unicef
for every child

EVALUATION

REPORT 3

**Impact Feasibility Assessment of UNICEF Mental Health and Psychosocial Support interventions**

# Impact Evaluation Design, Demonstration Case

## Parenting for Child Development (P4CD) – Papua New Guinea

Please contact:
UNICEF
Evaluation Office
3 United Nations Plaza
New York, NY 10017, USA
Email: evalhelp@unicef.org
Website: www.unicef.org/evaluation/

This report was authored by Dr. Manasi Sharma and Dr. Yasemin Kisbu, with guidance from Eduard Bonet Porqueras, Beth Ann Plowman and Zlata Bruckauf, from the UNICEF HQ Evaluation Office.

Cover photo: © UNICEF/UN0385710/Simons
Design and layout: Elena Panetti
Copy-editing:

# Table of Contents

01

# Motivation and intervention description

Stage 3 of the Mental Health and Psychosocial Support (MHPSS) Impact Feasibility Assessment (IFA) proposes evaluation designs to serve as demonstration cases, showcasing how interventions implemented at various levels can be evaluated rigorously.

Specifically, the selected interventions for Stage 3 include one community-level intervention, one school-level intervention, and two individual-level interventions, of which one is a digital intervention. By offering these illustrative examples, the goal is to provide guidance for future impact evaluations in different countries with similar interventions. This demonstration case does not necessarily imply that an impact evaluation is being initiated; this will depend on continuity of the programme, will and opportunity.

The selected Parenting for Child Development (P4CD) intervention in Papua New Guinea (PNG) is an individual-level intervention. Approximately 41 per cent of PNG's population are children and adolescents aged 0–18 years, and estimates from the Global Burden of Disease 2019 study reveal that mental health conditions account for 10 per cent of the total burden of disease among children aged 10-19 years, and that 2.5 per cent of adolescents aged 15–19 years had a depressive disorder and 5.3 per cent an anxiety disorder in 2019.[1] Among younger adolescents (aged 10–14 years), conduct disorder (at 3.3 per cent) and anxiety disorder (at 4.5 per cent) were the most prevalent.1 In a 2010 regional survey conducted by UNICEF, self-reported general happiness and self-esteem among children and adolescents in PNG were relatively low when compared with other countries. Slightly more than a quarter (28 per cent) of children and adolescents reported feeling happy most of the time, while seven per cent reported they were never or almost never happy.[2] The COVID-19 pandemic has heightened the need for MHPSS due to the significant impact on education, social connectedness, family stressors, inequality and disruption of essential services.[3,4]

Additionally, violence against children (VAC) defined as any physical, mental, or sexual violence, injury or abuse, neglect or maltreatment, and exploitation of any person under 18 years of age, is a key issue of concern in PNG. The *Unseen, Unsafe: The Underinvestment in Ending Violence Against Children in the Pacific and Timor-Leste: Case Study of PNG* report showed that: (i) nearly 70 per cent of children reported feeling scared and in pain in their community; (ii) 27 per cent of parents/caregivers used physical punishment; and (iii) over 50 per cent of parents and caregivers reported calling their child lazy, stupid or something similar, with nearly two-thirds reporting

1 Vos, T., Lim, S. S., Abbafati, C., Abbas, K. M., Abbasi, M., Abbasifard, M., ... & Bhutta, Z. A. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*, 396(10258), 1204-1222.

2 World Health Organization Western Pacific Office. Health of adolescents in Papua New Guinea. Manila: WHO Regional Office for the Western Pacific, 2011.

3 UNICEF East Asia and Pacific Regional Office. A Snapshot: UNICEF's approach to mental health during COVID-19 in East Asia and the Pacific. Bangkok: EAPRO, 2020.

