

---

## TERMS OF REFERENCES

### INSTITUTIONAL CONTRACT

#### EVALUATION OF UNICEF SUPPORTED COVID RELIEF CASH TRANSFER

---

### 1. BACKGROUND AND JUSTIFICATION:

Since the beginning of the pandemic caused by the virus COVID-19, governments throughout the world have been taking measures such as lockdown and quarantines to stop its spread. While helpful in preventing the virus spread, such measures have had immediate adverse effects on the global economy and socio-psychological wellbeing of people. It is estimated that in the early stages of the pandemic, 55 million domestic workers were already impacted by COVID-19 worldwide<sup>1</sup>. Job loss for many is inevitable as strict measures limiting human physical interaction and closing businesses, especially for those who depend on daily wages are prolonged.

In response to the pandemic, Nepal has also taken strict measures since March 2020. As in many countries, these measures caused social and economic hardships such as job loss, disruptions in public services and limited access to healthcare and education. These visible impacts of the pandemic entail a range of issues that children might encounter in their homes staying with perpetrators of violence, including issues such as food insecurity due to job loss of caregivers; lost education due to closure of schools and unavailability of caregivers to help them learn, and many other issues which can only become evident with thorough monitoring of their situation.

The likelihood of negative economic impacts by a protracted pandemic in the country is high. The World Bank<sup>2</sup> estimates that economic growth will fall to a range between 1.5 % and 2.8 % in FY2020 and it is likely the country will not experience significant economic growth in 2021. The country's economy, mostly reliant on tourism industry, remittances and agricultural activity, which have weakened from all angles, including declined imports and exports, is most likely to increase poverty among the population.

With the population of 29 million, Nepal is estimated to have 20.7 million people of the working age, and the employment rate is 11.4 %.<sup>3</sup> As ILO estimated in the last Labour Force Survey (LFS)<sup>4</sup>, 62.2 % of Nepali people work in the informal sector, 17.5 % in trade industry, 14 % in construction, 24 % in service and sales and 20 % in elementary occupations. Agriculture, as the biggest employment industry, provides a job to one in every five people in Nepal. The private sector and agricultural industry, including people depending on daily wages are hit the hardest by the pause of economic activities during lockdown in Nepal.

A significant proportion of migrant Nepalis has been contributing to the country's GDP – more than 8 billion remittances accounting for 28% of GDP were recorded in FY2018/2019 and women account for 8.5 % of the migration flows.<sup>5</sup> Most Nepali migrants work in India, and other popular countries are Qatar, Malaysia, Saudi Arabia, UAE and Kuwait. Most of an estimated 587,646 Nepalis in India's work in service sector and 86% of them depend on daily wages.<sup>6</sup> Due to economic setbacks and health concerns caused by the pandemic in many host countries, it is likely that hundreds of thousands of migrants will have to return to Nepal.

---

<sup>1</sup> *Impact of the COVID-19 on loss of jobs and hours among domestic workers*, International Labour Organization (Geneva, 2020).

<sup>2</sup> *Nepal at a glance*, World Bank (<https://www.worldbank.org/en/country/nepal/overview>)

<sup>3</sup> *Labour Force Survey (NLFS III)*, National Planning Commission, Central Bureau of Statistics, International Labour Organization (Nepal, 2018)

<sup>4</sup> *Ibid.*

<sup>5</sup> *Impact of COVID-19 on Nepali Migrant Workers: Protecting Nepali Migrant Workers during the Health and Economic Crisis*, International Labour Organization, (Nepal, 2020).

<sup>6</sup> *NLFS III*.

To understand the evolving situation of households with children during the pandemic, UNICEF initiated a multi-series nationwide telephonic survey under title, “Child & Family Tracker (CFT)”. The survey was carried out with over 7,000 households with children across Nepal, every 4-6 weeks since May 2020. In conducting this survey UNICEF’s main purpose was to increase visibility of children’s issues and ensure evidence-based response to their needs, in addition to assessing and raising awareness on social and economic impacts of the pandemic. The survey<sup>7</sup> found that in May 55% of households with children experienced income and livelihood loss. The percentage of households with income and livelihood loss increased to 51% in July and slowly began decreasing in August. As of October, 45% of households reported loss of income and livelihood loss.

Many households with earnings or livelihood loss became food insecure during lockdown. Households living in rural areas and almost half of low-income families became food insecure, with over 20% changing diets and reducing food intake of their children. A considerable increase is observed among households reporting children’s education as one of their top immediate needs at present. The other two are financial support and employment, and food for 30% of households remains as immediate need.

The survey results from different rounds support evidence coming from the households declaring children’s education as an immediate need. In the first two months of lockdown in Nepal, children in 95% of households stopped going to school and children only in 29% of the households studied at schools which offered distant learning (DL). Yet only half of children who studied at such schools could take advantage of an opportunity to continue their studies remotely. At present, children in 70% of the survey’s households are studying at home, and 22% of them are using DL. Nearly 40% of them encounter issues with the internet connectivity and electricity failures.

In response to growing needs of the households, UNICEF initiated a one-off cash transfer in the amount of NPR 2,000 per child and covering up to 2 children in each household. There is an additional and important objective of drawing up lessons for national replication and presenting alternate models of delivery of cash within both humanitarian and developmental social protection system. Regular social protection system channels funds manually, though bank transfers have been also increasing. There is a digital cash transfer policy as well now, which both MoHA and development partners have been challenged to implement. Mobile transfers are yet to be actioned by the policy makers, making Nepal a laggard in use of mobile money. An emphasis on shock responsive social protection necessitates setting up of flexible and expandable and fast modalities of cash transfer, a goal which has so far not been pursued within any national experiment. UNICEF’s supported Child and Family COVID Cash Transfer is aimed at addressing this gap in experimentation and knowledge.

This response is based on evidence coming from the CFT and robust analysis of the situation of children and their families. Since all CFT participants have children, the cash transfer of UNICEF does not entail any specific criterion, but the participants were asked if they would be willing to receive cash to meet the needs of their children. In total, 5,700 households from various income group, who experienced income/livelihood loss and are struggling for food and with increasing needs of their children, volunteered to become a recipient of a cash transfer.

The cash transfer is purely based on evidence of identified households who have been voluntarily participating in the CFT every 4-6 weeks since May 2020. This project is expected to alleviate financial burden of the households in meeting the needs of their children and improving their food security. The cash will be transferred to bank accounts and e-Sewa digital wallets of the households. The project is designed with expectation that its successful implementation will serve as an example for scaling up similar interventions and to provide immediate response to shocks caused by emergencies.

