METHODOLOGICAL REPORT

Pluralistic Memories Project – Pilot Survey – Sri Lanka
2015

Sumedha Jayakody
Doctoral Researcher – University of Lausanne

Dammika Herath
Kandy Consulting Group – Sri Lanka

Table of Contents

1. Sampling strategy

1.1 Survey areas

The areas for conducting the pilot study were carefully selected by the Sri Lankan local research team in order to simulate some of the realistic challenges that might arise during the implementation of the project's island wide longitudinal survey. These areas were selected based on two main criteria. Firstly they had to be reflective of the country's socio-demographic diversity with regard to factors such as language, religion, urbanism, socioeconomic conditions, and past conflict exposure. Secondly they had to be closer in terms of proximity in facilitating supervisory and monitoring requirements.

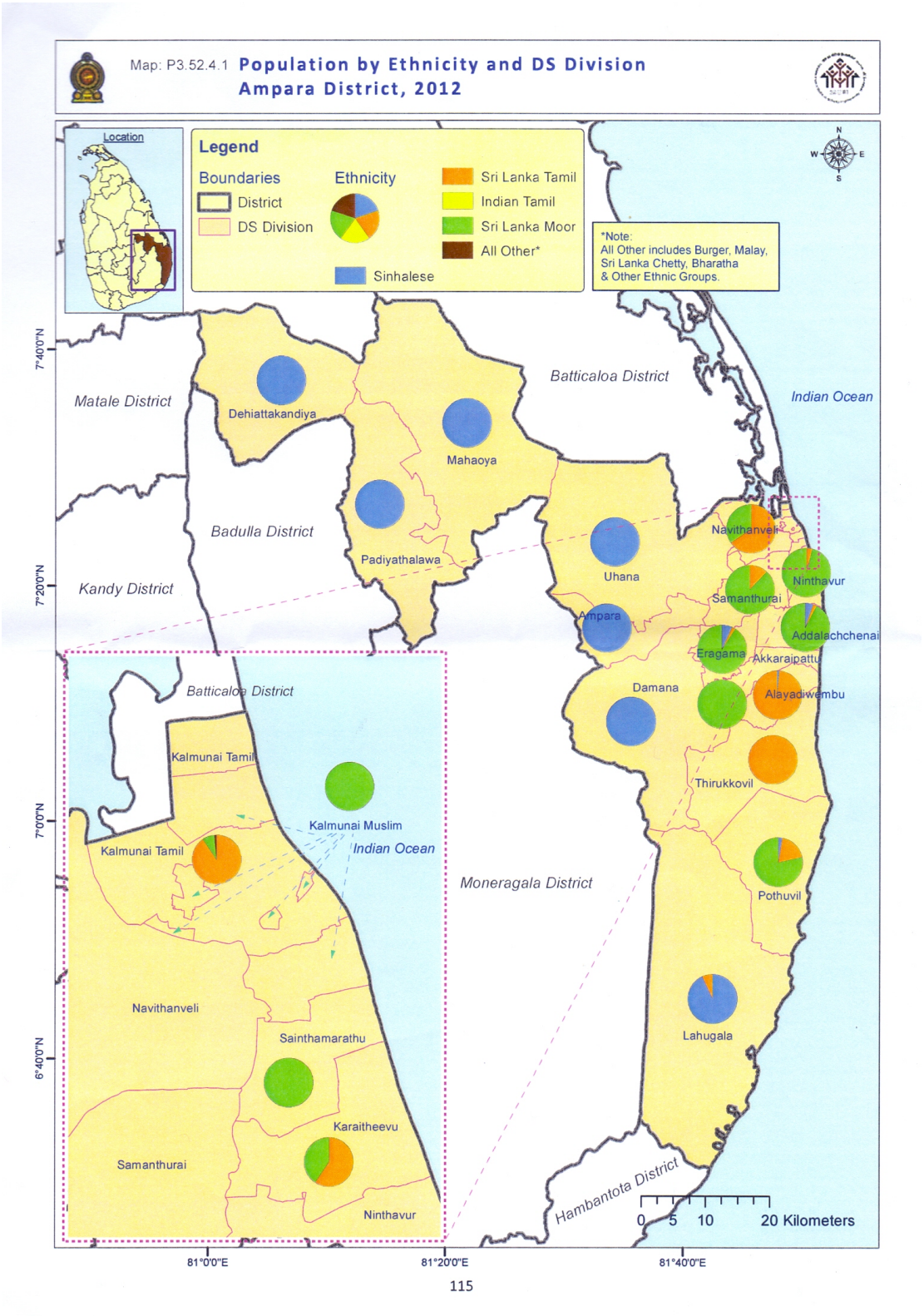
Based on these criteria, two districts, Ampara & Matale were selected. Ampara situated to the Eastern border of the country was a district that was directly affected by the war whereas Matale a district situated in the central highlands of the country, which was not directly effected by the proceedings of the war.

Within these two districts, 12 local areas were selected for the implementation of the pilot study based on the requirements of the PMP central team. The requirement was to have local areas that have been defined at an intermediate scale between small neighbourhoods and major administrative divisions, within which inhabitants were likely to move for daily activities. According to this requirement the definition of a local area applicable to Sri Lanka was a divisional secretariat division (DS division) a medium sized administrative division, which roughly comprises of 200 Km^2. The 12 DS divisions that were selected in Ampara were Uhana, Ampara, Damana, Irakkamam, Akkaraipattu and Alayadiwembu. The DS divisions selected for Matale included Yatawatta, Matale, Pallepola, Ukuwela, Ambanganga Korale and Rattota. Both districts also have DS divisions in their name, which comprise of townships reflecting commercial hubs of those respective districts. Figure 1 & 2 depicts the two districts and their local areas with a statistical breakdown in terms of ethnicity.

In the selection of areas, priority was given to linguistic diversity. As a result 12 of the local areas were equally divided among 6 Sinhalese speaking and 6 Tamil speaking areas. With the added pressure of covering as much territory as possible, locations with a sizeable population of a targeted ethnic group speaking a particular language (either homogeneously or heterogeneously dispersed in terms of ethnicity) were selected. This resulted in interviews being conducted only with members of one linguistic group per DS division.

This definition of a local area (DS division) within the sampling strategy has been consistently used alongside the definition of a local area within the pilot survey questionnaires and life calendars. Their precise boundaries have been predefined by the government of Sri Lanka and have been clearly communicated throughout the pilot project's geographical information via (1) an inventory of all local area units, and (2) a country map displaying the drawing of their area boundaries. All these documents are publicly available through Sri Lanka's department of census & statistics.