4 Sharma, M., Idele, P., Manzini, A., Aladro, C. P., Ipince, A., Olsson, G., ... & Anthony, D. (2021). Life in Lockdown: Child and Adolescent Mental Health and Well-Being in the Time of COVID-19. *UNICEF Office of Research-Innocenti*.

shouting, yelling or screaming at their child.[5] Similarly, the prevalence of domestic violence, and intimate partner violence (IPV) are of grave concern in PNG. For example, the Demographic and Health Survey (DHS) showed that 59 per cent of women (aged 15-49 years) experienced physical and/or sexual violence in their lifetime.[6]

PNG has made important efforts to address mental health, with a strong focus on reducing the impact of risk factors, such as gender-based violence, domestic violence and early pregnancy. The P4CD intervention was developed in 2016 in response to these issues and was designed for parents and caregivers with children aged 3-10 years. The programme improves parents' knowledge and awareness of children's development needs and helps them to learn alternatives to harsh and coercive styles of discipline. P4CD aims to bring about improvements in parenting and both family and parent-child relationships in early to middle childhood. The programme was first developed in a cycle of research, consultation, training, and an evaluation of the implementation of the programme undertaken by Menzies Health Research in 10 pilot communities in PNG. In 2016 and 2017, the programme was piloted in four additional provinces, specifically, the Western Highlands (WH), Jiwaka, Chimbu, and Madang, in partnership with seven different faith-based organizations (FBOs). In 2018-2019, the programme was expanded to National Capital District (NCD) and Morobé provinces. The programme, with a budget of US$1,141,140.64, has reached a total of 74,512 participants. It is in its final phase of implementation and will be scaled up.

The intervention has a six-session format, training facilitators on positive parenting of children aged 3-10 years, with the objectives of reducing: (1) violence against children, and (2) intimate partner violence. An initial formative evaluation (non-experimental) of the intervention was previously conducted to identify emerging good practices which could be replicated in the programme's future scale up. The evaluation found that the P4CD programme achieved most of its objectives by improving parents' knowledge and skills which has led to the reduction of violence, abuse, and neglect of children knowledge and skills which has subsequently led to the reduction of violence, abuse, and neglect of children. The evaluation also assessed the criteria of equity, human rights, and gender equality, and the recommendations were designed to strengthen the P4CD programme operations and inform its scale-up.

The programme is led by two Catholic Archdioceses and one Evangelical Lutheran Church who work with a coordinator in each province. There is also an assistant coordinator who provides support in monitoring programme progress and the skills development processes of both the programme volunteers and targeted parents. A parish priest identifies a team leader and community-based volunteers, or facilitators, who work directly with parents in their respective villages/parishes.

**children**

**70%**

feel scared and in pain in their community

**parents/caregivers**

**27%**

used physical punishment

**+50%**

calling their child lazy, stupid or something similar

5 Save the Children, World Vision (2019). Unseen, Unsafe: The Underinvestment in Ending Violence Against Children in the Pacific and Timor-Leste", Case study: Papua New Guinea, page: 6.

6 GoPNG, (2016). DHS, page 289. Violence in the DHS is measured in the following form: **Physical spousal violence**: push you, shake you, or throw something at you; slap you; twist your arm or pull your hair; punch you with his fist or with something that could hurt you; kick you, drag you, or beat you up; try to choke you or burn you on purpose; or threaten or attack you with a knife, gun, or any other weapon; **Sexual spousal violence**: physically force you to have sexual intercourse with him even when you did not want to, physically force you to perform any other sexual acts you did not want to, or force you with threats or in any other way to perform sexual acts you did not want to; **Emotional spousal violence**: say or do something to humiliate you in front of others, threaten to hurt or harm you or someone close to you, or insult you or make you feel bad about yourself.

Before conducting any workshops, the facilitators and team leaders collect demographic data on parents who may be eligible to participate. Once 15-20 male and female caregivers have been identified, parent workshops are delivered across six full-day workshops, one workshop per week, by trained facilitators and team leaders. The group format encourages hands-on learning and social support between parents through interactive learning methods. The workshops provide opportunities to explore and learn about positive parenting practices that give children the best start in life, while also highlighting risks and dangers to child development, including harsh parenting styles. The programme also encourages parents to share experiences and to learn from each other. Workshops one through three provide a background in knowledge and awareness of children's social and emotional development and learning from birth to early and middle childhood. Workshops four through six focus on how parents respond to child behaviour, can use positive discipline, and improve parent and family well-being.

Box 1

## How was the P4CD intervention in PNG selected?

The P4CD intervention was selected for Stage 3 of the Impact Feasibility Assessment via a standardized process consisting of: 1) an analysis of **country context** and 2) an **intervention mapping** of MHPSS programming supported by UNICEF in the 21 selected countries interviewed during Stage 2 of the MHPSS IFA.

Key considerations in the analysis of country context were aspects like the political interest and will from partners in generating evidence, prioritization of knowledge gaps, national data and evaluation capacity, as well as the state and burden of mental health in the country.

