequent mood

changes") and the externalizing subscale (10 items, e.g., often lies or cheats) will be used to measure problem behaviors. Items on both subscales are summed to create a total score of child behavior problems.

17. Three Problem Rating Scale – Parent Report

The Three Problem Scale-Parent Report was adapted from developed by Scott [40] to assess whether specific concerns or issues parents are having with their children change during an intervention. Parent-defined concerns from the Three Problem Scale that is consistent or larger than other measurements such as the Strengths and Difficulties Questionnaire or Child Behavior Check List. Parents are asked to identify up to 3 areas of concern with their children that are causing them the most distress and to rate each problem area from 1 to 10 (1 = "not a problem," 10 = "couldn't be worse"). The same problem areas are then asked at post-assessment to assess whether these problems have changed. Each single item will be analysed separately as well as a total problem rating score for the 3 issues.

18. Parenting efficacy

Parenting efficacy will be assessed using the Efficacy Subscale of the Parenting Sense of Competence Scale (8 items; PSOC-ES) [41]. The PSOC has been widely used in studies to evaluate parenting self-esteem, efficacy, or competence [41]. The PSOC Efficacy Subscale measures parental perception of competence, problem-solving ability, and capability in the parenting role (e.g., "I honestly believe I have all the skills necessary to be a good mother/father to my child"). Each item is rated on a 6-point scale that ranges from 1 (strongly disagree) to 6 (strongly agree). Items are summed to create a total score of parental self-efficacy.

19. Parenting stress

Parenting stress will be assessed using the Parental Stress Scale (PSS; 18 items) [42]. PSS has been widely used to measure parenting stress, including in LMIC, such as Pakistan [43] and China [44]. The scale has also been used with non-parent caregivers such as grandparents [45]. Caregivers report current positive attitudes (n = 8, e.g., "I feel close to my child") and negative attitudes (n = 10, e.g., "I feel overwhelmed by the responsibility of being a parent") related to parenting stress based on a five-point Likert scale (0 = strongly disagree; 4 = strongly agree). Positive items are reversed and then all items are summed to create a total parenting stress score (range 0 to 90).

20. Self-efficacy in managing emotions

Self-efficacy in managing emotions is measured using an adapted version of the Regulatory Emotional Self-Efficacy Scale [46]. The anger-irritation subscale includes four items. Parents indicate how well they think they can control their emotions (e.g., How well can you manage negative feelings when reprimanded by significant others?) using a 5-point scale (1 = Not well at all to 5 = Very well). Scores are averaged to create a total score for self-efficacy in managing emotions.

21. Parental depressive symptoms

Parental depressive symptoms will be measured using the adult version of Moods and Feelings Questionnaire (MFQ) [47, 48]. The scale includes 13 items indicating depressive symptoms (e.g., I didn't enjoy anything at all). Parents indicate how they have been feeling or acting in the past two weeks using a 3-point Likert scale (0 = Not true, 1 = Sometimes True, 2 = True). Items are summed to create a total parental depression score.

22. Child depressive symptoms

Child depressive symptoms will be measured using the short parent report version of Moods and Feelings Questionnaire (MFQ) [47, 48]. The scale includes 13 items indicating depressive symptoms (e.g., I didn't enjoy anything at all). Parents indicate how their child has been feeling or acting in the past two weeks using a 3-point Likert scale (0 = Not true, 1 = Sometimes True, 2 = True). Items are summed to create a total child depressive symptoms score.

23. Parent exposure to IPV and intimate partner coercion

Adult self-report of experiencing intimate partner violence over the past month will be assessed using an adapted version of the Revised Conflict Tactics Scale Short Form (CTS2S, 8 items) [49] and adapted items

from the WHO Multi-Country Study Questionnaire on Women's Health and Life and Domestic Violence against Women (WHO, 10 items) [50]. The CTS2S scale includes 2 items on the frequency of negotiation (e.g., "partner suggested a compromise for a disagreement"), 5 items on physical assault (e.g., "partner pushed, shoved, or slapped me"), and 1 item on psychological aggression (e.g., "partner insulted, shouted, yelled, or swore at me"). The adapted WHO questionnaire includes 7 items on coercion and emotional violence (e.g., "my partner tried to keep me from seeing my friends" and "my partner insisted on knowing where I was at all times"), and 3 items on restriction of financial autonomy (e.g., "my partner took my earnings or savings from me against my will.") All answers are coded on a frequency scale of 0 to 3 (0 = never happened; 1 = once or twice; 2 = 3-5 times; 3 = more than 5 times). Both the CTS2S and the WHO will determine an overall indication of intimate partner violence on a level of severity (sum of items across both scales) and prevalence (dichotomous variable indicating experience of conflict or not), as well as for each subscale.

24. Family functioning

Family functioning will be measured using the Burmese Family Functioning Scale [51]. It includes subscales on family cohesion (12 items, e.g., having understanding towards one another), family communication (7 items, e.g., expressing love through words or actions), and negative family interactions (4 items, e.g., having a lot of bad feelings in the family). Parents indicate how much they agree with statements regarding their family in the last four weeks on a 4-points Likert scale (1 = Almost Never True; 4 = Almost Always True). Items are summed to create total frequency score as well as for each subscale.

25. Parental Support for School Scale

Parental support for school will be measured using an adapted Parental Support for School Scale asking how often the parent engages in behaviours that support learning (6 items, e.g., talk to my child about the importance of finishing high school, support my child's schoolwork in any way that I can) using a 5-point Likert scale (1 = never, 2 = hardly ever, 3 = sometimes, 4 = most of the time, 5 = always). Higher scores reflect more parental support and value for school.

26. Educational aspirations and expectations

Educational aspirations will be measured one item asking the parent "How far would like your child to go in school?" rated on a 5-point scale (1 = finish some high school, 2 = graduate from high school, 3 = graduate from a 2-year college, 4 = graduate from a 4-year college, 5 = graduate from law, medical or graduate school).

Educational expectations will be measured using one item asking the parent how far they think their child will actually go in school rated on the same 5-point scale (1 = finish some high school, 2 = graduate from high school, 3 = graduate from a 2-year college, 4 = graduate from a 4-year college, 5 = graduate from law, medical or graduate school).

27. Community violence exposure

Exposure to community violence will be measured using 6 items from each of the subscales of the Violence Exposure Scale: Witnessing Community Violence and Experiencing Community Violence (12 items total) [52]. Parents report on whether they witnessed or experienced community violence in the past month based on 4-point Likert scale (0 = never; 4 = Often, more than 5 times). They also have the option to report that the incident has occurred before but not in the past month. Items from each subscale are summed to create a total subscale score as well as to create an incidence rating of exposure to community violence.

b) Child questionnaire

Sociodemographic information

1. *Basic child demographics*

Basic child demographic information will be asked using items from the UNICEF Multiple Indicators Cluster Survey (MICS) Household Survey [15]. The MICS was developed to monitor the situation of children and women on a global level and is based on Demographic and Health Surveys. It has been used widely throughout low- and middle-income countries (LMIC) including the Philippines. It assesses caregiver/child age, gender,

marital status, employment status, education level, basic literacy, child's relationship to caregiver, presence of child's biological parents (including reasons for absence), and other household members' age, gender and relationship to caregiver. It also assesses other household socio-demographic characteristics including household structure, family employment, and whether or not the family receives any government grants.

2. *Basic necessities*

Relative poverty will be assessed using the Basic Necessities Scale (8 items). Developed by the Center for South African Social Policy in the 'Indicators of Poverty and Social Exclusion Project,' the Basic Necessities Scale measures levels of economic deprivation by identifying basic household items that families are unable to afford [19]. These include food, toiletries, clothes, shoes, and school uniforms, equipment, and fees. Items are coded dichotomously for positive or negative responses and summed to create an overall score. Previous studies in have used this scale in multiple settings with high internal reliability (α = .84) [20, 21]. This information will be collected from both children and caregivers.