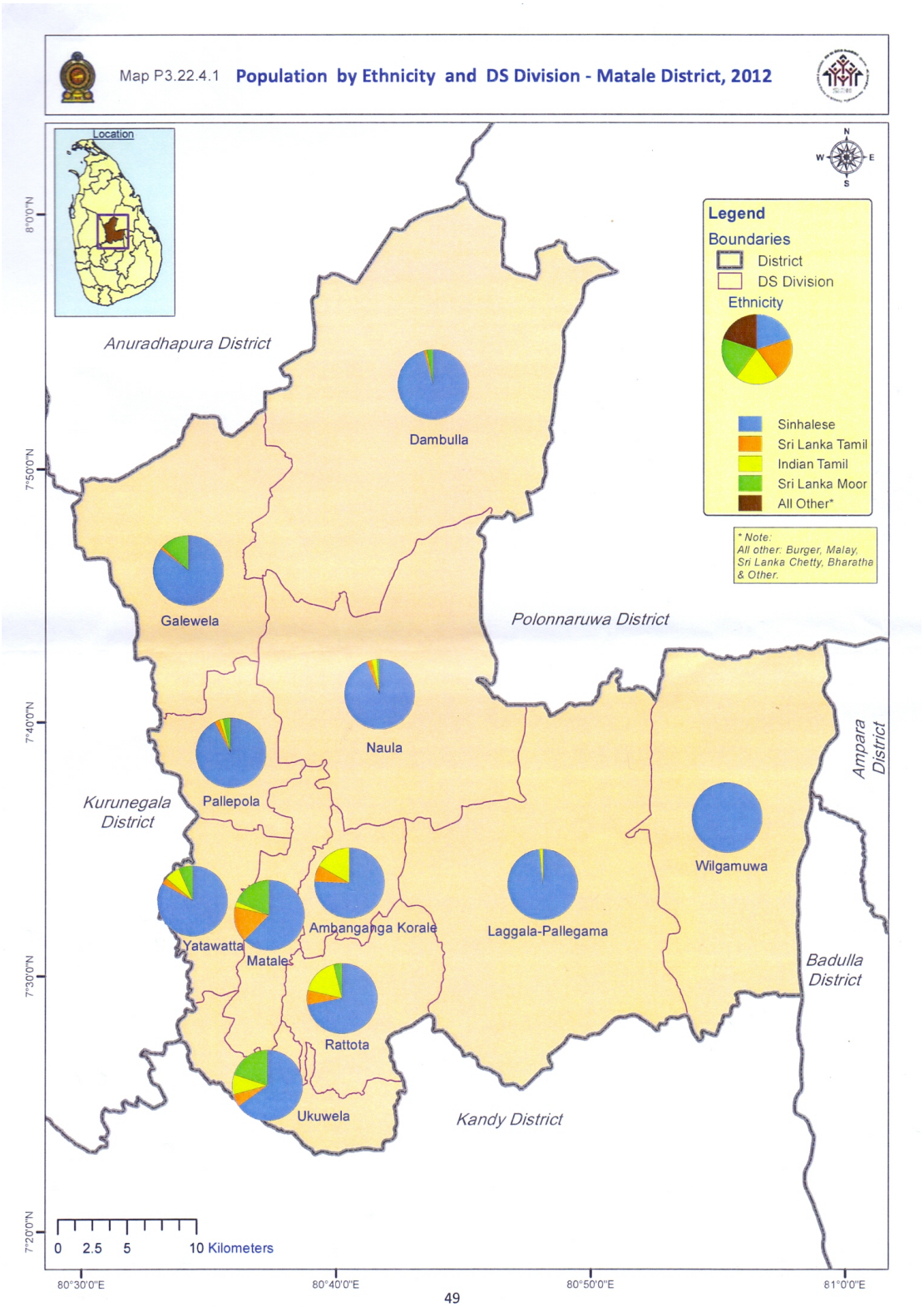
Figure 1 – Map of Ampara with its DS division boundaries and ethnic breakdown per DS division (local area boundary)



115

http://www.statistics.gov.lk/PopHouSat/PopulationAtla_2012/04_DSLevelMaps/Ma%20P3.52.4.1%20Ampara%20-%20Population%20by%20Ethnicity%20by%20DS.pdf

Figure 2 - Map of Matale with its DS division boundaries and ethnic breakdown per DS division (local area boundary)



http://www.statistics.gov.lk/PopHouSat/PopulationAtla_2012/04_DSLevelMaps/Map%20P3.22.4.1%20%20Matale%20-%20Populat%20ion%20%20by%20Ethnicity%20by%20DS.pdf

1.2 Network Sampling

The sampling methodology used within this pilot survey was network sampling and a cardinal point of interest for the PMP central team was to evaluate the future effectiveness of implementing this sampling methodology within the local setting through this pilot study. Been utilized as a method of selecting and studying sub-populations of interest, network sampling tends to continuously recruit respondents while the survey instrument is been administered until a point of saturation predefined by the researcher is achieved.

The field survey team recruited by the Kandy Consulting Group (the 3rd party survey consultancy firm contracted by the local PMP project team) comprised of 6 enumerators / research assistances (or RAs according to Kandy Consulting Groups terminology) and 2 field coordinators (or supervisors). The recruitment was performed keeping in mind the 2 official languages operating in Sri Lanka, which are Sinhalese & Tamil. Hence out of the 6 enumerators, 3 were native Sinhalese speakers and 3 native Tamil speakers. Similarly among the supervisors, 1 was a native Sinhalese speaker and the other a native Tamil speaker. All four members of the Sinhala speaking survey field team were males whereas all members of the Tamil speaking team were females. When recruiting the native Tamil speaking positions of the field research team, focus was given to recruit bilingual Sinhala speakers. A main reason for this was because majority of KCG staff & members of the PMP local research team having no proficiency in Tamil. It is worthwhile noting that finding Tamil native speakers bilingual in Sinhala is a quite difficult exercise. Sinhala native speakers with working proficiency in Tamil are almost non existent creating a high demand for the existing few Tamil bilinguals.

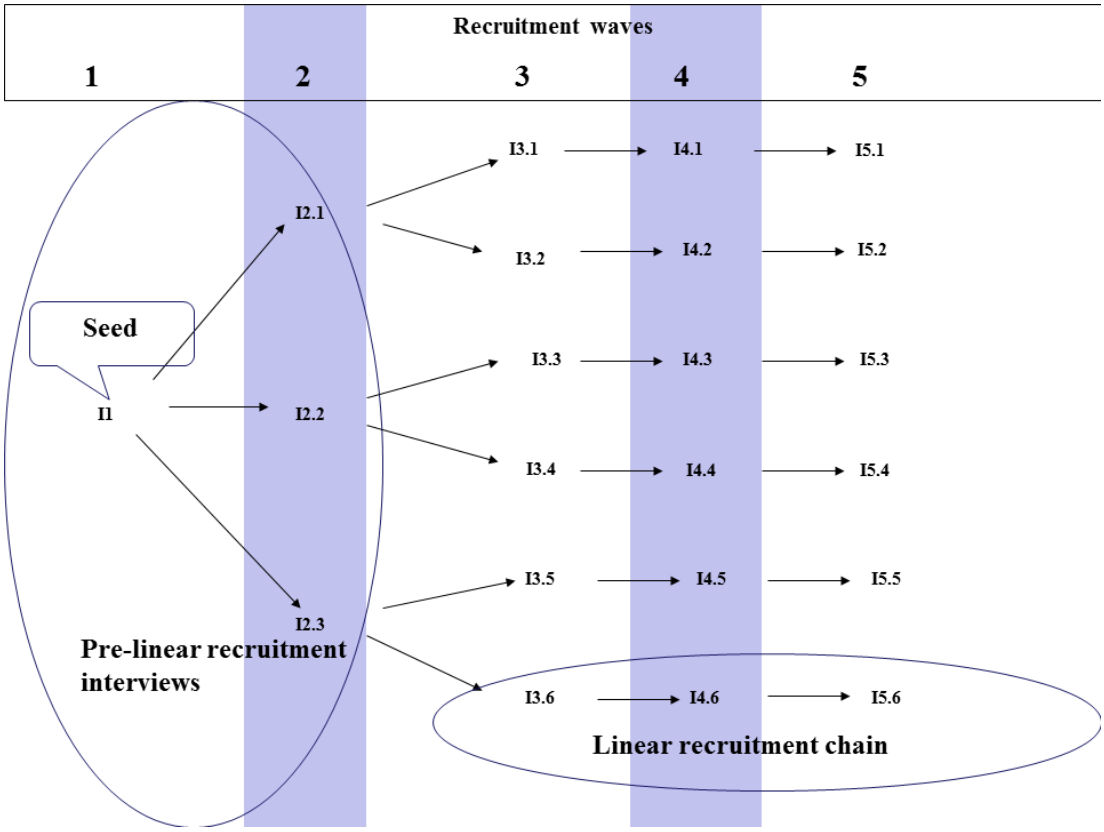
The basic sampling unit used within this pilot survey was a network cluster defined by the central PMP team as a set of 22 target interviews which span across 5 successive recruitment waves. The cluster starts with one single seed within the initial recruitment wave. Then the seed introduces three more respondents during the second recruitment wave and the network continues to grow through referrals up to a saturation point of 22 completed interviews by the end of the fifth recruitment wave. Figure 3 clearly depicts the proliferation of the network cluster throughout the five recruitment phases.