---

<sup>7</sup> *Child Family Tracker Series*, UNICEF, 2020 (CF)

UNICEF will evaluate the cash transfer results to understand how it works efficiently, assess the extent to which it helped children and their families and gather lessons learned. To evaluate the project UNICEF is seeking a qualified team of experts in evaluation of cash transfers and social protection intervention in development and emergency settings, including the pandemic contexts.

---

## 2. PURPOSE AND OBJECTIVES:

The purpose of further evaluation is to contribute to enhancing shock responsive cash transfer and social protection schemes, with specific objectives to:

- Evaluate relevance, efficiency, effectiveness and sustainability of the cash transfer project.
- Generate lessons learned and provide feasible recommendations.

The audience of the evaluation and policy briefs are children and their families, the Government of Nepal, UNICEF, development and implementing partners, and other stakeholders who will make use of the evaluation results. The evaluation results will be used in advocacy for shock responsive and innovative social protection interventions for addressing the needs of children, most vulnerable and marginalized.

---

## 3. SCOPE OF WORK:

The evaluation will focus on the implementation of the cash transfer and its effects on household recipients. In total, 5,700 households are enrolled in the cash transfer project. These households are participants of the CFT. The sample size of the households is representative across Nepal. For comparison analysis on possible effects of the cash transfer, non-recipient households will be sampled from an existing database. Impact of the cash transfer will not be evaluated as it is a one-time short intervention.

---

## 4. EVALUATION FRAMEWORK AND METHODS

### *Evaluation framework and criteria*

The evaluation will be based on mixed-methods approach and remote data collection. The project will be evaluated against four (4) of OECD-DAC criteria: relevance, effectiveness, efficiency and sustainability.

#### **4.1. Relevance** of the cash transfer in response to social and economic impacts of Covid-19.

4.1.1. Is the cash transfer an adequate response to the needs of children, especially the most vulnerable and marginalized, in the context of Covid-19?

4.1.2. Is the cash transfer an appropriate intervention which fits the social protection framework and policies of Nepal?

#### **4.3. Effectiveness:** the extent to which the cash transfer improved the situation of children and their families?

4.3.1. To what extent did the cash transfer improve coping mechanisms of the households?

4.3.2. Did the cash transfer increase food security of children and their families?

4.3.3. To what extent did the cash transfer contribute to financial security of the households?

4.3.4. To what extent did the cash transfer improve children's situation in terms education?

4.3.5. Did the cash transfer improve children's access to health care and change decisions of parents to take their children with fever to hospitals instead of self-treatment at home?

4.3.6. How did the cash transfer help the households in general? What are the main uses of the cash transfer by the households?

**4.4. Efficiency:** the extent to which operations of the cash transfer were implemented in a timely manner and cost efficiently.

4.3.1. To what extent did the budget, logistics and operations of the cash transfer deliver outcome in a timely manner and economically?

4.3.2. Is the cash transfer faster and easier, or more difficult, in comparison to other e-Sewa transfer activities outside of the CFT recipients; SSA transfers of the Government and other service providers.

4.3.3. Is the amount of the cash transfer sufficient to make a substantial difference in the situation of 5,700 households?

**4.5. Sustainability:** the extent to which the effects of the cash transfer are likely to be sustained.

4.5.1. To what extent the cash transfer results are sustainable in the long run?

4.5.2. To what extent can the cash transfer be scaled-up in the context of Covid-19 in Nepal?

### *Sampling*

To measure the effects of the cash transfer, baseline data, including several rounds of data collected at different points in time since May 2020, will be available. A comparison group will be identified preferably based on PSM or other appropriate approaches. For measuring effects, the evaluation will use an appropriate sample size to test hypothesis of expected changes as a result of the cash transfer between treatment and comparison groups. The evaluation will use the sample size such that the probability the probability of a type 2 error is minimized. More details of the sampling and power calculations are in Annex I.

### *Data collection:*

The evaluation will utilize remote telephonic or online data collection with the participants and stakeholders. To measure various indicators of the cash transfer qualitative and quantitative data analysis are required. Preferred data collection activities are:

1. Telephonic interviews/surveys with the households
2. IVRs with the households
3. Interviews with the implementing partners:
  - e-Sewa, responsible for transferring cash to the households
  - ShareCast Initiative, responsible for registering and verification of the cash transfer recipients
4. UNICEF Social Policy, Evidence and Evaluation Team
5. Government partners who collaborate with UNICEF on social protection interventions (the list of interviewees will be provided through the consultation with the evaluation team during the inception phase)

The evaluation will also use the CFT data:

1. CFT results in May 2020
2. CFT results in July 2020
3. CFT results in August 2020
4. CFT results in October 2020

#### 4. CFT results in December 2020

#### 5. CFT results in January 2021

The presentations of the survey findings are available on: <https://www.unicef.org/nepal/reports/covid-19-child-and-family-tracker-baseline-findings>. More presentation will be uploaded in December 2020 - January 2021.

The evaluation team will have access to the CFT raw data that can be used for in-depth analysis and triangulation of the evaluation findings.

#### *Data Analysis*

In evaluating the cash transfer project, major questions to address are about positive and negative changes, and the way it improved the situation of children and women. The cash transfer targets the needs of children in all aspects which help their wellbeing, but it is also important to assess and analyze the household spending using the cash for other means which can directly and indirectly affect children. Many households, especially in the bottom two income groups report borrowing money to cope with financial struggles caused by the pandemic. The analysis should explain if the cash helped improve households' coping mechanisms and whether it had any effect on alleviating their immediate needs.

Significant findings of the CFT include children's decreased access to education and their difficulties to study at home. Especially children living in poor households had the least access to DL and are more likely to experience problems while studying remotely. The analysis is therefore expected to explain possible effects of the cash transfer on children's access to education coupled, food security and child protections issues, including child labour.

In terms of operational mechanisms of the cash transfer, the analysis must unpack issues related to possible errors which might have affected the speed of the project's implementation.

Data must be disaggregated by gender, ethnicity, castes, age, wealth and locations.

#### *Evaluability and Limitations*

UNICEF has enough data to evaluate the cash transfer project. UNICEF has been gathering data on social and economic impacts of Covid-19 on children and their families since May 2020, every 4-6 weeks. The first round of the survey can be used as the baseline and the 6<sup>th</sup>, last round, is the end-line.

Limitation of the evaluation are safety and risk issues in traveling to target locations for field-based data collection. This evaluation will therefore be conducted remotely. The cash transfer does not have Theory of Change, but it will be developed by the evaluation team.