Primary outcome

3. *Child maltreatment – physical, emotional, and sexual abuse*

Physical abuse (including abusive discipline) and emotional abuse will be measured using 26 items from an adapted and expanded version of the ISPCAN Child Abuse Screening Tool-Trial Adolescent version (ICAST-TA) [24]. The ICAST-TA measures child reports of the incidence of abuse perpetrated against them over the past month using a frequency score on a scale of 0 to 7, or 8 or more times (e.g., "In the past 4 weeks, how often did your caregiver push, grab, or kick you?"). This study will assess incidence of child maltreatment for physical (10 items), emotional (10 items), and sexual abuse (6 items), as well as an overall indication of previous child abuse (0 = no abuse; 1 = previous abuse). We will also assess frequency of overall abuse by summing all of the subscales as well as for each individual subscale.

Proximal outcomes

4. Parenting behavior

Parent-child interaction, will be measured using two subscales from the Alabama Parenting Questionnaire – Child Report [30]: Positive Parenting (8 items, e.g., "your parent/caregiver praises you when you behave well") and Parent Involvement (8 items, e.g., "your parent/caregiver takes you child to a special activity"). The APQ has been shown to have moderate to strong internal reliability for both parent and child reports (Cronbach's α = 0.50 to 0.89). It has been used widely including in LMIC such as South Africa and Mexico [31-36]. Caregivers report on the frequency of parenting behavior based on a 4-point Likert scale (0 = never; 3 = Often, more than 5 times). Items are summed to create total frequency score (range 0 to 24) as well as for each subscale.

5. *Three Problem Rating Scale – Child Report*

The Three Problem Scale-Child Report was adapted from the Three Problem Scale for parents developed by Scott [40]. This scale measures whether specific concerns or issues parents are having with their children change during an intervention. Parent-defined concerns from the Three Problem Scale that is consistent or larger than other measurements such as the Strengths and Difficulties Questionnaire or Child Behavior Check List. This study adapted the parent-report scale for children to report on specific issues or problems they are having with their parents. Children are asked to identify up to 3 areas of conflict with their parents that are causing them the most distress and to rate each problem area from 1 to 10 (1 = "not a problem," 10 = "couldn't be worse"). The same problem areas are then asked at post-assessment to assess whether these problems have changed. Each single item will be analysed separately as well as a total problem rating score for the 3 issues.

6. Attitudes to punishment

Attitudes toward punishment will be assessed using one item from the UNICEF Multiple Indicator Cluster Survey (MICS) 5 Child Discipline module[15], five items from the ISPCAN Child Abuse Screening Tool-Intervention Efficacy and Attitudes sub-scales (ICAST-I)[25]. The MICS item asks the child: "In order to bring up, raise up, or educate a child properly, the child needs to be physically punished." Children will report whether they disagree or agree with the statement based on a 5-point Likert scale of 0 to 4 (0 = Disagree

strongly; 4 = Agree strongly). Four of the ICAST-T Attitudes sub-scale items refers to a scenario in which a child is "always getting into trouble," and asks about the level of effectiveness of various disciplinary responses based on a 5-point Likert scale of 1 to 5 (1 = Very ineffective; 5 = Very effective). The fifth item on this sub-scale then asks "Which of the above approaches do you feel is most effective for disciplining children?" and permits a single select response.

Secondary outcomes

7. Child behavior problems

Child behavior problems will be measured using the Child and Adolescent Behavior Inventory (CABI) [39]. The CABI assesses a wide range of internalizing and externalizing symptoms in children and adolescents and is relatively shorter than the Child Behavior Check List (CBCL), making it a practical and reliable tool for measuring behavior problems. Children report on their own behavior during the past month (0 = Not True, 1 = Somewhat or Sometimes True, 1 = Very True). The irritability subscale (4 items, e.g., "has frequent mood changes") and the externalizing subscale (10 items, e.g., often lies or cheats) will be used to measure problem behaviors. Items on both subscales are summed to create a total score of child behavior problems.

8. Child prosocial behavior

Child prosocial behavior will be measured using a Prosocial Behavior Scale [53]. Children report how often they engage in prosocial behaviors (16 items, e.g., "I share things I have with my friends") using a 5-point Likert Scale (0 = Never/Almost Never; 5 = Almost Always/Always True). Items are averaged to create a total score for child prosocial behavior.

9. Child risk behavior

Child risk behavior will be measured using a Risk Behavior Scale [54]. Children report how many times they have engaged in risky behaviors in the past month (9 items, e.g., "drinking beer or wine") using a using a frequency score on a scale of 0 to 7, or 8 or more times. Items are summed to create a total score for child risk behavior. This measure will be administered using audio-CASI to increase response rate.

10. Child depressive symptoms

Child depressive symptoms will be measured using the short child report version of Moods and Feelings Questionnaire (MFQ) [47, 48]. The scale includes 13 items indicating depressive symptoms (e.g., I didn't enjoy anything at all). Children indicate how they have been feeling or acting in the past two weeks using a 3-point Likert scale (0 = Not true, 1 = Sometimes True, 2 = True). Items are summed to create a total child depressive symptoms score.

11. Parental monitoring

Parental monitoring and supervision practices will be measured using an adapted parental monitoring scale (11 items) [37]. Children report on their parents' solicitation of information (5 items, i.e., "How much do your parents try to know…") concerning their child's activities and friendships (e.g., who you spend time with) using a 3-point scale (0 = They don't try; 2 = They try a lot); rule-setting, or how often their parents set rules or limits about the same items using a 4-point scale (0 = Never; 3 = Always). Four items on monitoring of online activities were added to the solicitation of information subscale (e.g., "How much do your parents try to know the websites you visit?) and two items were added to the rule-setting subscale (e.g., "How much do your parents set rules or limits on how long you use a device?), derived from the Global Kids Online survey [38]. Mean scores on the two subscales are standardized and summed to create a total parental monitoring score, with higher scores indicating higher levels of monitoring.

12. Neglect

Neglect will be assessed using an adapted version of the ICAST-T Child (mentioned above) Neglect subscale [24]. This subscale has 6 items for assessing medical, physical, and educational neglect, including "In the past month, how often were you not taken care of when sick or injured, even when your family was able to do so and could afford it?" Items are added to create a total neglect score as well as to create an incidence of neglect in the previous month.

13. Parental Support for School Scale (child-report)

Parental support for school will be measured using an adapted Parental Support for School Scale asking how often the parent engages in behaviours that support learning (6 items, e.g., talk to my child about the importance of finishing high school, support my child's schoolwork in any way that I can) using a 5-point Likert scale (1 = never, 2 = hardly ever, 3 = sometimes, 4 = most of the time, 5 = always). Higher scores reflect more parental support and value for school.

14. Educational aspirations and expectations

Educational aspirations will be measured one item asking the child "How far would you like to go in school?" rated on a 5-point scale (1 = finish some high school, 2 = graduate from high school, 3 = graduate from a 2-year college, 4 = graduate from a 4-year college, 5 = graduate from law, medical or graduate school).

Educational expectations will be measured using one item asking the child how far they think they will actually go in school rated on the same 5-point scale (1 = finish some high school, 2 = graduate from high school, 3 = graduate from a 2-year college, 4 = graduate from a 4-year college, 5 = graduate from law, medical or graduate school).

15. *Witnessing family violence*

The incidence and frequency of witnessing family violence will be assessed by using 2 items from the ICAST-T Child (e.g., "How many days in the past month were there arguments with adults shouting in your home?") [24]. Items will be summed to create a total frequency of witnessing family violence as well as dichotomized to create an incidence variable.

16. *Community violence exposure*

Exposure to community violence will be measured using 6 items from each of the subscales of the Violence Exposure Scale: Witnessing Community Violence and Experiencing Community Violence (12 items total) [52]. Children report on whether they witnessed or experienced community violence in the past month based on 4-point Likert scale (0 = never; 4 = Often, more than 5 times). They also have the option to report that the incident has occurred before but not in the past month. Items from each subscale are summed to create a total subscale score as well as to create an incidence rating of exposure to community violence.

17. Family functioning

Family functioning will be measured using the Burmese Family Functioning Scale [51]. It includes subscales on family cohesion (12 items, e.g., having understanding towards one another), family communication (7 items, e.g., expressing love through words or actions), and negative family interactions (4 items, e.g., having a lot of bad feelings in the family). Children indicate how much they agree with statements regarding their family in the last four weeks on a 4-points Likert scale (1 = Almost Never True; 4 = Almost Always True). Items are summed to create total frequency score as well as for each subscale.