Both the seed as well as subsequent network members had recruitment criteria to satisfy prior to have been enrolled within a particular cluster. These criteria were well defined prior to the implementation of the pilot project by the PMP central team. For seed respondents, compliance with the following criteria as far as possible was essential:

- Being aged 18 years or more at the time of the survey
- Being living in the local area for the last 5 years
- Being interested in the topics of the survey, feel comfortable to take part in the interview, and display a cooperative attitude
- Being a frequent mover throughout the local area, and have acquaintances at different location in the local area
- To be having an extended social network (in order to facilitate the further recruitment process)

- To not be personally acquainted in any way to the field coordinator or any other member of the field team
- To have no role as a formal leader within the local area (e.g., as a religious leader or political representative)

Figure 3 – A Basic sampling unit within a local area (a cluster of 22)



Following the selection of a seed, respondents throughout sampling iterations 2 to 5 were selected among eligible network members of respondents from the previous recruitment wave. A precise procedure defined by the central PMP team controlled such recruitment. During a survey interview, respondents are asked to mention all individuals with whom they've talked about the past. After noting down names of such individuals, eligible network members are selected based on the following criteria, which automatically becomes the recruitment criteria for all non-seed respondents within a cluster. They are:

- Network members who, according to the referring respondent, are aged 18 or more at the time of the survey
- Network members who live in the same local area (such as defined under 1.1.)
- Network members who speak the same native language as the current respondent

*(This criteria was enforced due to limited budgetary allocations resulting in one particular enumerator speaking one native language continuing to interview the entire cluster)

Once eligible network members have been identified, they are numbered from 1 to a maximum of 18, out of which 3 of them are randomly selected. The random procedure used within the pilot study was drawing lots. Precise instructions on this

procedure were provided within the network part of the pilot survey questionnaire.

Once drawn, enumerators ask permission from the respondent to contact the 3 prospective eligible network members by recording precise contact information of them such as their names, addresses, phone numbers, and/or a detailed description of the place and time when and where they can most easily be reached. The rank of the listed target participants corresponds to the order in which they have been randomly drawn (by drawing lots).

The budgetary allocations for the pilot study encompassed resources only for a single enumerator to be working within a particular DS division. Since all enumerators were native speakers of one local language, it was decided based on feedback from the field that, in the event a network member speaking a language that is foreign to the enumerator is been recorded, he/she is authorized to request the respondent to substitute the name with another network member speaking the same native language as the enumerator.

In order to ensure confidentiality, as per procedure, contact details were recorded only once, on a separate contact sheet, which is related to the questionnaire only by way of an anonymous code. Field staff members were not allowed to copy contact details, nor create any electronic files including contact details. Unlike anonymous questionnaires, contact details of respondents are to be destroyed at the end of fieldwork. Until then, they are securely stored.

The total number of targeted survey interviews budgeted within Sri Lanka's PMP pilot project was 400. This number of 400 was subdivided to include 200 interviews each from the districts of Ampara & Matale. As mentioned in section 1.1 these areas represent diversity in terms of impacts of the ethnic war. Within each district the 200 targeted survey interviews would equally comprise of 100 Sinhala questionnaires and 100 Tamil questionnaires. By doing so it was intended that the diversity among the different ethnic groups living within these areas would be captured. These would then be subsequently distributed among the 3 Sinhala and 3 Tamil enumerators resulting in each enumerator being responsible for approximately 33 survey interviews per district and a maximum of approximately 66 interviews within the entire pilot survey project. The reason for ensuring that each enumerator received a minimum of 66 interviews (and a maximum of 68) was due to uncertainties involving the time taken for network proliferation. With such uncertainties it was rational to base enumerator expenditures based on piecework rather than a weekly or monthly payment. The breakdown of the 400 survey interviews among the six enumerators within the two districts based on the ethnicity of the respondents (targeted through language use) is displayed through figure 4.

By providing each enumerator an approximate number of 33 survey interviews per district a previous pilot sampling gridline was violated. This guideline stated the following:

> - No interviewer should carry out alone more than half of the total number of interviews by cluster, i.e. the maximum of full recruitment chains to be attributed to one interviewer within one cluster is 3.

- No interviewer should carry out more than half of his overall personal workload within one single cluster. The personal workload of an individual interviewer should not exceed 50 interviews in total."

However, due to budgetary restrictions involving the pilot survey, in ensuring that the costs were minimized a piecework payment method for enumerators were introduced. In ensuring a sufficient number of interviews per enumerator, the number of enumerators recruited and trained was reduced to six. In keeping with time and travel cost constraints it was agreed upon with the central PMP team to provide each enumerator with an approximate number of 33 interviews per district overriding the previous two sampling guidelines.

In the case of an instance of non-cooperation or non-eligibility of a respondent within a cluster the following procedure was to be adapted according to PMP central team instructions. In the case of the 3 target respondents at iteration 2 were to be replaced by a new seed. The new seed respondent would now be a part of the previous cluster and replaces the missing respondent at iteration 2 in the cluster's branching scheme. The reason for this is because the loss of any one respondent at iteration 2 would result in the non-inclusion of 6 further respondents resulting in a substantial decrease of the final sample size. Instances of non-cooperation or non-eligibility at any other recruitment iterations would not be replaced due to low impact on reduced sample size. However the PMP pilot survey project in Sri Lanka recorded zero instances of non-cooperation or non-eligibility due to high skills and commitment of the survey field team.

Because minimum targets of 33 survey questionnaire interviews were set for each enumerator per local area, per district, under ideal circumstance (zero instances of non-cooperation or non-eligibility) it was expected that each enumerator would complete 1.5 clusters per district (22 survey interviews per cluster + an additional 11 in another cluster). However, since data collection in the field was time bound for a period of 4 weeks and the uncertainties involved with network proliferation, enumerators were provided with more freedom in terms of completing the 33 survey questionnaire interviews. For instance they weren't required to be linear in their approach in completing each cluster. They were given the opportunity to finish off the target of 33 interviews as soon as possible within a single local area in one district and move on to the next local area in the next district based on respondent availability. This meant that some enumerators would be linear in their approach completing 1 cluster and then moving on to the next, having 2 uncompleted clusters (i.e. 17 to 16 completed interviews in each cluster pursued on respondent availability), or having 3 or more than 3 uncompleted clusters (due to non-cooperation and or non-eligibility). However as mentioned previously the survey field team were highly engaged and persuasive with respondents and performed exceptionally well resulting in zero recorded instances of non-cooperation or non-eligibility. However, unfortunately they did not accurately record the number of attempts made in making contact with a particular respondent and also the number of contacts made before completing a survey interview, regardless of this being communicated during enumerator trainings. Also the lack of an option such as "targeted interviews completed or cluster completed" motivated enumerators to mark participants who did not take part in the survey as 'No contact could be established' (under 'Outcome of visit') within the contact sheet. This would ideally pose a problem in distinguishing between

respondents who were not contacted at all from those who were contacted but were not reached. However since zero recorded instances of non-cooperation or non-eligibility were achieved, instances marked as 'No contact could be established' should be counted as respondents who were not contacted.

2. Survey questionnaires and calendars

2.1. Translation

Enumerators were instructed to carry with them the following sampling documents and material to the field when engaging respondents.