### 5. TASKS, PAYMENT PLAN AND DELIVERABLES

#	Item	Duration
<b>1</b>	<b>Inception Phase</b>	
1.1	Desk review and secondary data analysis	10 days
1.2	Consultation with the SPEE Team	2 days
1.3	Consultations with partners	2 days
1.4	Draft inception report	5 days
1.5	Final inception report and data collection tools	3 days
	<b>Subtotal</b>	<b>22 days</b>
<b>2</b>	<b>Data Collection</b>	
2.1	Interviews and surveys	14 days
2.2	Triangulating findings and combining analysis	

2.3	Data analysis	
<b>Subtotal</b>		<b>14 days</b>
<b>3</b>	<b>Report Writing</b>	
3.1	Draft evaluation report	10 days
3.2	Final evaluation report	
3.3	Policy brief on reducing prevalence of stunting, wasted and underweight children	5
3.4	Policy brief on reducing prevalence of women with chronic energy deficiency	5
<b>Subtotal</b>		<b>20 days</b>
<b>Total number of days for evaluation report</b>		<b>56 days</b>

**6. DELIVERABLES:**

<b>1</b>	<b>Inception Report</b>
<p>The inception report is the outcome document of the research findings, interviews with UNICEF and stakeholders, and the review of the programme documents. The report must present the overall evaluation approach, detailed evaluation methodology, theory of change, evaluation matrix, process monitoring framework, final evaluation questions, data collection and sampling approach for the different data collection stages and types (including statistical power calculations for the household surveys) and timeframe for each proposed data collection method. The evaluation team must submit the first draft report by a required deadline. The draft report will be reviewed by UNICEF and the detailed list of comments, if any, will be shared with the evaluation team in 7 or 10 days. The evaluation team is expected to respond to the comments and revise the report in 7 days. Depending on the quality of the inception report, the evaluation team may be required to revise it more than once until it meets the UNICEF standards. The entire inception phase is the time when the Evaluation Team and UNICEF verify that the inception report covers every detail and clarify expectations.</p>	
<b>2</b>	<b>Summary of Initial Findings from the Interviews, Desk Reviews and Secondary Data Analysis</b>
<b>3</b>	<b>Copies of the Data Files and Analysis</b>
<b>4</b>	<b>Evaluation Report</b>
<p>The first draft evaluation report submitted to UNICEF will be reviewed and it usually takes from 7 to 10 working days to provide the Evaluation Team with comments. The Evaluation Team will be responsible for revising the report and resubmitting it within the requested timeline. Similarly, to the inception phase procedures, the writing phase of the evaluation report will include reviews and feedback by UNICEF. The timeline for the review and feedback on the first and second draft reports will take about 7 and 10 working days. The evaluation team must respond to all comments and revise the report in a required timeframe by UNICEF. The report should include background, detailed description of methodology, analysis of data which address each of the key evaluation questions and conclusions. Comments and suggestions gathered during the meeting shall be integrated into the draft final report. The length of the evaluation report must be between 40-60 pages. TOC, evaluation matrix, data collection tools, tables and graphs illustrating evaluation findings must be included in the report as annexes.</p>	
<b>5</b>	<b>Presentation of the evaluation findings to UNICEF and partners</b>
<p>The evaluation findings will be presented to UNICEF, and it will include a PowerPoint presentation summarizing the evaluation process and initial findings.</p>	

**7. DURATION: The start date is 01 February 2021. The contract expected to be concluded in 56 days.**

**8. WORKING LOCATIONS: REMOTE WORKING**

**9. PROPOSED PAYMENT SCHEDULE:**

No	Payment schedule	Percentage
1	Upon delivery of the final Inception Report	30 %
2	Upon completion of secondary data analysis	30 %
3	Upon delivery of the final Evaluation Report and Policy Briefs	40 %

**10. CONTRACT SUPERVISION:**

The evaluation team will be supervised by Evaluation Specialist. Evaluation Reference Group (ERG) will be formed and consisted of experts from the Government, UNICEF and other relevant partners. ERG members will provide support in quality assurance of deliverables through providing expert advise on the course of the evaluation direction, contents in the analysis, verification of findings and applicability of recommendations to the context of Nepal.

**11. QUALIFICATIONS AND EXPERIENCE REQUIRED**

The team should be composed of one Team Leader and team members to ensure the successful implementation of the assignment. Team members proposed in the submitted proposal must be available for the duration of their assigned tasks.

Team Leader will be the principal evaluation expert with an extensive experience in evaluating cash transfers and social protection programmes, strategies and policies. The Team Leader will lead the evaluation and oversee the entire process and be responsible for deliverables of excellent quality. The Team Leader must have:

- a. Advanced degree in public economics, sociology, and other social science related field.
- b. At least ten years' experience in managing, designing and conducting evaluations of cash transfers and social protections programmes.
- c. An extensive experience in social policy and strategies in developing countries, preferably in Nepal and South Asia.
- d. In-depth knowledge of human rights, equity and gender-based approaches to social protection programming and cash transfer schemes, policies and strategies.
- e. Demonstrated ability to deliver high-quality written reports in English and to engage effectively with stakeholders at all levels.
- f. Excellent writing and communication skills.
- g. Fluence in English is required.

The Evaluation Team Members will support the Team Leader in the secondary data analysis and interviews. They will be responsible for timely and accurate delivery of results.

- a. Members of the Team should have master's and bachelor's degrees in statistics, social policy, sociology, human rights, anthropology, and other social science related field.
- b. Data analyst must have at least 5 years of experience in statistical analysis and data visualization; qualitative analysis and data visualization and producing high quality research reports.
- c. Team Members must have experience in analyzing and interpreting cash transfer and social protection data, and strong knowledge of social policy and programmes for children.
- d. Team Members must have strong excellent writing and communication skills.
- e. Team must be gender balanced.
- f. Fluency in English is essential.

## 12. APPLICATION AND EVALUATION PROCESS:

In making the final decision, UNICEF considers both technical and financial aspects. The proposal that offers the best value for money will be recommended for award of the contract.

### **The Technical Proposal should include but not be limited to the following:**

- **Work Plan**  
Proposed work plan showing detailed sequence and timeline for each activity and man days of each proposed team member
- **Team Composition**  
Title and role of each team member
- **CV's**  
CV of responsible officers and field coordinators (including qualifications and experience)  
Ensure to include information related to the qualifications and experience of each proposed team member as required and outlined in item 10 of this document.
- Any project dependencies or assumptions

### **The Financial Proposal should include but not be limited to the following:**

Bidders are expected to submit a lump sum financial proposal to complete the entire assignment based on the terms of reference. The lump sum should be broken down to show the detail for the following:

- **Human Resource cost:** This should include the cost related to project planning and coordination operational cost.
- **Data Entry Cost:** The data on Social Protection beneficiaries are collected by the local government and the firm should make data entry into the MIS system closely with the local government. This should include the cost of data entry per beneficiaries as the estimated number of data entry may vary.