Table 1. Summary of parent/caregiver-report outcome evaluation tools and questionnaires for MaPa Kids and Teens pilots

Outcome	MaPa Kids Pilot Measurement	Items	MaPa Teens Pilot Measurement	Items
Demographic factors (33/33)				
Parent, child and family demographics	General demographics questions [15]	16	General demographics questions [15]	16
Household ladder	MacArthur Subjective Scale of Social Status [16]	1	MacArthur Subjective Scale of Social Status [16]	1
Household income	Income measure [17]	1	Income measure [17]	1
Household assets	Basic Necessities Scale [18]	8	Basic Necessities Scale [18]	8
Household poverty	The Hunger Scale Questionnaire [19]	3	The Hunger Scale Questionnaire [19]	3
Parent history of child maltreatment	ISPCAN Child Abuse Screening Tool-Retrospective Version [20]	5	ISPCAN Child Abuse Screening Tool-Retrospective Version [20]	5
Parent alcohol abuse	Self-report on alcohol use	1	Self-report on alcohol use	1
Primary outcome (21/20)				
Child maltreatment	ISPCAN Child Abuse Screening Tool-Intervention – Physical and Emotional Abuse Subscales [21]	21	ISPCAN Child Abuse Screening Tool-Intervention – Physical and Emotional Abuse Subscales [21]	20
Proximal outcomes (57/22)				
Positive parenting	Parenting of Young Children Scale [22]	21	Alabama Parenting Questionnaire [23]	16
Dysfunctional parenting	The Parenting Scale [24]	30	N/A	
Attitudes to corporal punishment	Multiple Indicator Cluster Survey [15]	1	Multiple Indicator Cluster Survey [15]	1
Attitudes to corporal punishment	ICAST Attitudes Subscale [25]	5	ICAST Attitudes Subscale [25]	5
Secondary outcomes (117/144)				
Child neglect	ISPCAN Child Abuse Screening Tool-Intervention – Neglect– 3 items [21]	3	ISPCAN Child Abuse Screening Tool-Intervention – Neglect– 3 items [25]	3

Parental monitoring	N/A		Parental Monitoring Scale [26] Global Kids Online [27]	10 6
Child behaviour problems	Child and Adolescent Behaviour Inventory [28] Three Problem Rating Scale [29, 30]	14 3	Child and Adolescent Behaviour Inventory [28] Three Problem Rating Scale [29, 30]	14 3
Parenting efficacy	Parenting Sense of Competence, Efficacy Subscale [31] ICAST – Efficacy Subscale [21]	8 2	Parenting Sense of Competence, Efficacy Subscale [31] ICAST – Efficacy Subscale [21]	8
Parenting stress	Parenting Stress Scale [32]	18	Parenting Stress Scale [32]	18
Self-efficacy in managing emotions	Adapted from Regulatory Emotional Self-Efficacy Scale [33]	4	Adapted from Regulatory Emotional Self-Efficacy Scale [33]	4
Parental depression	Moods and Feelings Scale - Adult [34, 35]	13	Moods and Feelings Scale - Adult [34, 35]	13
Adolescent depression	N/A		Moods and Feelings Scale - Child [34, 35]	13
Parent exposure to IPV	Revised Conflict Tactics Scale Short Form [36]	8	Revised Conflict Tactics Scale Short Form [36]	8
Intimate partner coercion	WHO Multi-Country Questionnaire on Women's Health and Domestic Violence against Women (WHO) [37]	10	WHO Multi-Country Questionnaire on Women's Health and Domestic Violence against Women (WHO)	10
Family functioning	Burmese Family Functioning Scale [38]	23	Burmese Family Functioning Scale [38]	23
Witnessing and exposure to community violence	Violence Exposure Scale [39]	12	Violence Exposure Scale [39]	12
Parent support of education			Parent Support of School Scale	6
Educational aspirations			Aspirations and expectations	2

Process evaluation outcomes				
Implementation fidelity	Facilitator check-lists	n/a	Facilitator check-lists	n/a
	Parenting for Lifelong Health Facilitator Assessment Tool (PLH-FAT) [40]	51	Parenting for Lifelong Health Facilitator Assessment Tool (PLH-FAT) [40]	51
Adherence/exposure/engagement	Attendance registration	n/a	Attendance registration	n/a
	Activity check-list (parent report)	n/a	Activity check-list (parent report)	n/a
Satisfaction/acceptability	Parent overall satisfaction	40	Parent overall satisfaction	40
Acceptability and feasibility	Parent FGDs	n/a	Parent FGDs	n/a

Table 2. Summary of child-report outcome evaluation tools and questionnaires MaPa Teens pilot

Outcome	Measurement	Items
Demographic factors (potential subgroups and moderators)		**18**
Child demographics	General demographics questions [15]	10
Household assets	Basic Necessities Scale [18]	8
Primary outcome		**20**
Child maltreatment (Physical and Emotional Abuse)	ISPCAN Child Abuse Screening Tool-Intervention [21]	20
Proximal outcomes		**25**
Parenting	Alabama Parenting Questionnaire	16
	Adapted from Three Problem Rating Scale [29] [30]	3
Attitudes to corporal punishment	Multiple Indicator Cluster Survey [15]	1
Attitudes to corporal punishment	ICAST Attitudes Subscale [25]	5
Secondary outcomes		**86**
Child behaviour problems	Child and Adolescent Behaviour Inventory [28]	13
Child prosocial behaviour	Prosocial Behaviour Scale [41]	16
Risk behaviour	Adapted from Risky Behaviour Scale [42]	9
Child mental health	Moods and Feelings Scale [34, 35]	13
Parental monitoring	Parental Monitoring Scale [26]	10
	Global Kids Online [27]	6
Neglect	ISPCAN Child Abuse Screening Tool-Intervention [21]	3
Family functioning	Burmese Family Functioning Scale [38]	23
Witnessing family violence	ISPCAN Child Abuse Screening Tool-Intervention [21]	2
Witnessing and exposure to community violence	Violence Exposure Scale [43]	12
Parent support of education	Parent Support of School Scale	6
Educational aspirations	Aspirations and expectations	2
Process evaluation outcomes		**20**
Satisfaction/acceptability	Teen overall satisfaction	20
Acceptability and feasibility	Teen FGDs	n/a

C.4 Process data collection

a) Quantitative process evaluation

1. Implementation fidelity

Implementation fidelity by program facilitators of the MaPa program will be measured using self-report checklists by program implementers will examine the extent to which core intervention components are delivered. These checklists will include specific activities for each session, such as home practice discussion and role-playing exercises. Then, in order to produce a basic level of fidelity, a ratio of program implementation to program design will be created for both self-report and observational scores [44]. According to Borrelli and colleagues, a standard of 80% program fidelity will be considered as "high treatment fidelity" [45].

2. Quality of delivery

Quality of delivery will be assessed using the Parenting for Lifelong Health Facilitator Assessment Tool (PLH-FAT) [40]. The PLH-FAT was developed by the study investigators and program developers to assess the proficiency of program delivery by facilitators as a prerequisite to certification. Seven standard behavior categories are grouped into two scales based on the core activities (23 items) and process skills (28 items) as outlined in the program manual [5]. Assessment of core activities includes quality of delivery during home activity review (14 items, e.g., "identify specific challenges when shared by at least one parent"), illustrated story discussion (7 items, e.g., "discuss possible solutions for negative stories"), and practicing skills (10 items, e.g., "debrief with the participants about experience and feelings"). Assessment of process skills includes modelling skills (6 items, e.g., "give positive, specific, and realistic instructions"), collaborative facilitation approach (5 items, e.g., "accept participant responses verbally by reflecting back what the participant says"), encouragement of participation (7 items, e.g., "participants appear comfortable and involved in session"), and leadership skills (6 items, e.g., "use open-ended questions during group discussions").