Key considerations in the intervention mapping include scale and scalability of programming, plans for future expansion, existence of previous impact evaluations, type of programming and humanitarian setting context.

The results of the assessment for PNG were as follows:

- **Country context:** The total burden of mental health in PNG is high, while the access and provision of MHPSS services is low. Modelled estimates from the 2019 Global Burden of Disease Study also indicated that the mortality rate due to self-harm in PNG for adolescents aged 10–14 years was 0.5 per 100,000 population and 3.1 per 100,000 population for those aged 15–19 years. Boys had an excess risk of suicide when compared with girls. In addition, the evidence gap for mental health interventions in PNG is high, specifically there were zero primary studies of MHPSS interventions from the UNICEF MHPSS Evidence and Gap Map.
- **Intervention mapping:** The P4CD intervention was considered a priority in PNG, as it met all assessment criteria, including those around scale, future expansion, key thematic topics for high knowledge gain, parenting intervention, target population of early and middle childhood years, and no existing impact evaluation.

The P4CD intervention is one of four interventions short-listed as a priority for Stage 3 (alongside Benin, Brazil, and Nepal) for which impact evaluation plans are being developed.

*For more information, see: Intervention mapping and analysis of country context for the Impact Feasibility Assessment of UNICEF Mental Health and Psychosocial Support interventions, Stage 2 (2024).*

02

# Impact pathways and key outcomes

The P4CD programme targets parents and caregivers (as the duty bearers) and aims to reduce child maltreatment and violence against children (rights holders) in home settings. It aims to reduce aggressive, hostile and verbally abusive interactions between parents and children.

The intervention logic in brief is that parent participation in the programme leads to improvements in (a) parental knowledge of child development, (b) reductions in harsh and neglectful parenting, (c) increases in positive parenting practices, and (d) improvements in parent-child and in family relationships, which then leads to reductions in child maltreatment and violence against children.

The Theory of Change (ToC) can be stated as follows:

**IF** there is an intervention designed to correct cognitive misattribution through:

1. The provision of parenting management skills and resource materials to parents ultimately decreasing negative parenting behaviours; and
2. The promotion of responsive parenting using play, thereby reinforcing prosocial behaviours,

**THEN** a reduction in aggressive, hostile, and verbally abusive interactions between parents and parents and children will follow.

Therefore, the key outcomes of the intervention were:

- Parents' violent behaviours against their children
- Parents' positive parenting behaviours
- Parental mental health (using an overall mental health and well-being measure)

These will be achieved through intermediate outcomes, including parental knowledge and skills for parenting and child development:

- Parental knowledge of child development
- Parents' skills of responsive parenting
- Parents' problem-solving skills
- Parents' communication skills
- Inter-parental violence

**reduce**

aggressive, hostile and verbally abusive interactions between parents and children

03

# Evaluation questions

The quantitative impact evaluation will seek to answer the following main evaluation questions:

- Does P4CD lead to a reduction in parents' violent behaviours against their children?
- Does P4CD lead to an improvement in parents' positive parenting behaviours?
- Does P4CD lead to an improvement in parent mental health and well-being?

In addition to these main evaluation questions, it will be important to determine the effect of the intervention on the intermediate outcomes to investigate the mechanisms suggested in the intervention logic:

- Does P4CD lead to an improvement in parental knowledge of child development?
- Does P4CD lead to an improvement in parents' skills of responsive parenting and positive disciplining?
- Does P4CD lead to an improvement in parents' skills to problem-solving and communication in the family?
- Does P4CD lead to a reduction in inter-partner violence?

In addition to answering the above questions, it will be important to investigate the "for whom does it work?" question by separately estimating effects for fathers and mothers and across child age groups (e.g., aged 3-5 years, aged 6-10 years) (*subgroup analysis*).