- **Data verification cost:** It is recommended to provide per beneficiaries’ data verification cost as exact number of entries may vary from estimation.
- **Travel Costs**  
All travel costs should be included as a lump sum fixed cost.  
For all travel costs, UNICEF will pay as per the lump sum fixed costs provided in the proposal.  
A breakdown of the lump sum travel costs should be provided in the financial proposal.  
Bidders are required to estimate travel costs in the Financial Proposal. Please note that i) Air travel costs shall be calculated based on economy class fare regardless of the length of travel and ii) costs for accommodation, meals and incidentals shall not exceed the applicable daily subsistence allowance (DSA) rates, as propagated by the International Civil Service Commission (ICSC).
- **Any other costs (if any)** Indicate nature and breakdown

**Technical evaluation criteria:**

The research organization/team has a strong record of evaluations of nutrition programmes	5
a. Advanced degree in public economics, sociology, and other social science related field. At least ten years’ experience in managing, designing and conducting evaluations of cash transfers and social protections programmes. a. An extensive experience in nutrition policy and strategies in developing countries, preferably in Nepal. b. An extensive experience in social policy and strategies in developing countries, preferably in Nepal and South Asia. c. In-depth knowledge of human rights, equity and gender-based approaches to social protection programming and cash transfer schemes, policies and strategies. d. Demonstrated ability to deliver high-quality written reports in English and to engage effectively with stakeholders at all levels. e. Excellent writing and communication skills. f. Fluence in English is required.	20
Members of the Team should have master’s and bachelor’s degrees in statistics, social policy, sociology, human rights, anthropology, and other social science related field; Team Members must have experience in analyzing and interpreting cash transfer and social protection data, and strong knowledge of social policy policies and programmes for children; Team Members must have strong excellent writing and communication skills.	10
Data analyst must have at least 5 years of experience in statistical analysis and data visualization; qualitative analysis and data visualization and producing high quality research reports.	10
Technical proposal is of excellent quality and includes: <ul style="list-style-type: none"> <li>• Sampling</li> <li>• Data collection methods</li> <li>• Detailed workplan</li> <li>• Risk mitigation and quality assurance measures</li> <li>• Ethical procedures</li> </ul>	25
<b>Total</b>	<b>70</b>



## Annex I. Sampling of the households with children for the CFT

The population sampling for telephonic interviews is composed of three stages:

1. **Grid sampling** for selecting target locations and ensuring representatives of the population. Existing databases<sup>8</sup> with telephone numbers was used to select households with children below the age of 18. A household with children is the only criterion used for selecting households. Grid based sampling approach ensured geographic coverage throughout Nepal. A selection of 250 wards (out of 6744) was made by randomly selecting wards with each grid of the map of Nepal through distance and density optimization algorithms. The selected number of wards is expected to be statistically robust at the 10% margin of error and includes 50 wards for replacement purposes (in case of majority non-response from small wards) or for further expansion of coverage of the survey. Within each ward in Nepal, it is expected there are about 1,500 children. Given this, it is recommended that 300 children be sampled from each selected ward (to get a reasonable margin of error). This brings the total sample size to  $300 \times 20 = 6000$  (or  $300 \times 25 = 7500$ ). The map below shows the locations of wards where households will be sampled:

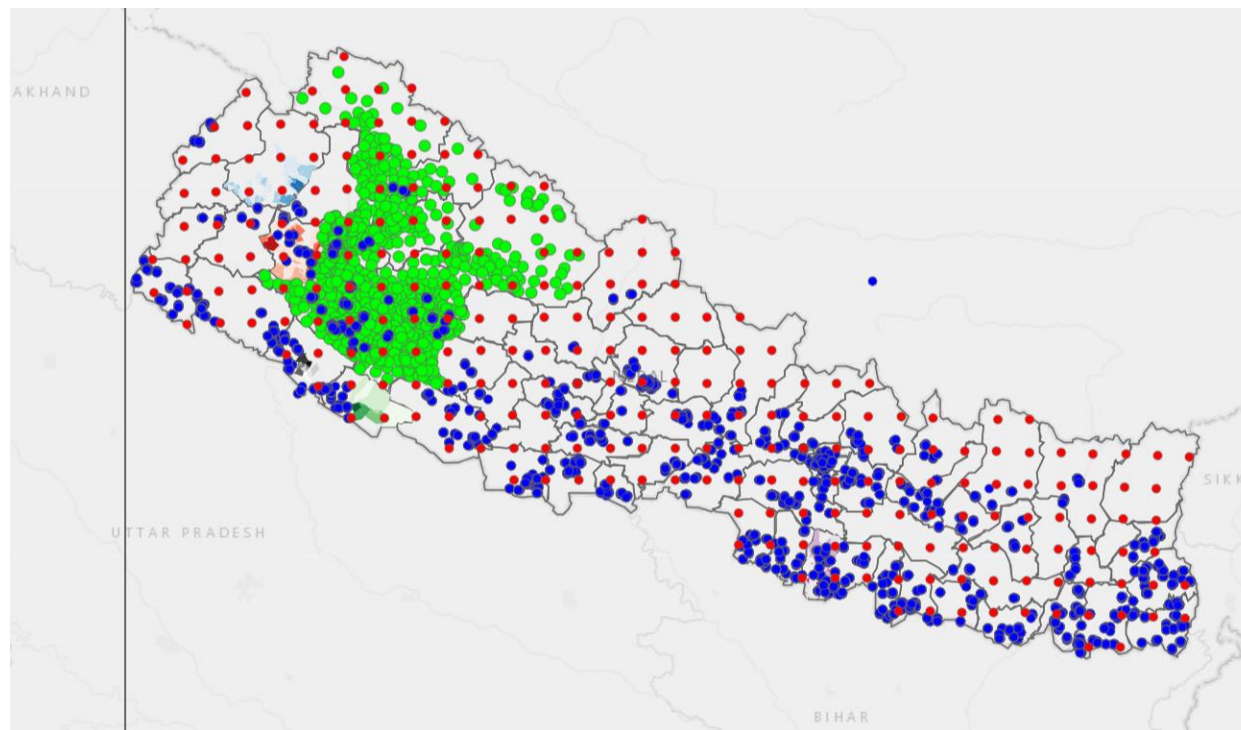


Figure 1: Map of Nepal. Red dots =grids; blue dots=locations of wards based on ShareCast Initiative database; green dots =locations of ward secretaries based on UNICEF database.