3. Program adherence

Program adherence will be assessed by examining rates of enrolment, attendance, dropout, completion, and engagement of home activities. Enrolment rates will be based on the ratio of those completing baseline assessments to those who attend at least one session. Mean attendance rates for enrolled participants will be determined based on the ratio of number of attended sessions to the total number of program sessions. Dropout rates for enrolled participants will be defined as the percentage of participants who fail to attend at least three consecutive sessions and do not attend any sessions at a later stage. Completion rates for the entire allocation group will be determined based on the number of enrolled participants who attend a cut-off threshold of at least 66% of the program [46].

4. Participant satisfaction

Participant satisfaction will be examined for families who attend at least one session of the MaPa program [47]. Mean and standard deviation scores will be reported for the overall participation satisfaction scale (40 items), as well as for subscales (i.e., whether the program fulfilled their expectations, acceptability of delivery and teaching methods, acceptability of theoretical parenting techniques, and evaluation of program facilitators). Responses will be based on a 5-point Likert scale (1 = Very unhelpful; 5 = Very helpful).

b) Qualitative process evaluation

5. Program acceptability and participation

The study will conduct qualitative focus group discussions (FGDs) with all parents/caregivers who participated in the program (n = 4 FGDs; 15 participants per FGD), adolescents who participate in the program (n = 2 FGDs; 15 participants per FGD), and facilitators and coaches who deliver the programs (n = 2 FGDs; 5 participants per FGD).

Qualitative data collection will occur after program delivery. All data collection instruments will be translated into local languages, and the translations checked by back-translation. Ateneo de Manila researchers assisted by trained research assistants will conduct all of the interviews and FGDs with parents, adolescents and service

providers in local languages. FGDs will last approximately 120 minutes and occur in local community centers. Interviews will last 60 minutes and occur onsite at the community centers or in participant homes as appropriate. Participants will be provided lunch and transportation to focus groups.

Interviews and focus groups will occur during the post-evaluation and examine the following themes:

1. Participants observed change in parent-child relationships at home during program;
2. Cultural acceptability and appropriateness of program materials, delivery, and key program components; and
3. Existing barriers to participation during sessions and engagement in home practice and other activities.

Focus groups and interviews will be captured on digital recorders with written notes as a backup. All equipment will be tested onsite prior to use. Data will be transcribed verbatim and will be translated into English and back-translated into local languages to verify accuracy.

C.5 Process data analysis

Qualitative process data analysis

Qualitative data analysis will explore views and experiences of participants and facilitators using a thematic analysis approach with open, axial, and selective coding. Two independent raters will assess and code data from focus group transcripts. The study will investigate emergent themes regarding 1) participants observed change in parenting practices and child behavior at home during program; 2) acceptability and appropriateness of program materials, delivery, and key program components; and 3) existing barriers to participation during sessions and engagement in home practice and other activities.

Quantitative process data analysis

Quantitative process data will also be analyzed to assess: 1) implementation fidelity, recruitment, and retention; and 2) cultural and contextual relevance, including acceptability and satisfaction. These assessments will rely on self-reports by parents/primary caregivers, weekly facilitator self-report checklists, program session attendance registers, and data from outcome evaluation questionnaires. Analyses will be conducted using SPSS 25.0 or STATA 14.2.

Implementation fidelity will be analyzed by examining data from the facilitator self-report checklists. In order to produce a basic level of fidelity, a ratio of program implementation to program design will be created for both self-report and observational scores [44]. According to Borrelli and colleagues, a standard of 80% program fidelity will be considered as “high treatment fidelity” [48]

Program adherence will be measured through assessments of rates of enrolment, attendance, dropout, completion, and home activity engagement. The enrolment rate will be determined by calculating the ratio of those parents/primary caregivers who have been successfully recruited into participating in the feasibility pilot, and those who attend a minimum of one session. Mean attendance rates for enrolled participants will be calculated based on the ratio of the number of attended sessions to eight sessions (the total number of program sessions). Dropout rates for enrolled participants will be calculated as the percentage of participants who do not attend at least three consecutive sessions and fail to attend any further sessions. Completion rates will be determined based on the number of enrolled participants who participate in at least 75% of the program (six sessions or more).

Participant satisfaction will be assessed using a satisfaction questionnaire administered post-intervention to parents/primary caregivers who have attended at least one session. Mean and standard deviation scores will be calculated for the overall participation satisfaction scale (45 items), as well as for the different subscales, including: fulfilment of participant expectations, impact on the parent/primary caregiver-child relationship, acceptability of delivery and teaching methods, acceptability of content themes and parenting techniques, and evaluation of program facilitators.

C.6 Outcome data analysis

Quantitative data analyses of the effects of the intervention will be in line with the procedures adopted by the PLH-Philippines study in Taguig, Philippines (ClinicalTrials.gov Identifier NCT03205449). Analyses will be conducted using SPSS 25.0 or STATA 14.2 in order to respond to the two main objectives of the feasibility pilot. For multiple imputation and pooling of estimates, the R package 'MICE' will be used, potentially together with Microsoft Excel to define inputs to the R imputation functions. Mixed models will be fitted using the R package 'nlme'.

Given that this feasibility pilot is a single-arm study and is not designed to test hypotheses about effectiveness, estimates of effects on primary and secondary outcomes will be reported as estimates with 95% confidence intervals and without P values, in line with CONSORT guidance [49]. All intervention effects will be analyzed using an intention-to-treat design, which includes all participants recruited into the study regardless of whether or not they complete the program or complete all assessment points [50]. Multiple imputations will be conducted at the item level in order to account for missing data [51], while Little's Missing Completely at Random (MCAR) test with an expectation maximization algorithm will be used for the purpose of assessing response rates and the randomness of missing data [52]. The Multiple Imputation by Chained Equations (MICE) method will also be used with a fully conditional specification and a Markov Chain Monte Carlo (MCMC) algorithm with 10 maximum iterations; based on the assumption that the tolerance for a preventable fall in power will be low, a total of 20 multiple imputed datasets will be used in this process [53]. Furthermore, in order to prevent a violation of the assumption of independence underlying common statistical methods, data analysis will consider within-group correlation by calculating intra-cluster correlation coefficients within the sample to measure the level of dependency of outcomes for members in the six parenting groups [51].

Effects on primary and secondary outcomes will be analyzed using the instruments described in Table 2. Internal consistency with means and standard deviation will be calculated for each measurement at baseline and immediate post-intervention. However, given that it is affected by the number of scale items, item inter-correlations, and dimensions within scales, it will be noted that Cronbach's alpha will only provide a limited assessment of reliability [54].

Paired t-tests and chi-square crosstab analyses will be used to compare differences between means for all outcome measures. Pearson's *r*, means, standard deviations, standard errors, confidence intervals, and Cohen's *d* will be reported.

Construction of outcome variables

A select number of composite scores will be constructed and analyzed, with each score consisting of the sum of measurements for several individual items that are either all continuous (Likert) scales or dichotomous responses. Some items will be reverse coded prior to summation, as appropriate. These scores will provide the primary and secondary outcomes for this study, with each outcome comprising a baseline measurement collected during the baseline interview, and a post-intervention measurement collected one month following the end of the intervention. In the primary analysis, each outcome will be separately assessed. In order to test reliability, Cronbach's alpha will be calculated for each score, with scores constructed from individual items within the data imputation process (described below).

Preliminary data exploration

Using descriptive statistics, baseline family, household, socioeconomic, and outcome variables will be described. Summaries of baseline characteristics will be reported for each group, along with data variability and average values for continuous variables [55]. For normally distributed outcome data, continuous variables will be summarized by observed mean and standard deviation; for data with skewed distributions, continuous variables will be described using the median and inter-quartile ranges. Categorical variables will be described using frequencies and proportions. Characteristics will be summarized and stratified according to whether subjects were lost to follow-up.

Treatment of missing data

Multiple imputations will be used at the outset to create multiple, plausible 'complete' datasets, which are then analyzed using a likelihood-based imputation model (described below). The pooling of the final results (the

estimated effect sizes) from the multiple 'complete' datasets will account for the imputation process (discussed below under 'Data analysis: Estimation of effect sizes.')

The multiple imputation of missing data will follow these steps: 1) inspection of the missing data; 2) application of the imputation model to create multiple 'complete' datasets; 3) Assessing convergence of the imputation model and checking the imputations. In addition, two sensitivity analyses will be conducted.