1. Survey questionnaires and answer cards
2. Life calendars printed on A3 format
3. A complete inventory and country map of local areas
4. Guidelines for field staff in approaching survey respondents
5. Contact sheet (template provided in Appendix 1)
6. Numbered lots to be drawn for random selection of network respondents
7. A coin to be tossed for random selection of questionnaire version to be administered

Out of them, it was the responsibility of the local PMP project team to provide local versions of sampling documents numbered 1 to 5. The initial versions that were handed down to the local PMP survey project team were in English.

The survey questionnaire included two versions. Version 1 comprised of 4 sections, which included questions on life events and personal information, network members, vignettes and respondent living conditions, community life, community leadership and identity. It contained a total of 141 items. Including all answer cards and the life events calendar, it spanned across 25 pages. The striking difference in Version 2 was the omission of the life events calendar. However this 25 paged questionnaire included 177 items across 5 sections which sought responses for personal information, network members, vignettes, community life, community leadership, identity, transitional justice, collective victim beliefs, reconciliation oriented development and collective action. The fifth section on victim beliefs, reconciliation oriented development and collective action, was developed by the quantitative doctoral researcher working for the local PMP survey team.

Translation of these documents posed a huge challenge on the local PMP project survey team due to the following reasons.

Firstly, as mentioned previously, there are two official local languages (Sinhala & Tamil) that are spoken by the major ethnic groups in Sri Lanka. This required the translation of the English versions of the documents into both Sinhala & Tamil and doing a back translation and comparison of the versions in order to ensure that both the versions communicated to respondents what was originally meant.

Apart from the translation, contextualization of the questionnaire posed a huge challenge. Major sections of the questionnaire were previously modelled and tested within a European context. They were complex, cognitively challenging, long and had

items ranging throughout a diverse set of fields requiring a clear contextual flow to be maintained throughout the length of questionnaire administration, which in some cases exceeded two hours. A simple direct translation of the original questionnaire would have excluded a lot of contextual meaning and would have lost a lot of validity in terms of some of the important and subtle constructs to be tested. It also ran risks of questionnaire rejection due to loss of interest and possibilities of interviewee fatigue. Contextualization needed to have been ensured between both the Sinhala and Tamil versions.

Appropriation of the questionnaire in terms of translation and contextualization involved detailed explanations of questionnaire sections and items in a more vernacular tone. This appropriation came with certain restrictions in terms of the number of pages and formatting of the questionnaire. Formatting required being similar across the three sites (Burundi, Palestine and Sri Lanka) and the high costs involved in printing in certain research sites resulted in the restrictions in questionnaire formatting.

A difficult balance had to be maintained by the local PMP survey team between a contextualized questionnaire that was more interesting, engaging and made sense to the respondents within the set restrictions of page formatting & numbering. A practical solution adopted by the team was to have a translation which was a more formal / written version of the questionnaire which couldn't be directly read to respondents. During training sessions, the enumerators were provided with a thorough training, which included adequate examples and scenarios in converting this formalized translation into a meaningful vernacular version during the survey interview. In future questionnaire design if restrictions are lifted in terms of number of pages and larger font sizes (the current usage of 2 columns), having lengthier vernacular versions of items in print may be more feasible.

Due to issues involving government scrutiny and related issues of security, risks associated with leaking of the PMP project's private in-house material, the lack of professional translation services that ensure privacy and confidentiality related to the above 2 factors and most importantly the difficulty in finding translators having knowledge of context in terms of the socio-political landscape of the country, the translations were performed in-house by the project coordinator and two PhD students (quantitative and qualitative). Having two PhD students conversant in each of the two native languages proved advantageous in this regard. Even though major proportion of the questionnaires were being translated by the local PMP project survey team, due to high work loads and time line constraints, certain proportions of the Tamil questionnaire translation needed to be handed over to a 3rd party. This 3rd party was discovered through known contacts and came highly recommended by the International Centre for Ethnic Studies the local organization hosting the project in Sri Lanka.

2.2. Difficult items and procedures

Following translations and thorough back translations and comparisons of the two translated versions of the questionnaire, the initial training for the enumerators occurred on the 27th and 28th of July 2015. Even though it was anticipated to be a

conventional training with members of the PMP survey team conducting the training with instructions and explanations on presenting each item of the questionnaires to the respondents, it eventually unfolded it self in the form of a focus group discussion between enumerators, KCG staff and members of the local PMP survey team. Experienced field staff exchanged valuable insights and decisions were made in terms of modifying or discarding items that would pose difficulties within the field, hampering successful data collection and threatening the security of field staff.

Items that would cause respondents to feel as if they were undergoing an interrogation, items narrowing respondents thoughts to a particular group when responding, items posing unrest and unwanted suspicion among different parties concerned within the geography of survey administration etc. were decided to be excluded. It was decided by the local team that items on war crimes on questionnaire version 2 ranging from 119 to 122 and 126 (e.g. Do you believe that war crimes were committed by the other side during the conflict?) to be removed as their inquiry or printing would pose unwanted problems in survey administration. Apart from the above, question number 23 in version 1 and question number 15 in version 2 on political affiliation was removed as the local field team considered it to be too sensitive to be asked provided the local political landscape and general elections being due on the 17th of September 2015.

Item number 11 in both questionnaire versions 1 and 2 (Did you become a combatant, carrying a weapon?) and item number 12 in version 1 (Did you use a weapon in the fighting?) were modified to prevent any implications mentioned above. They were modified as "Have you ever carried a weapon during a violent conflict?" and "Have you used a weapon during a violent conflict?".

During the focus group discussion it was agreed that in both versions questionnaire instructions for items pertaining to 'violent conflict' referred events involving the military only (i.e. In the following questions, 'violent conflict' refers to war events, military occupation or riots.). It was decided to change this bias in order obtain responses in a more equitable manner and to avoid emotional arousal from patronizes of the military (i.e. In the following questions, 'violent conflict' refers to war events (i.e., events of violence that are not interpersonal in nature).

For items 102 & 105 on public apologies in version 2, instructions written within parenthesis (perpetrated during the conflict) were decided by the field team to be discarded to avoid any unnecessary arousal of emotions. The idea was to prevent the shifting of the interpretation of these items from an apology to a conflict situation.

In questions 101, 102 & 104 of questionnaire version 2 (pertaining to public apologies) in order to avoid a situation where respondents might feel as if they have been forced into a category (i.e. a 'them versus us' situation), which might prevent people from responding due to fear etc. the field team decided that instead of both sides, the word all parties should be used (i.e. did the Commission adequately examine potential human rights violations by all parties involved in the conflict?).

In questionnaire version 2, it was raised by the field team that item 100 gave a fairly good explanation of what a truth commission is. However later it asks whether the respondent is familiar with a truth commission that occurred locally. It was suggested that if the respondent was unfamiliar, it would be illogical to ask question

101 evaluating the effectiveness of this truth commission. It was suggested that the respondent to be directed to the next item (102) relating to public apologies instead.

It was the view of the field team that the word 'group' within the instructions as well as the items numbered 74 to 85 was a bit limiting. It was thought that the word 'group' might refer to militant groups or the army more often than not and might limit the responses. Hence it was decided that the word 'ethnic group' instead of merely 'group' would broaden the definition of the group which might increase the willingness of participants to relate to and answer.