2. **Simple random sampling** of telephone numbers of 7,500 households from selected locations in the database.

<sup>8</sup> Sharecast Initiative and UNICEF databases

**3. Purposive sampling to complete the sampling of households.** During the first round of the survey, interviewers continued sampling through identifying households with children and subsequently registering them as participants based on informed consent. All selected households with children were registered in a special database intended to be used specifically for this project.

---

### Technical note for sample size determination for the CFT CT Pilot Program Evaluation<sup>9</sup>

One off cash transfer 2000 NPR per child (up to 2 children) for participants in the CFT through E-SEWA as a test of digital payment system. For a one-off transfer, the distinction between monitoring and evaluation becomes more nebulous. We propose a mixed-methods approach where quantitative and qualitative tools will be used to answer different hypotheses about the effect/impact of this pilot.

#### 1<sup>ST</sup> ORDER EVALUATION QUESTIONS

1. Speed of transfer broken down by mode of transfer
2. Error rate broken down by mode of transfer
3. Feedback from respondent about ease of transfer and difficulties
4. Comparison with other E-SEWA data about transfer efficiency outside the CFT participants
5. Comparison with government MoHA SSA transfers made through banking system
6. Comparison with other non E-SEWA service providers (e.g., Khalti)

#### 2<sup>ND</sup> ORDER EVALUATION QUESTIONS

For this part, the CFT itself would provide the baseline data as well as the next rounds (end-line) data for the treatment group. Furthermore, for the treatment group, future rounds of CFT questionnaires would also need to be changed to add in qualitative self-reflective assessments. A control group would have to be identified (based on PSM and/or other approaches) and tools (qualitative and quantitative) developed accordingly. The evaluation team may decide to administer separate instruments – especially for qualitative components of the evaluation.

7. What are the main uses of the cash transfer? (follow-up question)
  - a. Food & Nutrition
  - b. Non-food: Health, Education, Wash, Home repairs
  - c. Others: Debt
8. Sufficiency of the transfer?
9. Effectiveness of the transfer?
10. Use data from the CFT to analyze whether after the cash transfer there was any change in:
  - a. Coping mechanisms
    - i. Did the percent reporting borrowing change significantly? (49.3 %)
    - ii. Did the percent reporting depleting savings change significantly? (39.6 %)
    - iii. Did the percent reporting reducing expenditure change significantly? (22.9%)
    - iv. Did the percent reporting asking for support from friends/relatives change significantly (16.3%)?
  - b. Top 3 needs
    - i. Did the percent reporting food as a top three need change?
    - ii. Did the percent reporting financial support as a top three need change?
  - c. Was there any impact on the per cent of families struggling daily for food
  - d. Was there any impact on the per cent of children experiencing dietary changes
  - e. Was there any impact on the per cent of respondents who can buy a mask
  - f. Was there any impact on per cent of respondents who took their children to a hospital/health facility if they had fever for more than 7 days
  - g. Was there any significant change in the per cent of respondents reporting child labor?

---

<sup>9</sup> Prepared by A. Bonnerjee. For discussion purposes only.

## RECOMMENDED SAMPLE SIZE

Based on A/B testing to set up hypotheses to answer questions so that power is at least 80 per cent (the norm expected in social sciences). The optimal size would be around 4000 households per group. The sample would be sufficiently large to be able to detect *small* effects given that this is a 1-time transfer at a time where households have already faced protracted economic hardships. It would be highly powered if larger effects are detected post-fact.

## PROPOSED METHODOLOGY FOR 2<sup>ND</sup> ORDER EVALUATION

**Qualitative:** Tools to be developed

**Quantitative:** A/B Testing (also known as Hypothesis testing): Testing if there is a difference in observed outcomes between two independent population groups (A and B). Paired tests may be appropriate when comparing the same sample before and after situations. Joint hypothesis testing would also be appropriate when considering multiple hypothesis/questions to be answered. Reflexive valuation on the panel data (before/after) comparisons on select variables such as use of cash, impact on wellbeing and dignity etc. would also be included

**Examples:**

Suppose we wish to test if 1-time cash transfer affected the proportion of children experiencing dietary changes (captured by  $\mu$ ) of children. Our hypothesis is that in a short time, we may not see a difference at all. We construct the null hypothesis as  $\mu_1 = \mu_2$  where  $\mu_1$  and  $\mu_2$  are the measurements of nutrition either prior ( $\mu_1$ ) and post ( $\mu_2$ ) for a reflexive evaluation (paired testing), or,  $\mu_2$  measures outcomes of a completely independent control group that did not receive such an outcome. Either case, the alternative hypothesis is that the means changed, that is,  $\mu_1 \neq \mu_2$ .

$$H_0: \mu_1 = \mu_2$$

$$H_a: \mu_1 \neq \mu_2$$

Well known techniques (e.g, z-test/t-test/f-test, anova) can be used to test if the means (or proportions) are identical or not for the two groups. They also provide the appropriate criteria for rejecting or accepting the null hypothesis at an assumed significance level (0.05), or equivalently, the justification for accepting the alternative hypothesis.

## ERRORS, SAMPLING AND POWER CALCULATIONS

The evaluation will use an appropriate sample size such that the probability of a type 2 error is minimized. In hypothesis testing we encounter two kinds of errors. A type 1 error ( $\alpha$ ) is best understood as a false positive whereas a type 2 error ( $\beta$ ) is a false negative. For instance, if a person goes for a pregnancy test – is not pregnant – but the test finds the person pregnant – it is a false positive or a Type 1 error. On the other hand, if the person is pregnant but tests to be not pregnant, then we have a false negative or a Type 2 Error. The power of a test is  $1 - \text{Type 2 Error}$ . It minimizes the probability of a false negative. The higher the power of a test the higher the likelihood of finding an effect if there is any.

The power of a test is typically required to be over 80 per cent for credibility, which suggests that there is a 20 per cent change of committing a type 2 error. In other words, setting a value of  $1 - \beta$ , the power of the test, to 0.8 or higher would imply that the value of  $\beta$  must be 0.2 or less. The probability of a type 1 error,  $\alpha$  is typically assumed to be 0.05 and is also known as the significance level. It is acceptable to have a type 1 error rate of 5% globally. Statistically, these two types of errors trade-off against one another in any given sample. Efforts to reduce 1 would tend to likely increase the other type of error.

A critical question is about the requisite sample size to achieve credibility (power over 80 per cent). Too small a sample can lead to under detection of the effect. Too large a sample wastes resource. The balance is to find a sample size large enough that can detect an effect, if there is one, without wasting limited resources. Sample size depends on (and determines) three factors: effect size, significance level and power.<sup>10</sup>

**Effect size** is an important factor in selecting the optimal sample size.<sup>11</sup> In general, it refers to the magnitude of the perceived impact and typically calculated as the difference between the means of the two samples divided

<sup>10</sup> Given any three of these variables, the fourth can be calculated as they have a relational structure between them.