Inspection of data

The patterns of missing data will be examined and reported, including calculations of the frequencies of different patterns of missing data, and the plotting of the distributions of key variables stratified by whether other variables are missing.

Imputation model

The covariates to be included in the analysis are arm (control or intervention), sex (child and parent/primary caregiver), and age (child and parent/primary caregiver). If any outcome data is missing, the MICE method, also known as Fully Conditional Specification (FCS), will be used for imputation. MICE is used under the assumption that the variables used in the imputation procedure are Missing at Random (MAR), which means that the probability of observations being missing depends only on the observed data, but not on the underlying values of the incomplete variable [56]. MAR allows values for missing observations to be predicted using patterns found in the observed observations. A series of regression models will be run in which each variable with missing data is modelled conditional on other variables in the data; each variable is modelled according to its distribution, with binary variables modelled using logistic regression [51] and Likert-scale variables modelled using predictive mean matching (PMM) [57]. This is an iterative process that uses MCMC techniques.

The imputation model will be able to address one or more of the following reasons for which a subject may have missing outcome data:

1. The subject was lost to follow-up, resulting in no post-intervention measurements;
2. The subject did not provide all data on all items to construct a specific score (at baseline, post-intervention, or follow-up); and
3. The subject did not provide data for a subset of items that contribute to a specific score (at baseline, post-intervention, or follow-up).

For the purpose of ensuring that the imputation model does not change when follow-up data is analysed, only data collected at the same time point or at earlier time points will be used to impute missing values. When imputing values for a particular variable, the following predictors will be used:

- Child sex (male or female);
- Parent/primary caregiver sex (male or female);
- Module (2-9 years, 10-17);
- All other items used to construct the specific score at the same time point (baseline, post-intervention, or follow-up);
- All items used to construct the same specific score at baseline, if there is an earlier time point;
- Other scores falling into the corresponding group of scores, at that time point and any earlier time points; and
- All remaining primary and secondary scores at that time point and any earlier time points.

In the case of dichotomous variables, a dichotomous predictor will enter the imputation model as a dichotomous categorical variable, while Likert-scale predictors will enter as continuous variables. Flat-file

imputation will be applied to the 2 assessment points during the feasibility pilot. For flat-file imputation, a particular item/score which is measured at each time point will be included as separate variables, rather than as measurements of the same variable. If there is evidence of a group effect, the approach for including Group will be explored during the implementation of the imputation model (as a fixed effect categorical variable, random effect in a mixed effects regression model, or through data stratification). In the event that the size of the imputation problem becomes too large (which may exceed available computer resources or result in multicollinearity due to instabilities), the process will be simplified and the number of predictors will be reduced.

Self-reported and observed outcomes will be imputed separately. In addition, a random number generator will be used to initiate the imputation model.

Assessment of convergence and checking imputations

Plots of summary statistics (mean and standard deviation) of the imputed measurements by iteration number per variable will be used to identify model misspecifications and determine whether the number of iterations is sufficient. Plots of variable distributions, both observed and imputed, will also be generated.

Sensitivity analyses

In order to examine the sensitivity of results to the assumptions of the imputation model, two alternative imputation approaches will be conducted. The first will also adopt a MAR assumption, but will alter the underlying imputation model form. The second sensitivity analysis will extend the aforementioned multiple imputation model to encompass the assumption that data is Missing Not at Random (MNAR), meaning that the probability of observations being missing is systematically related to the hypothetical values that are missing [56].

For the main imputation model and each of the two alternative approaches, the effect size of interest and statistical uncertainty will be estimated.

Data analysis: Estimation of effect sizes

Given that the main focus of the pilot is to test the feasibility of the adapted parenting program content and study procedures in preparation for a RCT, and consists of a single-subject pre-post design with no control group, the estimation of differences between pre- and post-intervention scores will only be exploratory in nature.

Prior to fitting models to the data, the observed changes in measurements from baseline to post-intervention will be described. For each outcome, the average differences across subjects between post-intervention and baseline scores will be calculated with a 95% confidence interval and without P values [55].

Each of the outcomes will be analyzed in turn. The analysis using the substantive model will be applied to each of the imputed or 'complete' datasets; thus, the analysis is described below in terms of a single outcome and a single dataset. Despite the fact that each item in the dataset is categorical in nature, the composite score (the outcome to be analyzed) is constructed as the sum of a number of individual items. Therefore, the pseudo-continuous outcome will approximately follow a conventional parametric distribution and be normally distributed, allowing parametric regression to likely be used for data analysis. The distribution of the outcome will be visually assessed.

Paired-samples t-tests will be conducted to examine whether there are mean differences between pre- and post-tests. A paired-sample t-test is appropriate for comparing means when those means are from the same entities, and assumes a normal distribution or a curve that is bell shaped and symmetrical [58]. First, to correct the repeated-measures error bars in SPSS, the average scores for each participant and each outcome will be calculated, followed by the calculation of the grand mean of all scores. In order to equalize the means between participants, an adjustment factor will be calculated by subtracting each participant's mean score from the grand mean. Next, adjusted values for each variable will be calculated in order to eliminate any between-subjects differences.

Paired-samples t-tests will then be conducted, with the means, standard error, confidence interval, t test statistic, significance level, and Cohen's *d* reported for each outcome.

Pooled parameter estimates will be calculated by averaging the multiple imputed datasets. In secondary analyses (as discussed below), effects sizes by sex and age group will be reported. The appropriateness of the model assumptions and fit will be assessed using the plots of standardized residuals against the fitted values (and covariates) and histograms of the residuals. Model fit will be examined using one of the multiple imputed datasets per outcome.

For each outcome, analyses will be performed on multiple imputed or 'complete' datasets, which will result in multiple effect sizes estimates. Pooled effect size estimates and confidence intervals (which account for variability arising from imputing the missing data), will be calculated using Rubin's rules [59].

C.7 Data management

All non-electronic data including signed consent forms, transcripts from FGDs and quantitative paper questionnaires will be stored in a locked filing cabinet at the Bulatao Center at Ateneo de Manila University. Digital recordings of FGDs will be copied onto an external hard drive and also kept in this locked filing cabinet. All data collected on CASI tablets will be encrypted as soon as the questionnaire is finalized (i.e., completed by interviewer) and accessible only by senior research personnel. Access to general functioning of the tablets will be password protected, and each tablet will have a GPS tracking application installed and activated. This application will allow for remote deletion of all information on the tablet in case of theft. Tablets will be stored in a locked cabinet at the Bulatao Center office every week, which is under 24-hour security protection.

The study will use two methods to upload and store data at the end of each day of data collection. First, the CASI tablets will transmit encrypted data using a 256-bit encryption via wireless networks to the study's Open Data Kit server (www.opendatakit), which will be housed at a central server managed by the University of Oxford's Information Technology Services. Only senior research personnel will have access to this server (PIs, co-Is, and data management consultants). It has a robust security system with firewalls and frequent backing up of all data. Transmission will occur on a daily basis from the Bulatao Center site, which will have a reliable 3G wireless network (i.e., not dependent on electric power), and the local research manager in the Philippines will manually upload data from tablets onto a local server on a daily basis. This local server will also be password protected and serve as a backup of the central server system. Once data is uploaded onto the local server, it will be erased on a daily basis so that no information remains on the tablet.

Data cleaning will be conducted at the end of each data collection phase. Individual datasets from the baseline and post-test evaluations will be stored separately from the final merged dataset, so that data reference points are available in the data validation process. All electronic data, including transcripts of qualitative data as well as quantitative data, will be stored on at least two servers, each of which will be accessible by a password known only to members of the research team. Thus, data will be protected from both server failure and confidentiality breaches. Video recordings will be stored on a password-protected hard drive in a locked cabinet for a minimum of 5 years with access only to authorized members of the research team. Non-electronic data (for instance, signed consent forms) will be stored in a locked filing cabinet.