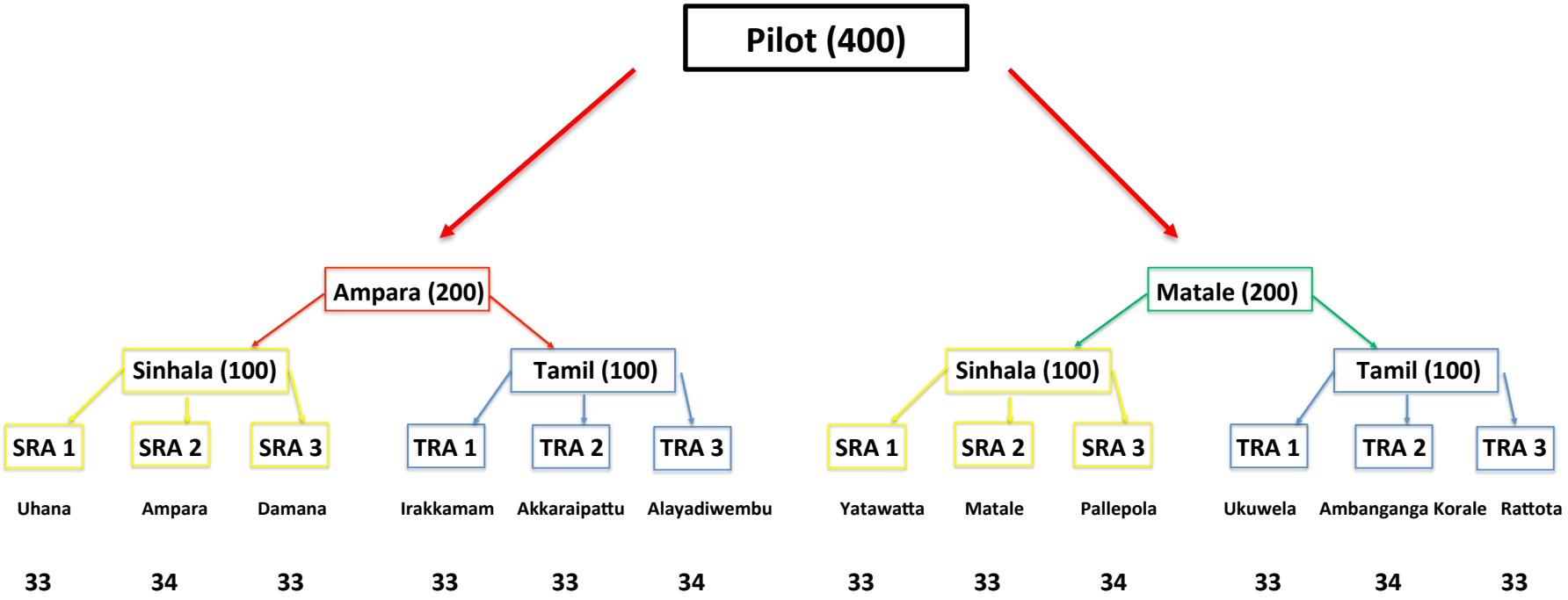
Question 7 in version 1 and Question 10 in version 2 inquired information from respondents about immediate family members who have been killed during violent conflict. The field team felt that respondents whom have immediate family members who have disappeared as a consequence of violent conflict may be neglected hence an opportunity to obtain valuable information unexploited. As a consequence an additional item was added to both versions inquiring about immediate family members who have disappeared as a consequence of violent conflict.

The local PMP survey team suggested the inclusion of media source type in addition to the name of the media source to item number 17 in questionnaire version 2 in order to make better sense in data analysis.

Furthermore, the need for a contact information sheet for seed respondents as well as the limitations of using three dice in randomly selecting three network members out of 18 (number 1 & 2 are never selected) was pointed out by the local PMP survey team.

Finally this whole team effort resulted in the reprinting of the questionnaires and a second training, which occurred on the 06th of October 2015 prior to the enumerators beginning fieldwork on the 10th of October 2015.

Figure 4 – A graphical representation of the sample breakdown across districts, among local areas, respondents and enumerators speaking different native languages



* Two Clusters (22 + 11) for each enumerator within a single DS division

3. Fieldwork

3.1. Fieldwork network

1. Pluralistic Memories Project's Local (Sri Lanka) Survey Team

Dr. Ramila Usoof – Co-investigator

Dr. Ramila Usoof is a co-investigator as well as a member of the steering committee of the Pluralistic Memories Project. She is also the anchoring figure coordinating and overseeing scientific research conducted through the project here in Sri Lanka. Dr. Usoof is also the main point of reference in communications taking place between the Kandy Consulting Group the survey agency in charge of field work in Sri Lanka, the PMP central team as well as the local Sri Lankan PMP survey team.

Sumedha Jayakody – Doctoral Researcher

Sumedha Jayakody is the quantitative doctoral researcher working for the Pluralistic Memories Project in Sri Lanka and coordinates the operational aspects of the project's longitudinal survey. Being a native Sinhalese speaker, Mr Jayakody performed all Sinhalese translations of the pilot survey and was in charge of conducting trainings for field staff regarding survey instruments as well as the survey methodology. Mr Jayakody was also involved in pilot survey debriefings and is also the author of this methodology report.

Esther Surenthiraraj – Doctoral Researcher

Esther Surenthiraraj is the qualitative doctoral researcher working for the Pluralistic Memories Project in Sri Lanka. Being a native Tamil speaker, her involvement in translating and back translating pilot survey documents, as well her role in mediating translations during survey trainings were extremely beneficial for the successful implementation of the PMP pilot survey in Sri Lanka. Ms Surenthiraraj's expertise within the field as well as her role within the project's testimonies collection made her a valuable asset in focus group discussions in improving survey documents.

Harini Dias Bandaranayake – Communication Officer

Harini Dias Bandaranayake is the communications officer within the local PMP team and played a major role in establishing links with trusted 3rd party translators.

2. Kandy Consulting Group (Pvt) Ltd. (KCG)

Senior Project Manager

The senior project manager of KCG overlooked the overall project and was responsible in the implementation of the field survey on behalf of the PMP local team.

KCG Research Assistants

Two KCG research assistants working on the project were responsible in preparing filed documents, keeping track of field work updates, coordinating with filed supervisors, field enumerators, data entry operators, and staff involved with STATA cleaning. They also were the main point of reference for KCG with regards to the PMP survey providing communication links between the local PMP team and the administration of KCG.

Field Enumerators

There were three native Sinhalese speaking and three native Tamil speaking enumerators that were used in both Ampara and Matale Districts. Initially they stared work in Ampara and then moved to Matale. Enumerators were to select respondents appropriately according to given instructions and complete survey interviews. They were to explain well about the study and obtain consent before initiating interviews. They should answer all necessary questions from respondents and provide background information. They were to take down additional information as notes. They had been instructed by KCG to complete all questionnaires with a BLUE pen and include signatures where required prior to submitting it to a field supervisor for further comments and checking. Once an interview was completed they were to offer the token gift as an appreciation of the respondent's valuable time and effort.

Filed Supervisors

Two field supervisors one each for supervising the Sinhalese and Tamil speaking enumerators were utilized by KCG. They too started work in Ampara and then moved to Matale. The main responsibilities of the supervisors were to oversee field enumerations, guide enumerators in selecting seeds / respondents during the initial stages of the pilot and constantly ensure enumerator adherence to given guidelines. In a nutshell they were to ensure that enumerators fully understood the network sampling methodology and adhered to it. They performed a 100% check on questionnaires submitted to them by enumerators. They were to randomly check on field attendance and timings of completed questionnaires. Once a questionnaire is checked 100% it was their task to deliver it to KCG's project office. Within their 100% check, they look for clarity, cleanliness and completeness. In instances where errors are found they note them down with BLACK ink and send them back to the enumerator for corrections. All of these are done based on precise deadlines.

Internal Research Assistants

A number of five internal research assistants were utilized by KCG for printing, gathering and distributing survey documents during the initial stage. They engaged in questionnaire scrutiny and coding and also transcribing and translating comments made in Tamil to Sinhalese during data entry. After being scrutinized in the field by field supervisors, internal RAs in the project office usually do the re-scrutinizing for completed questionnaires along with numbering and coding. In this pilot survey the RAs had to do the additional work of transcribing comments made in Tamil to Sinhalese. A native Tamil speaker (ideally a supervisor or an enumerator) who is familiar with the instrument read and explained the meaning in Sinhalese and the Sinhalese RA re-wrote the text on the instrument itself to ensure data entry operators understood the completed questionnaire. This complication arose due to the unavailability of data entry operators who were competent in both excel and Tamil typesetting within the given time constraints. Hence in order for the native Sinhalese speaking data entry operators to understand and convert the total data set into one language, this additional step of transcription was introduced. But in future survey rounds, KCG hope to avoid this step by recruiting Tamil typesetters well in advance. Finally during the last stages of the pilot survey they also performed phone calls for back check interviews.