Furthermore, qualitative [YK1] [YK2] data can be collected to further understand how the intervention works and investigate questions regarding its implementation and relevance to the needs of the targeted population. Key research questions are as follows:

- Was P4CD delivered as planned in terms of adherence to content, coverage, frequency, duration and quality? For example, did the facilitators' delivery of sessions on parenting skills development follow best practices? (*Implementation fidelity*)
- How is the parents' participation in the programme? How can it be improved? (*Compliance*)
- What are the factors influencing the achievement/non-achievement of the programme outcomes? Are there internal (e.g., robustness of design) and external (e.g., operational context) factors contributing to the effectiveness of intervention? How? (*Effectiveness*)
- Would there be a more efficient way to deliver the sessions? (*Efficiency*)
- How do the separate sessions contribute to the improvement in the intermediate outcomes? (*Understanding impacts*)
- Is the programme suitable to the needs and characteristics of the targeted population? Are there emerging topics/issues to be addressed? (*Relevance*)

More detailed questions should be developed/revised in collaboration with the programmatic team and other stakeholders, following the further detailing of the ToC and finalization of evaluation design.

04

# Evaluation design

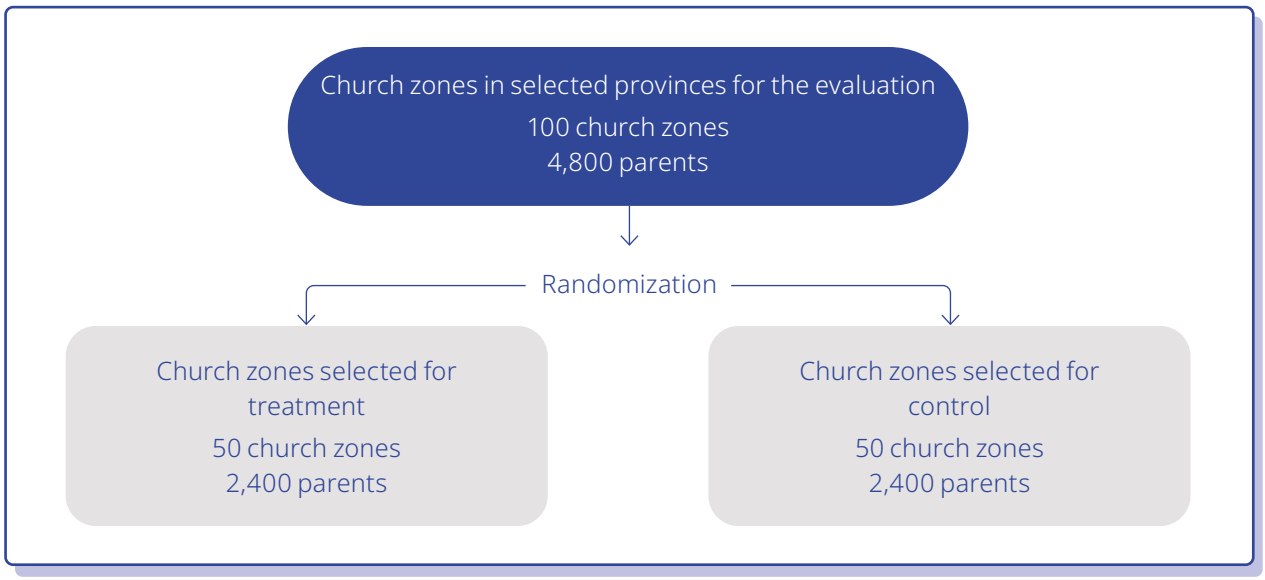
An initial formative evaluation (non-experimental) of the intervention was previously conducted to identify emerging good practices which could be replicated in the programme's future scale-up.

The suggested evaluation primarily aims to establish the causal effects, by testing whether the intervention causes any changes in the targeted outcomes. Two initial options for conducting a high-quality impact evaluation are suggested for the intervention. These options should be further discussed with the country office, especially to ensure how an impact evaluation can be integrated into the existing program with as little disruption as possible.

### A cluster randomized control trial (RCT) using church zones

A first viable option is to utilize the church zones through which the communities participating in the intervention sessions are reached, for setting up the trial. The church zones can be randomly assigned into treatment and control groups. This design will allow for identifying both the impact of the programme on fathers and mothers, and across different child age groups. This design will cover a minimum of 50 control church zones and 50 school zones with P4CD intervention. The total sample of fathers and mothers will need to be determined but will likely be 24 fathers and 24 mothers per church zone, amounting to a total of 4,800 parents.