Non-anonymized paper data will be kept for 3 years in a locked storage cabinet at the Bulatao Center at the Ateneo de Manila University after which it will be shredded and disposed of by the lead researchers. Anonymized data will be stored using United Kingdom Data Archive standards (www.data-archive.ac.uk/create-manage/planning-for-sharing). Anonymizing data will include removal of direct identifiers (names, addresses, postcode information, telephone numbers or pictures) as well as indirect identifiers (information on location, occupation or any other information that could be linked to a public source). This will include removing or aggregating variables or reducing the precision or detailed textual meaning of a variable in the dataset. Access to this data will be controlled and require authorization from the research team for further use.

D. Consequences of participation/ethical considerations

D.1 Research ethics

Ethical approval

Ethical approval will be obtained from the respective research ethics committees and/or institutional review boards from the following institutions prior to inception of the study: University of Oxford Central University

Research Ethics Committee, the Ateneo de Manila University's Research Ethics Committee, and the University of Cape Town Department of Psychology's Research Ethics Committee.

Informed consent procedures for parents and adolescents

Trained interviewers supervised by the research team will conduct informed consent procedures at either community centers or in the family households. Information sheets will be read to participants in Filipino to prevent illiteracy from hindering a participant's understanding about the methods and purpose of the study. All participants should be capable of giving their own consent, and we will not include any participants that are deemed incompetent. Special care will be taken to ensure that all participants are fully aware of and understand the research.

The informed consent process involves presenting a detailed verbal and written description of the study as it is described on the printed information and consent forms. Staff will emphasize that participation is voluntary for adults, and that participants can refuse to participate in either the intervention or research and/or can discontinue participation at any time without penalty, including their ongoing participation in the Department of Social Welfare and Development's 4Ps program. Interviewers will ask whether any participants have experienced any coercion to take part in the study; those that describe any pressure or coercion to participate will be excluded, and any other household members will also be excluded. Potential subjects will receive an item-by-item reading of the consent form by the study interviewer.

Participants will be informed of the procedures for ensuring their confidentiality, including: the use of unique non-personally identifying ID numbers instead of names on research materials, the video recording of sessions (and permission to refuse being recorded with no penalty to participation in either the program, study, or 4Ps), and maintenance of electronic and non-electronic data in locked computer databases and in locked filing cabinets in locked rooms. Tracking and contact procedures for scheduling follow-up interviews will be explained. Participants will be reminded that they may receive information regarding health or social services without participating in the study. All participants will be given the contact numbers of the co-PIs to answer questions about the study or one's rights as a human subject, as well as a 24-hour site contact number. All consenting and assenting participants will be offered a copy of the informed consent form.

Participants may consent to participate only after having the information sheet been read to them and there has been an opportunity for questions. The research team will assure that they are fully informed about the study and have had the chance to ask any questions that might have arisen.

Participants will have the opportunity to consider consent for up to a week, before interviewers return for the baseline assessments. However, when we have previously offered participants periods of 24 hours to 1 week to consider consent, the vast majority have actively requested to participate immediately. In light of this, we propose that participants are offered the choice of whether to consent or refuse immediately, or to have 1-7 days to consider whether they choose to consent.

Informed consent procedures for facilitators and coaches will be similar to those described above for parents.

A random subset of informed consent procedures will be reviewed by senior research personnel.

Informed consent procedures for parenting experts

Informed consent for parenting experts will follow a similar procedure as above. Potential Working Group members may consent to participate only after being provided with a written copy of the information sheet in Tagalog or English according to their preference, and having an opportunity to ask questions prior to the first focus group discussion. Participants will have the opportunity to consider consent for up to a week, before the research team will contact them to ask about their interest in joining the Working Group. The participant and the co-Investigator, Joscon, will then sign the Participant Consent Form on the same date prior to the start of the first Working Group discussion. However, when senior members of the research team have previously offered study participants periods of 24 hours to one week to consider consent, the vast majority have actively requested to participate immediately. The research team will ensure that participants are fully informed about the study and have had the chance to ask any questions.

All participants will be informed that they have the right to decline to participate and to drop out at any time. They will also be informed that all the points made during Working Group discussions will remain confidential and only reviewed by the research team, unless information is shared that a child (person under aged 18) has been seriously harmed or is at risk of serious harm. In the event of such a disclosure, the research team will consult with one another and decide on an appropriate course of action in line with the 2003 Child Protection Act.

It will also be explained to participants that participation in the study will not be related to any future potential funding opportunities from the donor.

Prior to the submission of reports and manuscripts for publication, the research team will consult with each Working Group member as to whether his/her role in the project can be specifically acknowledged by name and affiliation. Members will have the option to remain anonymous

Confidentiality

All precautions will be taken to ensure confidentiality for all participants. Participants will be given an individual research identification number to assure that their personal names are not disclosed. Research assistants will not be able access information from the electronic tablets as soon as they have completed and finalized the questionnaires. The electronic tablets will be password protected with the capability to remotely erase all data stored on them in the event of theft. Electronic data will be stored on a secured, password-protected, and encrypted server accessible only the principal investigators. Finally, all assessment data will be anonymized prior to statistical analysis.

Other ethical issues

We recognize that answering long surveys can potentially be a burden to research participants as well as have an adverse effect on the accuracy of self-report data [60]. We have estimated that each visit will last approximately 90 minutes per respondent. At approximately halfway through the assessments, we will offer participants the opportunity to stretch and take a break from answering questions during which refreshments will be provided.

Computer Assisted Self-Interviewing (CASI) methods have also been shown to reduce respondent burden while improving the efficiency and accuracy of data collection. The CASI technology was shown to: have a high degree of acceptability to research assistants and participants; decrease data collection time; and be more accurate than paper questionnaires [61].

Ethical issues regarding staff

All staff will be experienced in working on community research projects in the same or similar communities to the one in our study. Nevertheless, it is important to secure the mental wellbeing of interviewers and other research staff as well as mitigate any potential harm caused by working on the study. Thus, it is one of our utmost priorities. We will ensure that all research assistants are trained in awareness and safety measures. Staff will not undertake assessments in any situation in which they feel uncomfortable or unsafe, and are encouraged to travel in pairs in areas that are less safe.

In addition, we are aware that research with vulnerable families can result in increased stress due to the disclosing of difficult personal information and the demands of meeting deliverables. We will also conduct weekly debriefing meetings with research personnel to discuss any potentially distressing events that may have occurred during data collection. Staff will also participate in one-on-one reviews at the end of each data collection wave to assess their performance and psychological needs. The local investigator, Dr Alampay, is trained in providing psychological support and counselling.

D.2 Potential of harm

The focus of this study is to strengthen and evaluate the impact of a community-based parenting program. We believe that the overall potential benefits of the research and intervention in reducing the risk of child maltreatment and improving family wellbeing will outweigh potential risks of harm against participants. Nevertheless, we must consider two levels of potential harm: participation in the study and participation in the intervention.

Potential harm from the study

It is important to consider potential harm from participation in the study, no matter how small. There is a possibility that the opportunity to discuss parenting and parent-child relationships during FGDs may prompt distress in the respondents [62].

This study has an obligation to plan for the possibility of participants becoming distressed during FGDs. All interviewers will be trained and experienced in working with vulnerable families. The research team will include a qualified doctor (Dr Madrid) skilled in paediatrics, child health, public health, and community health as well as qualified psychologists (Dr Alampay, Dr Jocson) who will be able to discuss or supervise discussion of any issues with participants following the interviews. If there is a need for a participant to access more extensive support (such as seeing a counsellor or attending a clinic) referrals will be made. This includes support from the Department of Social Welfare and Development, Ateneo Bulatao Center, UP-PGH Child Protection Unit, and Philippine Ambulatory Pediatric Association (Please see appendices for child protection protocol).

All research personnel will be trained in ethical procedures and protocols concerning research with human subjects. During the consent stage, we will inform all participants that everything said will be confidential unless it becomes clear that they are at risk of significant harm or of putting anyone else at risk of harm. We will also inform participants that they do not have to answer any questions that they feel uncomfortable with and that they can stop the interview at any time without any negative consequences. In the case of participants not completing the baseline assessment, they will be excluded from the study.

Disclosure of harm

There is the possibility that participants may disclose harsh parenting practices that would reflect potential abuse or neglect of children. This study recognizes that researchers have a responsibility towards children who may be at risk of abuse or neglect or any other risk of severe harm. It is to be noted that it is an ethical principle to provide help for children whom the research identifies as in need.