<sup>11</sup> Informally can be thought of as a proxy for the ‘impact’ or standardized level of change.

by their pooled standard deviation. Large effect sizes require smaller samples and smaller effect sizes would require larger samples.

Hence in order to calculate the appropriate sample size we would need to have an estimate of the means for each of the samples and the pooled standard deviation (to calculate effect size), the requisite power (0.8+), the desired significance level (0.05). Given these, it is straightforward to calculate the optimal sample size.<sup>12</sup> Note that when there are many questions being asked (i.e. many hypothesis) the maximum sample size is selected to ensure that all the tests have adequate power.

**EXAMPLE 1 (SAMPLE SIZE SELECTION & POWER)**

Prior to the cash transfer the proportion of children experiencing changes to their dietary intake  $\mu_1 = 0.25$  (25 per cent). The research issue is if this changed in the next month. In a short time, we may not see a significant change, so our null hypothesis is that there is no change. However, we would like to be able to detect even a 10 per cent decline, so that  $\mu_2 = 0.225$  (22.5 per cent) if it is found to be statistically significant. Hence our hypothesis tests can be formulated as:

$H_0: \mu_1 = \mu_2 = 0.25$  (Null hypothesis) or  $H_0: \mu_1 - \mu_2 = 0$

$H_a: \mu_1 > \mu_2$  (Alternative hypothesis: 1 tailed) or  $H_1: \mu_1 - \mu_2 > 0$

$H_a: \mu_1 \neq \mu_2$  (Alternative hypothesis: 1 tailed) or  $H_1: \mu_1 - \mu_2 \neq 0$

Consider the two-tailed variation. The alternative hypothesis is that there is a difference between the observed mean ratio or the difference between the means does not equal zero.

- Given the value of 0.9 as acceptable power, and 0.05 as significance level, we would need a total sample size of 3,043 (in each group) to be able to detect a significant difference between  $\mu_1$  &  $\mu_2$  if in fact there is one. (Figure 1)
- If we allow for lower power (80 per cent), then a smaller sample would be needed: 2,273 in each group (2).
- If we expect a much large magnitude of change, and we assume  $\mu_2 = 0.20$ , then we would need 731 in each group for 90% power and 546 in each group for 80% power (Figure 3 & Figure 4). Hence expected ‘effect size’ plays a major role in the optimal sample size.
- If we allow for a higher margin of type 1 error (10%) but expect at  $\mu_2 = 0.225$  we get a lower sample size (Figure 5).

These are shown in the figures below:



**Figure 1: Optimal sample size (90% power)**



**Figure 2: Optimal sample size (power=80%)**



**Figure 3: Optimal sample size (Larger effect size)**



**Figure 4: Optimal sample size (Larger effect and power=80%)**



**Figure 5: Optimal sample size (Type 1 error=10%)**

Considering that across all our options the maximum sample size required to detect a small size effect is about 7750 (3,875 per group) which would be yield enough power to detect small changes if in fact there is one (it is

<sup>12</sup> In this note, all the calculations were performed using the pwr package in R and scipy.stats package in python.

significant). In practice, due to non-response, it would be safer to take 10-15% more. Table 1 shows a variety of scenarios for different variables (note – the values are assumptions).

**Table 1: Optimal Sample Sizes under different scenarios**

Case	Variable	Null	Alternative	Type 1 Error	Power	Sample Size
1	Dietary $\Delta$	$\mu_1 = 0.25$	$\mu_2 = 0.23$	$\alpha = 0.05$	0.9	3,503*2=7006
2	Dietary $\Delta$	$\mu_1 = 0.25$	$\mu_2 = 0.23$	$\alpha = 0.05$	<b>0.8</b>	2,273*2=4546
3	Dietary $\Delta$	$\mu_1 = 0.25$	$\mu_2 = \mathbf{0.20}$	$\alpha = 0.05$	0.9	731*2=1462
4	Dietary $\Delta$	$\mu_1 = 0.25$	$\mu_2 = \mathbf{0.20}$	$\alpha = 0.05$	<b>0.8</b>	546*2=1092
5	Dietary $\Delta$	$\mu_1 = 0.25$	$\mu_2 = 0.23$	$\alpha = \mathbf{0.10}$	0.9	2,408*2=4816
6	Struggling $\Delta$	$\mu_1 = 0.19$	$\mu_2 = 0.17$	$\alpha = 0.05$	0.9	3,875*2=7750
7	Struggling $\Delta$	$\mu_1 = 0.19$	$\mu_2 = \mathbf{0.16}$	$\alpha = 0.05$	0.9	1,683*2=3366
8	Borrowing $\Delta$	$\mu_1 = 0.35$	$\mu_2 = 0.32$	$\alpha = 0.05$	0.9	2,600*2=5200
9	Reliance $\Delta$	$\mu_1 = 0.16$	$\mu_2 = 0.14$	$\alpha = 0.05$	0.9	3,346*2=6692
10	Food $\Delta$	$\mu_1 = 0.50$	$\mu_2 = 0.45$	$\alpha = 0.05$	0.9	1,048*2=2096
11	Mask $\Delta$	$\mu_1 = 0.06$	$\mu_2 = 0.04$	$\alpha = 0.05$	0.9	1,236*2=2096

---

Reference documents

**ANNEX III: Ethical Procedures for Evaluations**

The following principles<sup>13</sup> guide the ethical procedures for evaluations conducted at UNICEF:

1. **Utility:** evaluations are designed to address and effectively serve the needs of the target groups and decision-making needs of intended users.
2. **Necessity:** evaluations are commissioned where they are necessary, and the effort justified in terms of the benefits likely to accrue from the evaluation exercise.
3. **Independence:** evaluation is free of bias. Evaluators exercise independent judgement and are not influenced by the views or statements of any party.
4. **Impartiality:** evaluations produce a comprehensive presentation of strengths and weaknesses of the policy, program, project or organizational unit being evaluated, taking due account of the views of a diverse cross-section of stakeholders.
5. **Credibility:** evaluations are based on reliable data and illustrate evidence based on verified findings, balanced judgement and lessons learned.
6. **Conflicts of Interest:** evaluators disclose in writing any past experience and their relationships, including of their families and friends to an evaluated subject.
7. **Honesty and integrity:** evaluations are conducted by evaluators who truthfully represent their level of skills, knowledge and present unbiased findings.
8. **Accountability:** responsible UNICEF Staff Members and evaluators are accountable for the completion of evaluations; ensure that evaluation expenditures are properly accounted for.
9. **Obligations to participants:** evaluators are obliged to receive consent from participants and respect their right to provide information in confidence. It is mandatory that evaluators inform participants make participants aware of the scope and limits of confidentiality. Any sensitive information must be ensured that it cannot be found by anyone and participants are protected from retaliations.
10. **Respect for dignity and diversity:** evaluations are conducted according to UNEG ethics code of conduct, ethics procedures and with respect to different cultures, customs, religious beliefs, gender roles, disability, age and ethnicity.
11. **Human rights, equity and equality:** evaluations illustrate fair representation of vulnerable groups, gender roles and ethnic groups.
12. **Avoidance of harm:** to negative effects of evaluations on participants, evaluators prepare risk management plans to minimise unnecessary harms.