**Figure 1.** A randomized control trial (RCT) at the church zone level

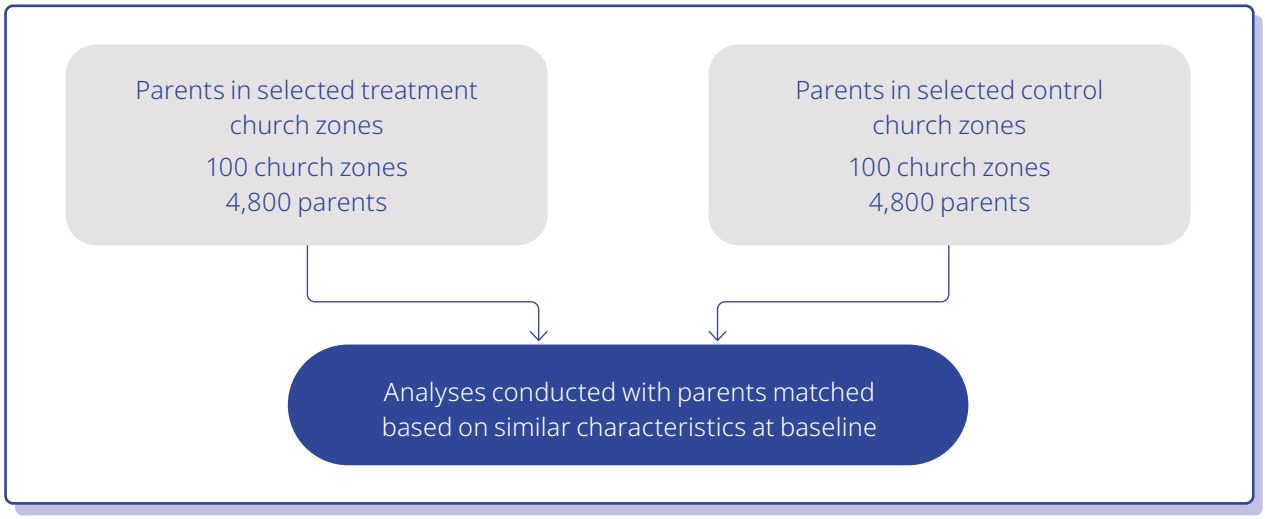


### A quasi-experimental, matched difference-in-differences design

In the case where it is determined that selection of some treatment and control zones cannot be done randomly, or if the timing does not allow for randomization, another option for the evaluation design is a quasi-experimental method. In this case, control church zones will be identified that have similar characteristics as the treatment church zones, especially regarding factors that may influence the targeted outcomes (e.g., socio-economic status of residents).

The sample size for this design will depend on several factors but needs to be more than the sample size required for an RCT. To demonstrate the need for a high sample size in a matched design, a sample size that is double of the sample size suggested in the RCT design is suggested, as depicted in Figure 2 (i.e., a total sample size of 4,800 parents for the matched design). As with the above design, ideally the baseline survey would include this needed sample size, with only those who best match kept into the study for subsequent follow-ups.

**Figure 2.** A quasi experimental, matched design




Several options exist for the design of the qualitative component of the impact evaluation, likely involving a range of data collection methods – semi-structured key informant interviews (KIIs) with programme implementers and other stakeholders; and focus group discussions (FGDs) with programme participants.

Qualitative data will be collected during a follow-up period to capture the perceived experiences with the programme for both the programmatic team and intervention participants. The qualitative follow-up can also take place after the preliminary analysis of quantitative impact evaluation data, to be used to probe findings from the quantitative study.

05

# Key outcome indicators

The primary outcome indicators to measure are directly related to the intervention ToC. The targeted outcome variables are harsh parenting behaviours and positive parenting behaviours.

But it is also important to measure the intermediate outcomes which are the parental knowledge of child development, parents' skills of responsive parenting and positive disciplining, parents' skills to problem-solving and communication in the family, and inter-partner violence in the family. These will be measured at the individual parent level (i.e., both mothers and fathers).