As a result, the following protocol is proposed to mitigate any potential harm to children or adults that might occur during the study:

1. If information is disclosed that suggests that any member of the household is at risk of significant harm, the researcher will discuss concerns with the respondent at the end of the interview;

2. If the household member at risk of harm is a child, the researcher will discuss with the parent the possibilities for referral to child welfare, health organizations, and other services;

3. If the harm is considered to be significant, the research staff will inform local child protection services via the DSWD;

4. If severe abuse is disclosed in data collection, children will be immediately referred to social or medical services and the participant will be automatically excluded from the study;

5. If the decision is made to act, the participant(s) will be informed and referral will be made (*Please see attached draft referral form)*;

6. All staff will also receive additional training from the research team on how to respond to these situations in alignment with the study's referral protocols;

7. Weekly supervision meetings with all field interviewers will allow discussion of issues that arise concerning harm to research subjects and children;

8. Finally, if we determine that respondents or their families have experienced significant harm as a result of participation in the research study (i.e., severe abuse, suicidality, intimate partner violence, or other potential psychological or physical injuries), we will cease further activities until these issues can be addressed adequately.

Consistent with the abovementioned protocols, the informed consent form will indicate that relevant information may have to be disclosed without respondents' consent if the participant or a household member poses a serious danger to the self or to others, or if there is evidence to suggest child abuse or neglect (See Appendices 11 and 12 for Child Protection Protocol and Referral Form).

There is the possibility that participants may disclose harsh parenting practices that would reflect potential or actual abuse or neglect of children. This study recognizes that researchers have a responsibility towards children who may be at risk of or experiencing abuse or neglect, or any other risk of severe harm. It is to be noted that it is an ethical principle to provide help for children whom the research identifies as in need.

All staff will also receive additional training from the research team on how to respond to these situations in alignment with the study's referral protocols. Weekly supervision meetings with all field interviewers will allow discussion of issues that arise concerning harm to research subjects and children. If the research team determines that respondents or their families have experienced significant harm as a result of participation in the research study (i.e., severe abuse, suicidality, intimate partner violence, or other potential psychological or physical harm), activities will be stopped until these issues can be adequately addressed.

D.3 Mitigating potential harm from the parenting modules

We have also considered the potential risk of harm from participating in the intervention, and will be monitoring this throughout the project [63]. There may be potential psychological harm as a result of participation in the parenting program as a result of bringing up difficult experiences in caregivers' own childhoods or confronting intimate partner violence at home. However, decades of research on parenting interventions, including many randomized trials in LMICs [64], have not shown any evidence of harm from these interventions with plenty of evidence of benefit for parents and children, and high parent satisfaction. As a result, this study will be explicitly looking at potential benefits and risks of the parenting programs on both children and parents. Our statistical analyses plan will use two-tailed tests for differences between groups to examine potential negative and positive intervention effects. We will also monitor research subjects to assess whether participation in the intervention is placing any individuals at potential risk of harm. In addition to post-test assessments, monitoring will occur at specific time points when we are also monitoring implementation fidelity.

We do not anticipate any direct harm as a result of withdrawing from the intervention. Participation in the program will be completely voluntary with no direct penalties for refusing to participate. Furthermore, there is no evidence of harm from termination of parenting interventions in numerous other trials, including a number of evaluations in other low-resource settings [64]. Finally, if there is any indication of significant harm from either intervention condition at post-test, we will cease implementation of the program until harm has been adequately addressed and the programs have been adapted accordingly.

D.4 Self-referral for participants

In addition to our rigorous referral procedure described above, we will provide a self-referral document detailing available services to all participants as part of the informed consent stage as suggested. This information will include services for family and child support, substance use, gender-based violence and rape, child abuse and protection, physical, mental, and sexual health, government financial support, and contact details for available helplines. Self-referral forms will also be available during subsequent assessments at the request of participants. We will also provide up-to-date information regarding self-referral services at the end of the final wave of data collection.

D.5 Benefits

Adult participants will be offered a gift card (PHP 500 or approximately £7 per participant) after baseline and again after post-intervention assessment. Child participants will receive a token worth PHP 200 (approximately £3 per participant) after baseline and again after post-intervention assessment. Refreshments (i.e., juice/soda, tea, sugar, and bread) will be provided during data collection. All participants will receive a certificate of completion at the end of the program. Any participant who chooses to leave the study early will also receive a certificate stating how many sessions s/he attended. Participants of focus group discussion will be offered refreshments (i.e., juice/soda, tea, sugar, and bread) during the sessions.

Facilitators and coaches will be provided with a daily fee as confirmed through a written contract with you. The daily fee for facilitators will be PHP 2,500 and PHP 3,000 per day for coaches. Participants will also be reimbursed for transportation costs to the parenting programme venue and focus group discussion.

D.6 Access to data

Only the Principal Investigators (Lachman and Alampay) and Co-Investigator (Jocson) will have access to raw datasets. Access to cleaned, anonymized datasets will be provided to the other investigators. Anonymized datasets will also be stored online with password protection.

References

1. UNICEF, *National Baseline Survey on Violence against Children: Philippines*. October 2016, Council for the Welfare of Children and UNICEF Philippines.
2. Mikton, C., Butchart, A., *Child maltreatment prevention: A systematic review of reviews.* Bulletin of the World Health Organization, 2009. **87**(5): p. 353-361.
3. Desai, C.C., J. Reece, and S. Shakespeare-Pellington, *The prevention of violence in childhood through parenting programmes: A global review.* Psychology, Health & Medicine, 2017.
4. Ward, C.L., et al., *Parenting for Lifelong Health: From South Africa to other low-and middle-income countries.* Early Childhood Matters: Responsive Parenting: A Strategy to Prevent Violence, 2014. **49**.
5. Lachman, J.M., et al., *Masayang Pamilya Para Sa Batang Pilipino Parenting Program: Facilitator manual, Second Edition*. 2017, Manila: Parenting for Lifelong Health.
6. Lachman, J.M., et al., *Integrating parent management training within a conditional cash transfer system for low-income families: Results from a small-scale randomized controlled trial in the Philippines.* forthcoming.
7. World Health Organization, *INSPIRE: Seven strategies for ending violence against children*. 2016, Geneva: World Health Organization.
8. Cluver, L., et al., *Parenting for Lifelong Health: A pragmatic cluster randomised controlled trial of a non-commercialised parenting programme for adolescents and their families in South Africa.* BMJ Global Health, 2018. **3**(1): p. e000539.
9. Wainberg, M.L., et al., *Targeted ethnography as a critical step to inform cultural adaptations of HIV prevention interventions for adults with severe mental illness.* Social Science and Medicine, 2007. **65**(2): p. 296-308.
10. Brislin, R.W., *Back-translation for cross-cultural research.* Journal Cross-Cultural Psychology, 1970. **1**(3): p. 185-216.
11. Phillips, A.E., et al., *A systematic review and meta-analysis of quantitative interviewing tools to investigate self-reported HIV and STI associated behaviours in low- and middle-income countries.* International Journal of Epidemiology, 2010. **39**(6): p. 1541-55.
12. Davies, M., *Using computer-assisted self-interviewing (CASI) questionnaires to facilitate consultation and participation with vulnerable young people.* Child Abuse Review, 2005. **14**(6): p. 389.
13. Connolly, P., *Children, assessments and computer-assisted interviewing.* Child Abuse Review, 2005. **14**(6): p. 407.
14. Aspland, H. and F. Gardner, *Observational measures of parent-child interaction: An introductory review.* Child and Adolescent Mental Health, 2003. **8**(3): p. 136-143.
15. UNICEF, *Multiple indicator cluster survey manual, 2005: Monitoring the situation of children and women*. 2005, New York: UNICEF.
16. Adler, N. and J. Stewart, *The MacArthur Scale of Subjective Social Status.* MacArthur Research Network on SES & Health. Retrieved from http://www.macses.ucsf.edu/Research/Psychosocial/subjective.php, 2007.
17. Jocson, R.M. and R. Cabello, *Resilience in low-income Filipino mothers exposed to community violence: Religiosity and familism as protective factors.* Psychology of Violence, 2018.
18. Wright, G., *Findings from the Indicators of Poverty and Social Exclusion Project: A profile of poverty using the socially perceived necessities approach. Key Report 7*. 2008, Department of Social Development: Pretoria.
19. Labadarios, D., et al., *National food consumption survey in children aged 1-9 years: South Africa 1999.* Forum Nutrition, 2003. **56**: p. 106-9.
20. Dunne, M.P., et al., *ISPCAN Child Abuse Screening Tools Retrospective version (ICAST-R): Delphi study and field testing in seven countries.* Child Abuse & Neglect, 2009. **33**(11): p. 815-825.