Credible evaluations represent accurate, complete and reliable findings and analysis. It is the responsibility of Programme Staff Members and Evaluation Specialist to ensure:

1. The purposes and content of evaluations are clear and precise.
2. Evaluations are conducted with methods and techniques that meet high standards and are based on validated information.
3. Evaluations present complete and fair assessment of programme's strengths and weaknesses.
4. Findings, conclusions and recommendations are based on the evaluation findings only.
5. Transparent communication with stakeholders on the purpose, methodology and criteria of evaluations, and intended use of findings.

---

<sup>13</sup> UNEG Code of Conduct for Evaluation in the UN System, UNEG, 2008.

6. Evaluation reports are made accessible to the intended users and participants in relevant languages. Evaluation reports are usually made public and will be withheld from publication for justified reasons.
7. Evaluation findings are presented and discussed with stakeholders before the evaluations are completed.

### **Ethical considerations for remote data collection in times of COVID-19 pandemic<sup>14</sup>**

Many UNICEF Country Offices are exploring virtual means to gather evidence about the impacts to and responses of the COVID-19 pandemic, given COVID-19 transmission risks related to direct person-to-person data collection. The 'do no harm' principle is priority when collecting data, as confirmed by the Technical Note by the UNICEF Evaluation Office.

Innovation in using remote data collection methods is encouraged and needed. At the same time, this should not let us lose sight of ethical standards in data collection and analysis, especially in times of increased vulnerability. This note presents some ethical considerations for evidence generation when planning remote data collection. The remote data collection methods considered are surveys via phone, sms, web-based (e.g. U-Report/RapidPro) or interactive voice response (IVR). For more information about the use of these technologies see [here](#) and [here](#).

Two initial considerations:

1. UNICEF's Procedures for Ethical Standards in Research, Evaluation, Data Collection and Analysis set the standards, regardless of the methods used, when data collection involves human subjects.
2. Ethics is not just about the procedure but about being reflective about possible ethical issues within the local context, taking a step back, discussing, creating mitigation strategies and adjusting plans before heading into data collection.

Other core ethical issues to consider and some suggested practices for remote data collection:

#### **1. Harms and benefits**

- A key ethical principle is for there to be clear benefits likely to arise from the evidence generation, and any harm to be avoided for participants, particularly when contacting vulnerable groups. Therefore, at the outset the question needs to be asked whether the remote survey will actually benefit the respondents and their communities. If not, consideration must be given to the modification (considering secondary data sources) or withdrawal of the data collection.
- This becomes even more critical when contacting people using their personal data, like phone numbers, for reasons that have nothing to do with the intended purpose for which these phone numbers were originally provided. For example, if phone numbers of participants in existing programmes are considered as a source of contact information for a remote survey that is unrelated to the programme, the purpose for using it needs to be well justified and consent clearly requested when contacting these persons, respecting any decision not to participate or requests that their numbers not be used for these purposes. Compliance with any legal/privacy conditions related to the use of personal data as stipulated by the terms and conditions of the programme is requisite.

---

<sup>14</sup> This note is authored by Tom Pellens, Multi-Country Evaluation Specialist at the UNICEF Regional Office Of South Asia. The support of Gabrielle Berman, UNICEF Senior Advisor, Ethics in Evidence Generation, in its formulation is gratefully acknowledged. It is intended for the use by UNICEF offices engaged in COVID19 related evidence activities. It may be freely shared within UNICEF, with other development agencies, and with our programme partners.

- Every question asked needs to have a clear purpose to contribute to intended benefit and avoid duplication of evidence, particularly given the current stressful situation respondents may be in. Keep surveys short and limit the data collected to what is absolutely necessary.

## 2. Personal and sensitive information

- Be cautious about asking potentially sensitive personal information or questions without clear purpose of use and benefit, especially since one cannot observe or control privacy conditions when asking questions through remote means, and the COVID-19 response measures may significantly negatively impact privacy.
- Protocols facilitating referral need to be in place when protection issues come up during phone surveys, similar to face-to-face surveys. In the case of sms/online/IVR surveys referral contact details can be mentioned at the end of the survey. Where support services or helplines are not available or unlikely to be reachable given demand, sensitive questions should not be asked.

## 3. Privacy and confidentiality

- When drawing on lists with contact details of participants in existing programmes, separate, if possible, the phone numbers from other identification data (e.g. names), and sample and conduct the survey based on de-identified data sets.
- If a panel survey would be valuable and more identifiable data is required, explicit consent should then be asked for follow up interviews, with clear indications of the nature, duration and frequency of these.
- Phone/sms/IVR surveys may be based on random selection of phone numbers of publicly available lists (or using a random phone number generator). These will likely be de-identified. Wherever possible and appropriate, avoid collecting personally identifiable information during the data collection process. This also applies for online surveys.<sup>15</sup>

## 4. Informed consent

- In respecting people's rights, informed consent must be sought when using remote means similar to any other data collection among human subjects. Practical tips in case of phone surveys can be found here. When using social media for data collection, the Innocenti Research Brief on Ethical Considerations in Using Social Media for Evidence Generation includes useful ethical guidance.
- Respondents need to be able to opt out of data collection, skip questions they do not want to answer, and have their name removed from the database for any future survey. This needs to be made explicit when conducting phone interviews. SMS/IVR/Online surveys should also include such options.

## 5. Other

- Remote surveys can be prone to biases because some groups do not have access to the technology used, restricted sample frames or low response rates. Since accurately presenting findings is core to the integrity of the evidence generation, one should acknowledge that bias may be present and be clear about the potential lack of representativeness of the data and all potential limitations.
- Remote data collection via phones may have a financial cost for the respondents. Compensation therefore needs to be considered (e.g. via phone credit or mobile money).