Data Entry Operators

Once questionnaires had been checked by filed supervisors and scrutinized by internal RAs, they were handed down to five data entry operators. They were provided with excel templates and codebooks with half a day's training conducted by the quantitative doctoral researcher of the local PMP team. They were responsible in entering data from field instruments following given instructions. Once completed they were to provide a soft copy in the form of an excel document after a 100% recheck of their entries. Subsequently this excel document would be sent for STATA checking.

3.2. Fieldwork timeline

Time Period	Milestones
2015	
February	Handing over of the survey instrument to be translated and tested (20th February 2015)
March	-
April	-
May	Translation of the 2 versions of the questionnaire & printing
June	
July	Training for enumerators prior to field work (27th & 28th July 2015)
August	Amendments to questionnaire based on feedback & reprinting
September	
October – November	2nd Training for enumerators following focus group discussions prior to field work (06th October 2015) Field work – Ampara (10th to 20th October 2015) 1st debriefing – Ampara (15th October 2015) Field work – Matale (24th October –15th of November 2015) Combined field visit by PMP local team & KCG staff (10th November 2015) 2nd debriefing – KCG (30th November 2015) Submission of all completed questionnaires (30th November 2015) Data entry training (17th November 2015)
December	Data entry (17th November to 23rd December 2015) Data set (first draft Sinhalese only - 08th December 2015) Data set (first draft Tamil only - 18th December 2015)
2016	
January	STATA checking (24th December 2015 to 06th January 2016) Data Set (cleaned final data set - 06th January 2016) Sorting & storing hard copies (19th January 2016) Data Set (corrected final data set – 22nd January 2016)

3.3. Interviewer selection and training

Interviewer selection

Two field supervisors (one native Sinhalese and one native Tamil speaker) and six enumerators (three native Sinhalese and three native Tamil speakers) were recruited by KCG for data collection. When recruiting enumerators for academic research related fieldwork, KCG maintains a minimum educational qualification for a recruit to be either an undergraduate, graduate or a postgraduate student. This condition was satisfied when recruiting Sinhalese speaking field staff. However, concerning Tamil speaking field staff, KCG had to lower the education criteria for a recruit to have passed the GCE Advanced Level qualification. However, the above was compromised by having Tamil speaking recruits who were bilingual and with high levels of field experience in similar capacities. KCG had used its professional links that it had developed within its past 10 years in finding candidates matching its stated criteria. For example the Tamil speaking supervisor is a professional translator and a social mobilizer with many years of experience in non-governmental organizations (NGOs) working in areas such as peace promotion, human rights, women rights and empowerment. One Tamil speaking enumerator was a social mobilizer and a teacher of Tamil language for native Sinhalese. The Sinhalese supervisor and enumerators were similarly with many years of field experience working with KCG and elsewhere. The project was greatly affected when it lost three experienced enumerators (two Sinhalese and one Tamil speaking) due to the delay in the starting of fieldwork.

Interviewer Training

Following a rigid translation process, completed questionnaires and other sampling documents were printed by KCG to be distributed among enumerators. Due to the complexity involved with the questionnaire, its contextual appropriation as well as the sampling methodology, the PMP local team in close collaboration with Kandy Consultancy organized a comprehensive training prior to the enumerators beginning work in the field. The survey instrument being worded in a formalized manner required adequate skills and examples to be inculcated by the enumerators. Additionally the fact that the training was to be conducted among both Sinhala speaking as well as Tamil speaking enumerators was another requirement for such a comprehensive training which resulted in it occurring on the 27th and 28th of July 2015. The logistical arrangements for this training were borne by KCG at its office premises in Kandy. Participants for the training included administrative staff from KCG handling responsibilities pertaining to the PMP's longitudinal survey, PMP local team members and enumerators and supervisors hired by KCG for the pilot survey. The training was conducted in Sinhala. KCG had hired bilingual native Tamil speakers as enumerators to work in sampled areas, which were predominated by Tamil speaking respondents. Each item and instruction was read out in Sinhala and the enumerators as well as the supervisors would acknowledge their understanding or provide feedback in terms of their interpretation or any other complication that might arise when asking such items / giving out such instructions to respondents within the field. The supervisor of the Tamil speaking enumerator team functioned as the main

translator and members of the local PMP team who are native Tamil speakers were able to facilitate the translation process.

Even though the training was anticipated to be quite straightforward where items and instructions would simply be explained to the field team, it turned out to be a focus group discussion where valuable feedback was provided by experienced field staff. Decisions were made in terms of modifying or discarding items that would pose difficulties within the field, hampering successful data collection and threatening the security of field staff.

Once modifications were performed on both Sinhalese and Tamil versions a second one day training on the questionnaire was organized in an identical manner to the first. This occurred on the 06th of October 2015 prior to the enumerators beginning fieldwork on the 10th of October 2015. Even though KCG had lost 2 native Sinhala speaking enumerators and 1 native Tamil speaking enumerator due to permanent employment between the time period of the two trainings, the new substitutes were provided with the questionnaire well in advance to the second training in order for them to prepare themselves for this one day intensive training. Both trainings were conducted by the quantitative doctoral researcher of the local PMP survey team.

3.4. Interviewer supervision and back-checks

Interviewer / enumerator supervision

A huge effort was made by KCG on enumerator and supervisor supervisions. Zero tolerance was shown over fabrication of data or any other misconduct in the field. In order to ensure that the enumerator is physically present in the field and conducting interviews with a stated respondent, rigorous checks were being implemented. Enumerators and supervisors often lodged together in the field, or met regularly. During the initial stages, supervisors accompanied enumerators and were fully involved in seed selection. The supervisors were physically present during the first couple of interviews conducted by each enumerator. By maintaining such close proximity, the supervisors were able supervise enumerators, ensure security and gain valuable insights to various issues within the field.

Once a questionnaire was complete, it was 100% checked by the supervisor and then sent to KCG's project office. In the case of an error such as a missed response, the enumerator was required to get in touch with the respondent by revisiting or by telephone. After the questionnaires were being submitted to KCG internal RAs or research assistants at KCG scrutinize them for potential errors. If errors were revealed, the process of enumerators getting in touch with respondents was initiated. During the pilot survey KCG reports that they hadn't come across any incident where enumerators had been sent back to the field to make amendments. However, KCG states that there had been many instances where incomplete questionnaires had been completed by contacting respondent over the phone. Enumerators had made several visits to KCG's project office in order to complete and clean their questionnaires during office scrutinizing and data entry phases.

Supervisor supervision

KCG had constantly been in contact with supervisors. They had verified the activities of the supervisors through their daily logs. These logs had been crosschecked with enumerators to verify whether adequate time had been spent within the field supporting enumerators in building their confidence to start working independently. KCG had also compared daily logs between enumerators and supervisors to check whether the entries match.

Random checks over the phone (Back-checks)

At least 10% of randomly selected respondents were given a phone call by KCG from its project office in order to ensure whether the interview genuinely took place. Some of the items that were checked included "Date of birth", "Name of the GN division" and the "Most frequently followed media source". Furthermore, respondents were asked about any difficult questions they experienced within the questionnaire, whether they had received a token of appreciation and whether there was any inconvenience caused due to the survey interview in general. Since the supervisor in the field did a 100% check of the questionnaire including validating the implementation of the sampling procedure, no requirement was seen to redo it over the phone.