21. Meinck, F., et al., *Adaptation and psychometric properties of the ISPCAN Child Abuse Screening Tool for use in trials (ICAST-Trial) among South African adolescents and their primary caregivers.* Child Abuse and Neglect, 2018. **82**: p. 45-58.
22. McEachern, A., et al., *Parenting Young Children (PARYC): Validation of a self-report parenting measure.* Journal of Child and Family Studies, 2011. **21**(3).
23. Frick, P., *The Alabama Parenting Questionnaire.* University of Alabama, 1991.
24. Arnold, D.S., et al., *The Parenting Scale: a measure of dysfunctional parenting in discipline situations.* Psychological Assessment, 1993. **5**(2): p. 137.
25. Runyan, D.K., et al., *The development and piloting of the ISPCAN Child Abuse Screening Tool-Parent version (ICAST-P).* Child Abuse and Neglect, 2009. **33**(11): p. 826-32.
26. Conger, R.D., et al., *Economic stress, coercive family process, and developmental problems of adolescents.* Child development, 1994. **65**(2): p. 541-561.
27. Global Kids Online. *Global Kids Online: Child questionnaire*. 2018; Available from: http://globalkidsonline.net/tools/survey/.
28. Burns, G., et al., *Child and Adolescent Behavior Inventory—Parent Version 1.0*. 2015, Pullman: Autor.
29. Scott, S., et al., *Multicentre controlled trial of parenting groups for childhood antisocial behaviour in clinical practice.* BMJ, 2001. **323**: p. 194-197.
30. Scott, S., *Deciding whether interventions for antisocial behaviour work: Principles of outcome assessment, and practice in a multicentre trial.* European Child & Adolescent Psychiatry, 2001. **10**(1): p. 59-70.
31. Johnston, C. and E.J. Mash, *A measure of parenting satisfaction and efficacy.* Journal of Clinical Child Psychology, 2010. **18**(2): p. 167-175.
32. Berry, J.O. and W.H. Jones, *The Parental Stress Scale - Initial Psychometric Evidence.* Journal of Social and Personal Relationships, 1995. **12**(3): p. 463-472.
33. Caprara, G.V., et al., *Assessment of filial, parental, marital, and collective family efficacy beliefs.* European Journal of Psychological Assessment, 2004. **20**(4): p. 247-261.
34. Angold, A., et al., *Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents.* International journal of methods in psychiatric research, 1995.
35. Messer, S.C., et al., *Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents: Factor composition and structure across development.* International journal of methods in psychiatric research, 1995.
36. Straus, M.A. and E.M. Douglas, *A short form of the Revised Conflict Tactics Scales, and typologies for severity and mutuality.* Violence and Victims, 2004. **19**(5): p. 507-519.
37. World Health Organization, *WHO multi-country study on women's health and domestic violence against women: summary report of initial results on prevalence, health outcomes and women's responses.* 2005.
38. Annan, J., et al., *Improving mental health outcomes of Burmese migrant and displaced children in Thailand: A community-based randomized controlled trial of a parenting and family skills intervention.* Prevention Science, 2016: p. 1-11.
39. Ruchkin, V., R. Vermeiren, and M. Schwab-Stone, *The Social and Health Assessment (SAHA): Psychometric developmental summary*. 2004, New Haven: Yale University.
40. Lachman, J.M., et al., *Parenting for Lifelong Health Facilitator Assessment Tool*. 2017.
41. Pastorelli, C., et al., *Measuring emotional instability, prosocial behavior and aggression in pre-adolescents: a cross-national study.* Personality and Individual Differences, 1997. **23**(4): p. 691-703.
42. Benthin, A., P. Slovic, and H. Severson, *A psychometric study of adolescent risk perception.* Journal of Adolescence, 1993. **16**(2): p. 153-168.
43. Richters, J.E. and W. Saltzman, *Survey of children's exposure to community violence.* National Institute of Mental Health, 1990.

44. Moncher, F., Prinz, R., *Treatment fidelity in outcome studies.* Clinical psychology review, 1991. **11**: p. 247-266.
45. Borrelli, B., et al., *A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research.* J Consult Clin Psychol, 2005. **73**(5): p. 852-60.
46. Webster-Stratton, C. and K.C. Herman, *Disseminating Incredible Years Series early-intervention programs: Integrating and sustaining services between school and home.* Psychology in the Schools, 2010. **47**(1): p. 36-54.
47. Martinez, C.R. and J.M. Eddy, *Effects of culturally adapted parent management training on Latino youth behavioral health outcomes.* Journal of Consulting and Clinical Psychology, 2005. **73**(5): p. 841-851.
48. Borrelli, B., *The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials.* Journal of Public Health Dentistry, 2011. **71 Suppl 1**: p. S52-63.
49. Eldrirdge, S.M., et al., *CONSORT 2010 statement: extension to randomized pilot and feasibility trials.* BMJ, 2016. **355**(i5239).
50. Hollis, S. and F. Campbell, *What is meant by intention to treat analysis? Survey of published randomised controlled trials.* Bmj, 1999. **319**(7211): p. 670-4.
51. Azur, M.J., et al., *Multiple imputation by chained equations: What is it and how does it work?* International journal of methods in psychiatric research, 2011. **20**(1): p. 40-49.
52. Li, T.J., et al., *Standards should be applied in the prevention and handling of missing data for patient-centered outcomes research: A systematic review and expert consensus.* Journal of Clinical Epidemiology, 2014. **67**(1): p. 15-32.
53. Graham, J.W., A.E. Olchowski, and T.D. Gilreath, *How many imputations are really needed? Some practical clarifications of multiple imputation theory.* Prevention Science, 2007. **8**(3): p. 206-213.
54. Cortina, J.M., *What is coefficient alpha? An examination of theory and applications.* Journal of Applied Psychology, 1993. **78**: p. 98-104.
55. Moher, D., et al., *CONSORT 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials.* British Medical Journal, 2010. **340**: p. c869.
56. Baraldi, A.N. and C.K. Enders, *An introduction to modern missing data analysis.* Journal of School Psychology, 2010. **48**: p. 5-37.
57. Morris, T.P., I.R. White, and P. Royston, *Tuning multiple imputation by predictive mean matching and local residual draws.* BMC Medical Research Methodology, 2014. **14**(75): p. 1-13.
58. Field, A., *Discovering Statistics Using SPSS*. 2009, London: SAGE Publications.
59. Rubin, D.B., *Multiple imputation for nonresponse in survey*. 1987, New York: Wiley.
60. Morsbach, S.K. and R.J. Prinz, *Understanding and improving the validity of self-report of parenting.* Clinical Child and Family Psychology Review, 2006. **9**(1): p. 1-21.
61. Brewster-Lee, D., Lukonge, H., Kaaya, N., Kagaruki, L., Raphael, G., Miller, C., *Faster Reporting, Reducing Errors, and Creating a Community Record: Improving OVC Programming with Mobile Reporting*, Catholic-Relief-Services, Editor. 2010.
62. Alderson, P. and V. Morrow, *Ethics, social research and consulting with children and young people*. Children, Youth and Environments. 2004, London: Barnardo's.
63. Lorenc, T. and K. Oliver, *Adverse effects of public health interventions: A conceptual framework.* Journal of Epidemiology and Community Health, 2014. **68**: p. 288-290.
64. Knerr, W., F. Gardner, and L. Cluver, *Improving positive parenting skills and reducing harsh and abusive parenting in low- and middle-income countries: A systematic review.* Prevention Science, 2013. **14**(4): p. 352-363.