---

<sup>15</sup> U-report data are anonymous in that the phone numbers are hidden from admin users. See U-Report Key Principles.

## Do's and Don'ts on Ethical Considerations for U-Report during the COVID-19 Pandemic

### Do's

- ✓ **Take into account the potential for limited privacy** during lockdowns and only ask questions that would be acceptable in low privacy settings
- ✓ **Think carefully about the sensitivity of the questions** and the potential for retraumatizing persons who may not be able to access support services or informal coping mechanisms and strategies (such as hanging out with friends).
- ✓ **Think about ways of framing sensitive questions that don't require children to revisit trauma** or soften any potential negative experience (even this more indirect line of questioning should only be asked if information is absolutely critical for current service delivery or policies). I.e. Using third person language, focusing on the dynamic with the parent rather than actions "have you become closer to your parents, do your parents get frustrated a lot, have you been spending more time with your parents, are you enjoying spending more time with your family?....
- ✓ **Always include local resources** and supports in relation to the subject matter (to ensure that we are not just extracting information but also providing important information). Where local resources are not available find reputable international resources.
- ✓ **See children as resilient** and not just victims and ask questions about their strengths and how they are coping.
- ✓ **Try and encourage children to also reflect on any positive aspects** of the situation and not just the negatives.
- ✓ **Try and end any poll with hopeful or empowering questions** such as "what is the first thing you will do when lockdown ends", "what would you recommend that governments do to help children during times like these?"
- ✓ **Clearly frame questions in child and age appropriate languages** and references.
- ✓ **Ensure appropriate and robust local supports and referrals** are available and accessible if you are going to ask any sensitive questions or if the poll is likely to raise further questions or concerns. This could involve having an expert available during the poll or referring children to a child helpline or helping a child to get in touch with local support services where they are available or accessible. (Provision of information is not adequate)
- ✓ **Make it clear that answering any particular question is voluntary** and that they can choose to answer all, none or some of the questions that are posed.
- ✓ **Ask children open ended questions**, particularly in relation to priorities, opinions, perspectives and experiences so that they are truly given a voice and not merely responding to our opinions, priorities and beliefs.

### Do Not

- **Use U-Report as a diagnostic tool**, there are sites by reputable population health organisations that provide expert advice and appropriate cautions etc. (refer children to these). This is particularly important in populations where persons cannot easily or quickly get access to tests. Being symptomatic or not (particularly for children) is not a conclusive test.
- **Ask sensitive questions in contexts where privacy is unlikely to be available**
- **Provide personal identifiable information from U-Report to third parties** without the explicit permission or on the request of the child and, if not placing the child in harm's way, and with the permission of a relevant guardian, particularly in contexts where this is legally required.
- **Ask sensitive questions that may upset or distress a child** and for which a child may have no resources, supports or access to formal or informal coping mechanisms (such as the ability to leave the house and walk around, or to spend time with friends outside of the house or access appropriate medicines)
- **Ask children about their direct experience of violence or abuse** in contexts for which they cannot (a) be removed or (b) receive psychosocial support or (c) may not have privacy.
- **Ask children leading questions** about their state of mind, attitudes, experiences and perspectives.
- **Collect any information that is not critical** or will not directly inform policy or programming. Ensure that each and every questions is necessary and will be used to further the rights and wellbeing of children, particularly during this difficult time.
- **Present U-Report findings as representative of the entire population of children.** Make clear in any report the limitations of the findings, the groups of young people that were not included and the implication in relation to findings for programming and policies of those that have been excluded.

## **ANNEX IV: Table of Contents for the Inception Report**

### **CONTENTS**

#### **1. INTRODUCTION**

- 1.1. Objective of the evaluation
- 1.2. Background and context
- 1.3. Scope of the evaluation

#### **2. APPROACH AND METHODOS**

- 2.1. Evaluation criteria and questions
- 2.2. Conceptual framework
- 2.3. Evaluability
- 2.4. Data collection methods
- 2.5. Analytical approaches
- 2.6. Risks and potential shortcomings

#### **3. PROGRAMME OF WORK**

- 3.1. Phases of work
- 3.2. Team composition and responsibilities
- 3.3. Management and logistic support
- 3.4. Calendar of work
- 3.5 Draft outline of evaluation report

#### **ANNEXES**

- 1. Terms of reference of the evaluation
- 2. Evaluation matrix (detailing criteria, main questions, subsidiary questions, what to look for, data sources, data collection methods)
- 3. Stakeholder map
- 4. Tentative outline of the main report
- 5. Interview checklists/protocols
- 6. Theory of change or outcome model
- 7. Detailed responsibilities of evaluation team members
- 8. Reference documents
- 9. ProgrammeProgramme
- 12. Detailed work plan

## ANNEX V: Resources for inception and evaluation reports

The evaluation will be conducted according to UNEG (United Nations Evaluation Group) Code of Conduct for Evaluation in the UN System (). Other documents to review before starting the evaluation are:

- UNICEF-Adapted UNEG Evaluation Report Standards. 2017:  
<https://www.unicef.org/media/54771/file>
- United Nations Children’s Fund. 2015. *UNICEF Procedure for Ethical Standards in Research, Evaluation, Data Collection and Analysis*,  
(<https://unicef.sharepoint.com/teams/OoR/Shared%20Documents/UNICEF%20Procedure%20on%20Ethics%20in%20Evidence%20Generation%20092015.pdf>).
- United Nations Children’s Fund. 2015. *UNICEF Procedure for Quality Assurance in Research*,  
(<https://unicef.sharepoint.com/teams/OoR/SiteAssets/SitePages/Procedures/UNICEF%20Procedure%20for%20Quality%20Assurance%20in%20Research.pdf>).
- Graham, A., Powell, M., Taylor, N., Anderson, D. & Fitzgerald, R. 2013. *Ethical Research Involving Children*, Florence: UNICEF Office of Research-Innocenti.
- The best UNICEF evaluation reports and good practices:  
[https://www.unicef.org/evaluation/index\\_60807.html](https://www.unicef.org/evaluation/index_60807.html)

### Other useful documents:

- United Nations Evaluation Group. 2008. *Ethical Guidelines for Evaluation in the UN System*,  
(<http://www.uneval.org/document/detail/102>).
- United Nations Evaluation Group. 2014. *Integrating Human Rights and Gender Equality in Evaluations*,  
(<http://www.uneval.org/document/detail/1616>).
- United Nations Evaluation Group. 2005. *Standards for Evaluation in the UN System*,  
(<http://www.uneval.org/document/detail/22>);

---

United Nations Evaluation Group. 2005. *Norms for Evaluation in the UN System*,  
(<http://www.uneval.org/document/detail/21>).