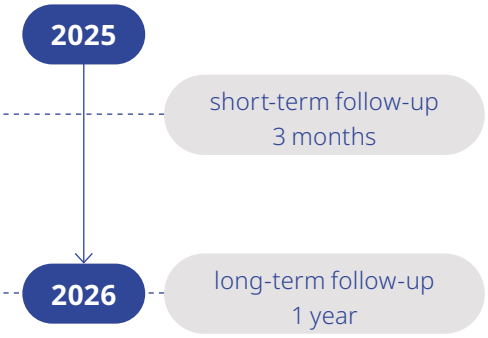
For each of these outcomes, culturally appropriate, valid and reliable measures should be included. Furthermore, in addition to child age-based measures of child development knowledge, it would be important to capture parenting practices that follow the age of the child. As the study requires longitudinal data collection and thus the age of children will change from one follow-up to the other, measurement invariance (i.e., whether the same construct is measured equally in different time points) across child age should be considered when selecting tools and tested for during data analyses. Since the key indicators align with the evaluation questions, target sample, and will be tailored to the unique evaluation plan, they will undergo additional refinement in a subsequent phase of the evaluation plan.

Furthermore, it is important to measure the uptake of treatment participants (i.e., the number of sessions attended recorded at individual participant level), that will enable additional estimations, such as the intervention effect estimate for parents who attended to most of the sessions (i.e., the treatment-on-the-treated effect) or the relationship between attendance rate and targeted outcomes (i.e., dosage effect).

06

# Timeline

Quantitative baseline data collection would ideally start in early 2025, with follow-up surveys (with the same parents at baseline) conducted during time periods yet to be determined (e.g., likely a short-term follow-up at three months and a longer-term follow-up at one year). Qualitative data collection would occur during the follow-up in the same period. The full timeline for the study would be approximately between 2025 and 2026.




07

# Data collection methods

The primary data collection methods for this study will be individual interviews which will include both quantitative and qualitative methods. Quantitative data will be collected longitudinally at three main points in time from parents. In both suggested designs, parents will be interviewed before the intervention begins and followed longitudinally over time. Appropriate survey questions and instruments, as well as ethical protocols, will be developed and implemented to ensure data collection follows ethical requirements.

Qualitative methods to be used include semi-structured KIIs with programme implementers and other stakeholders; and FGDs with programme participants.

08

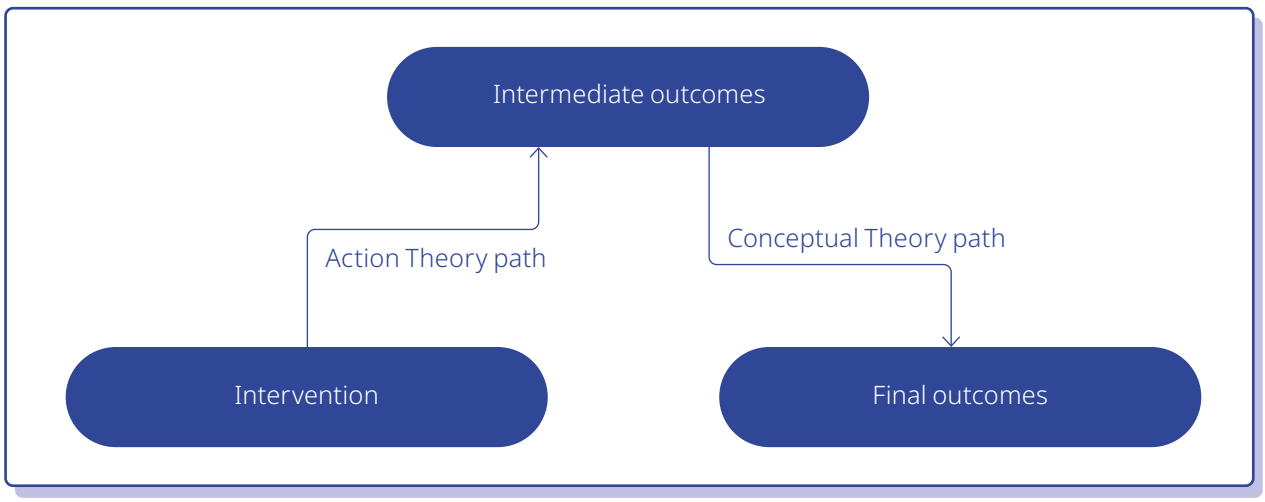
# Analysis methods

## Statistical methods should be employed to determine if there are differences between the treatment and comparison groups.

The most common statistical analysis used with a randomized design is regression analysis. The subgroup analyses (e.g., to estimate whether there are differential effects for fathers and mothers) will be employed through building interactions into the regression models. Clustering effects will be checked and appropriate statistical measures (e.g., robust standard error estimation) will be employed. Even though randomization is assumed to achieve balanced (i.e., similar) treatment and control groups at baseline, further analyses will be conducted to test potential baseline imbalance across groups at baseline; if groups are detected to be significantly different at certain characteristics at baseline, those characteristics will be controlled during the final analyses (i.e., through adding them as covariates into the regression model).