Some respondents had commented that they had been inconvenienced due to the extensive time consumed by the survey. Some had wanted to know whether they could gain access to any publications resulting from the information collected. KCG declares that not a single incident was reported on the absence of a physical interview or the delivery of the token of appreciation.

The process utilized by KCG for the random checks was as follows. Respondent ID's (rid's) of questionnaires completed by each enumerator were listed according to each district/DS division. A random number was generated for each rid and subsequently sorted according to descending order (largest to smallest). Then the rid's corresponding to the highest three numbers were called. If they weren't contactable i.e. not having a telephone or not answering the phone even after three attempts, the next respondent with the highest random value assigned would be selected. This process was repeated until 10% or 3 respondents per enumerator per district/DS were contacted.

STATA checking

Once data were entered and rechecked 100% by the data entry operators, the data sets and the questioners were passed on for STATA cleaning. The STATA cleaning research assistant has to merge all the data sets that are received from different data entry operators in order to make a single data set for each version of the questionnaire. This also includes the merging of different excel sheets within each data set from each data entry operator. Once a comprehensive data set has been generated for each version of the questionnaire, they are imported to STATA. The responsibilities of the STATA cleaning research assistant is as follows.

- The random stepwise inspection of variables for at least 10% of the entries made by each data entry operator.
- Ensuring the inclusion of the total number of questionnaires completed within the data set.
- Ensuring the sequence of responded IDs, and other numberings.
- Testing the logical flow of the questionnaires between variables and sections with identified as well as critical check points provided by the research team.
- Provide one cleaned data set for each questionnaire version using STATA in the form of a .dta or excel file as requested.

Interviewer Debriefings

Two interviewer debriefings were performed during different stages of data collection by the quantitative doctoral researcher with the main objective of investigating problematic issues that arose with regards to the content as well as the process of data collection. Debriefing sessions were also used as a platform for sharing best practices among enumerators. The initial debriefing occurred on the 15th of October 2015 which was during the 2nd week of data collection in Ampara. All members of the field team except for one enumerator were present at this debriefing session, which occurred as a group session at a public park in Ampara. The second debriefing occurred on the 30th of November 2015 at KCG office premises in Kandy. The enumerators had completed data collection at this point and were to hand over their competed questionnaires on that day. The debriefing meeting was organized as a final get together of the pilot field staff with a lunch reception at the end of the session. Valuable insights were obtained from both debriefing sessions, which are shared in detail within section 3.5 Difficulties and unexpected events.

3.5. Difficulties and unexpected events

Pre-fieldwork

Adhering to precise schedules and deadline during the pilot study was problematic due to the following complications. The novelty as well as the complexity involved with network sampling to the local project team as well as the survey agency resulted in a lot of time consumed in understanding the methodology as well as agreeing on budgetary allocations. Translating and contextualizing the questionnaires into two languages as well as getting them typesetted through trusted and competent sources required considerable amounts of time. Feedback received during the initial training session for the field staff resulted in modifications, a reprint of the questionnaires and an additional training. Within this time span the field team lost three of its enumerators due to receiving permanent employment elsewhere. The volatile political environment of the country during August 17th parliamentary elections and other reasons mentioned in detail resulted in the postponing of fieldwork from the initial planned 24th of April 2015 to its actual implementation on the 10th of October 2015.

Fieldwork

1. Issues pertaining to emotional trauma and interviewer / respondent fatigue

A unique finding from the field with respect to Ampara was that respondents often became very emotional with regard to collective marker events and vignettes. For many, personal marker events were irrelevant and spent a lot of time talking about collective events. The enumerators were faced with difficulty as it took on average two hours or more to complete a single questionnaire. They also found witnessing the re-traumatization of respondents somewhat difficult to bear. The strategy discussed within the team was not to interrupt respondents but let them ventilate their emotions. They were provided with an option if interested in sharing their experience in detail with the project's testimonies collection. In situations where respondents were over enthusiastic about sharing their experience, they were courteously reminded that the survey wasn't the opportunity, but to testify instead if interested. Hence the re-traumatization of respondents, traumatizing effects on enumerators and the competency in responding during trauma such as consulting emergency services were some concerns that arose from the field especially within areas that were directly affected by the war.

In relation to the issue of additional time taken in responding to the survey, there have been instances especially in Ampara where respondents have been reluctant to participate due to the sensitive nature of the survey. They have had bad experiences in the past where once they've responded, they've been subjected to harassment by law enforcement. In such instances of suspicion, the enumerators have experienced that the more they engage with the respondent, the more time they spend befriending and empathizing with them, the more likely they are to change their decision and participate in the survey.

2. Validity issues in relation to various interpretations associated with the survey instrument

At the onset of the survey, the introduction made to the respondents as well as the general nature of the areas studied by it seems to be creating an impression or a mind set within respondents that all responses need to be conflict related. There have been instances especially in conflict affected Ampara where respondents have replied when asked to mention personal marker events: "Do they have to be events which brought chaos to our family?" The response anticipated are events which are important to the individual not necessarily chaotic.

Another interesting example for respondents developing a certain mind set towards the questionnaire was evident among respondents from Matale. Even though Matale wasn't directly affected by the war, it was a district, which experienced a lot of violence during the Janatha Vimukthi Peramuna (JVP) insurrections. Formed in the 1960's by radical Marxist, the JVP recruited students and economically deprived youth from rural areas to revolt against the government during the early 70's and late 80's. The group terrorized the state machinery as well as anyone opposing its

ideologies by resorting to subversion, assassinations and raids made on military and civilian targets. Even though the state was successful in overcoming these insurgencies, these were times of increased brutality and extrajudicial killings. Amidst living through such difficult times, only a few respondents from Matale elaborated on their experience of JVP insurgencies. One reason for the lack of reference to the JVP era might be the chronological distance of such events in comparison to the ethnic war. However a few respondents had expressed their opinions after the end of the survey mentioning that this questionnaire was more suitable to the Northeast as Matale was never directly affected indicating their perceptions that the questionnaire was expecting responses related to the war. Again factors such as the introduction of the survey in a manner which is biased towards the ethnic conflict, items relating to transitional justice and vignettes including example of events which occurred only during the ethnic conflict might be probable reasons for developing such a mind set among respondents. The enumerators were also concerned that apart from the life events calendar, the questionnaires especially version 2 of the questionnaire had no space for respondents to share their experiences about the JVP insurrections. This might be an important point worth considering as many regions of the country such as the south which had little influence during the ethnic war were greatly affected during the JVP insurrections and the failure to capture them would be a huge loss in terms of valuable data.

Certain questions within the survey instrument seemed to be cognitively challenging and posing ambiguity in terms of interpretation. For instance the vignettes contain questions which require respondents to perceive a situation where statements of events related to conflict are being made by a member of the respondent's own ethnic group in front of a crowd i.e. question 26 in questionnaire version 2. The respondents is subsequently asked how most others in such a situation would react to the statement made. Respondents have had difficulty in interpreting the context of the 'others'. For instance respondents have informed enumerators that if the 'others' were from Ampara, they would consider the event to be true and quite frequently occurring. However if the 'others' were from outside of Ampara, they wouldn't believe the event to have been that frequent. Similarly the respondents had experienced difficulty in interpreting question number 25.3 & 25.4 (and their repetitions). A national politician or foreigner from Ampara would have experienced such events first hand but not someone outside of Ampara etc.

Within the section on identity characteristics, the two characteristics religion and ethnicity aren't far apart for respondents with a Muslim ethnicity. Hence enumerators have had difficulty asking and respondents responding in instances where both religion and ethnicity have been selected as important identity characteristics where the same questions are repeated for both characteristics.

The enumerators stated that they weren't quite sure behind the thinking of the respondents to the questions on 'the feelings they had experienced throughout the past two weeks'. For instance respondents would state that they weren't feeling very energetic and motivated because they were sick during the past couple of weeks. This could have easily resulted due to the sickness preventing the respondents from participating in their farming activities causing them to miss the season, costing their entire earnings for the next six months.