Furthermore, to quantitatively test the program ToC, statistical mediation analysis can be employed. Mediation analysis provides evidence of how a programme achieved results by testing hypothesized causal mechanisms and helps to understand how the programme can be refined. As illustrated in Figure 3, the analysis of the path on the left informs whether the programme activities succeeded/failed in changing the targeted intermediate outcomes (e.g., parenting skills), testing the so-called action theory. And the analysis of the path on the right informs whether the intermediate outcomes lead to changes in the final outcomes (e.g., parenting behaviours) as hypothesized in the intervention ToC, testing the so-called conceptual theory of the intervention. Furthermore, separate mediated effects are estimated for each pathway to get the relative effects on final outcomes through different intermediate outcomes (e.g., the effect of the intervention on changing harsh parenting behaviours through changing positive disciplining skills, versus, the effect of the intervention on changing harsh parenting behaviours through changing parental knowledge of child development).

**Figure 3.** Mediation model



In the case of the quasi-experimental matched design, statistical matching methods will be employed. It is important to recognize the limitation that a matched comparison group can only be balanced on observable characteristics, posing a risk of differences in unobservable characteristics and introducing bias into the overall estimates of the programme's impact. While collecting comprehensive data helps mitigate this risk, complete elimination is not feasible. Using a matched design requires high-quality data on observable characteristics for treatment and comparison groups, with for instance the propensity score matching (PSM) technique ensuring similarity. Two conditions, conditional independence and common support, are crucial for confidently estimating programme impact based on matched data. This involves including all baseline measurements of variables affecting the probability of treatment and selecting comparison groups closely matched to ensure compatibility within the same distributions. Conducting common support tests and balance tests is essential to verify the appropriateness of the matching process in creating a suitable comparison group.

Qualitative data will be analysed following an evaluation matrix, in which all evaluation questions are matched with the appropriate data source and data collection method. Regarding the evaluation questions on understanding how impact occurs, qualitative methods such as contribution analysis can be used.

09

# Estimated resources for data collection

Ideally, a professional survey firm should collect evaluation data. Costs associated with data collection usually cover the development of a study protocol, ethical review, creation of survey instruments (including translations), pre-testing, enumerator training, survey coding, data collection, data cleaning, and translation/transcription for qualitative data. Human resources often constitute the majority of data collection budgets. Quantitative data collection teams may comprise 20 to 40 enumerators, undergoing two weeks of training and pretesting. In contrast, qualitative data collection teams are typically smaller (six to eight interviewers), spending shorter periods collecting data (one to two months), but dedicating more time to transcribing and translating interview and focus group narratives. Furthermore, there may be security costs to be added for data collection in the programme implementation area. Estimations are as follows:

- **Quantitative data:** For a survey of 100 communities, with 48 parents per community (4,800 individuals in total) where all interviewing takes place in households = US$144,000 per wave = US$432,000 for three survey waves.
- **Qualitative data:** For data collection including KIIs with key stakeholders and FGDs with parents = US$40,000 for one data collection wave.

Thus, the estimated cost for a three-wave quantitative and one-wave qualitative data collection is around **US$472,000**. It is important to note that these costs are approximate and contingent on factors such as the geographic dispersion of communities/provinces (sample), the quality of the hired data collection firm, survey timing, and other variables. Adjustments in costs, either up or down, may be necessary based on the number of survey waves or study arms. The scope of each data collection could also be modified according to priorities or budgets, including additional components or questions. It is crucial to highlight that these costs exclude labour time for international or national principal investigators, as well as expenses for analysis, dissemination, or publication of study findings.

## Intervention website, documents and materials links

**Initial Evaluation Report**
https://www.unicef.org/png/media/1251/file/PNG-Reports-ParentingforChildDevelopment.pdf

**P4CD programme description:**
https://www.unicef.org/png/reports/parenting-child-development

**Burnett Institute report on MHPSS in PNG:**
https://www.unicef.org/eap/media/13021/file/UNICEF%20MHPSS%20in%20EAP%20region%20PNG%20report.pdf

October 2024