The enumerators were also a bit concerned that participants from Matale weren't grasping the meanings conveyed through the vignettes in a holistic manner. Hence were sceptical about the accuracy of their responses given to vignettes. Some would internalize partial statements such as 'I hid in the woods' and state "oh, we also hid in the woods during JVP times" and would respond to the vignette which is an event that took place during the ethnic conflict which is erroneous.

Some closed communities living in isolation such as the tea estate communities living in the mountainous regions of Matale have extremely low health, nutrition and education indicators. They find some items quite difficult to comprehend. On the other hand the secluded lifestyles they maintain results in them having an indifferent attitude towards some of the concepts discussed in detail within the survey (such as conflict and transitional justice).

3. Difficulties experienced by field staff in relation to network proliferation

The social dynamics of ethnicity and gender between the respondent and the enumerator seems to play a major role in determining the success of an interview outcome. This is especially prominent in rural closed communities that are rarely open to strangers. The social cultural norms are more conservative in such settings where woman working, travelling alone and interviewing respondents of an opposite gender might be perceived erroneously. There have been instances where female enumerators have been strictly advised by locals not to engage in survey activities and to adhere to particular dress codes. Similarly, male enumerators successfully surveying female respondents in a private setting is highly unlikely in such closed communities requiring a gender balance in the availability of enumerators for a particular area or sampling cluster. Under such conditions, female enumerators may face limitations in terms of the number of working hours available for them affecting the pace of data collection and network proliferation. Enumerators had taken notes of these instances and revealed them during field debriefings. In the future design of questionnaires it would be advisable to have a section on interviewer impressions where they could make spontaneous notes.

Similarly, the ethnicity of the enumerator seems to play a major role on the interview outcome. There have been instances where respondents have refused to answer the questionnaire due to the presence of a field staff member belonging to a different ethnic group. Differences in ethnicity are apparent in language use (accents) and dress code. Hence similar to gender, placing more emphasis on harmonizing the effect of ethnicity between the enumerator and the respondent might be important in ensuring the authenticity of responses. This can especially be an issue within border villages where respondents often have network members who belong to different ethnicities. There have also been instances where members of some communities have refrained from giving details of network members who belong to a different ethnic groups.

Daily records on network proliferation indicated that some enumerators were capable of performing up to five interviews per day whereas some weren't capable of finishing off five interviews within the first week of data collection. The reason for such variation was due to different social dynamics involving different respondents. In rural more collective and open communities respondent's network members would

be within the vicinity of a few hundred meters. There were instances where the seed respondent elected to take the enumerator to the next referral(s) in their bicycle. Some respondents would provide their network's contact numbers more readily than others. Some would call them up, explain to them the survey and their impressions and would introduce the enumerator voluntarily. However in more closed communities getting in contact with network members wasn't that easy. For instance the enumerator who couldn't finish off more than five interviews during the initial week had a government servant as a seed respondent. They typically have eight to five jobs and by the time their home they either wouldn't like to be disturbed or would be engaged in some other activity making them inaccessible during week days. In such instances there is a higher probability that their network members to have similar professions resulting in such slow network proliferation. In rural farming communities during cultivating seasons finding respondents during the day time even during weekends is difficult. Hence enumerators face immense challenges, working odd hours in rural areas with no transport and security threats not to mention threats from wildlife. This is especially challenging for female enumerators. Finding accommodation in rural closed communities is difficult resulting in higher transportation costs for enumerators. In some geographical locations, each household block looks exactly the same making it difficult for enumerators to locate respondents. In such instances instead of risking a scheduled appointment the enumerators would hire a taxi increasing their transportation costs further.

There were instances where enumerators had experienced male bias among respondent networks. Men would often have expanded networks as they would interact in public spaces such as religious institution and have conversations. However females seem to have had less social ties resulting in slow network proliferation in the case of a female respondent. In such instances the strategy was to find a new seed.

There were instances where enumerators were being questioned by officers of the civil security department who are often in charge of neighbourhood / village security. One enumerator working in a particular village at that time was lodging at a house where the landlord was an influential figure within the village (which he had got to know through personal contacts). When he mentioned that he had come to do a research and was staying at that particular house, the CSD officer had asked no further questions and had let him go. In another instance, an enumerator had been questioned by a police officer who had told him that information cannot be collected randomly without prior authorization. Fortunately the enumerator had informed both the Gramasewa Niladharee and the Divisional Secretariat officers (of the local administration) whom he knew through personal contacts. He had mentioned this to the police officer who had made an immediate check by calling the Gramasewa Niladharee. These examples emphasize the importance of having personal contacts in performing surveys of this nature and also the importance of prior authorization which isn't always practical.

One enumerator in his search for a seed respondent had got acquainted with a local villager and had got contact details of a potential seed. Once the enumerator had visited the potential seed and had revealed that he had got the details of him/her from that particular villager, the seed had got very suspicious and had asked a lot of details regarding the enumerator as well as his relationship to the particular villager. Later he

had revealed that the villager was a drug addict and was a person who regularly went to prison due to unlawful activities. These examples emphasize the importance of being familiar with the local context and the people present within the context in order to prevent threats to personal safety as well as jeopardizing network proliferation.

4. Maintaining privacy and confidentiality within the setting

In villages it is often difficult to have respondents alone. Others such as family members tend to participate within the interview. This can be facilitating in some instances, but inhibiting during some instances. For example there are instances where a wife might supplement a brief response made by the husband (who is the targeted respondent). Sometimes it is the supplementing family member a female in most cases who becomes emotional and expresses opinion. But in some instances other participants may prevent the respondent from answering. For example when asked about examples of mass graves a wife interrupted a husband by saying "don't talk about things you don't know". Also when there are multiple participants with contradictory responses, enumerators often find it quite difficult to carry on with the interview. There have also been instances where respondents have requested enumerators to meet them out of their homes. For example a principle of a school had told an enumerator to come meet him at the school library as his wife wouldn't be too keen in him answering a questionnaire of this sort.

5. Miscellaneous

Due to limited resources during the pilot study, a single enumerator was provided with a particular DS division to work in. As mentioned within the methodology section, in the event of selecting a network member speaking a language that is foreign to the enumerator, he/she was authorized to make another random selection in obtaining a network member speaking the same native language as the enumerator. However due to a miscommunication a Tamil speaking enumerator having full proficiency in Sinhalese had conducted interviews with two Sinhalese network respondents. The respondents had no access to show cards or the questionnaire as Tamil versions of them were used.

Responses in general from Matale in comparison to Ampara which is geographically distant to the ethnic conflict and closer in terms of the ideology to the dominant discourse of the capital seems to indicate that the survey invokes feelings of fear rather than anger or frustration which was initially anticipated during the planning of the fieldwork. Responses in general from Ampara was more open and comprehensive. Responses from Matale were general, ambiguous and safer. For example when asked 'Who do you think should still publicly apologize for past misdeeds in this country?' a typical response from Matale would be "all parties responsible".

Post-fieldwork

According to KCG the error rate in relation to questionnaire completion seemed to be fairly high for the Tamil enumerators. A possible reason for this could be the interviewer trainings being conducted in Sinhalese. In future it seems desirable to have trainings separately for the two languages.

KCG has had difficulty in finding data entry operators who are competent in both MS Excel and Tamil type setting, within the given time constraints. As a consequence they had introduced an additional transcription procedure to utilize the same Sinhala typesetters (data entry operators) for entering data for the Tamil questionnaires.

KCG had experienced significant difficulty with the Respondent ID ('rids') format which had been provided to them by the PMP central team which has a combination of numbers and strings. The inability to check string variables is a limitation in STATA and as a result the inability to check a unique variable (a primary key) becomes a huge set back in checking for errors. This had cost KCG considerable time and effort in checking for errors and rectifying them.