HUMAN FACTORS IN THE INFORMATION DISORDER AND FINDING MEASURES TO COUNTER

Nationally Representative Survey

Survey methodology note

Prepared by LIRNEasia

January 2025

LIRNEasia
Pro-poor. Pro-market.

LIRNEasia is a pro-poor, pro-market think tank whose mission is *catalyzing policy change through research to improve people's lives in the emerging Asia Pacific by facilitating their use of hard and soft infrastructures through the use of knowledge, information and technology*.

Contact: 9A 1/1, Balcombe Place, Colombo 00800, Sri Lanka. +94 11 267 1160.
info@lirneasia.net www.lirneasia.net

Table of Contents

List of tables

1. Introduction

1.1 Background

Media, both traditional (press, TV, radio) and new (digital including social media) play a key role in democratic process by providing a platform for free expression of ideas/opinions and in holding power to account. The problem of misinformation disseminated through media is as old as the media itself. Codes of ethics, press complaints commissions, provisions under various media licensing regimes, limitations over media ownership/concentration are some of the ways in which content, accountability and plurality of traditional print, TV, radio media was enforced. But the expansion of digital technology has enabled the production and broadcasting of information and opinion by individuals, groups and organizations well beyond the traditional media sources. The speed at which information, as well as misinformation, is created has increased drastically.

The spread of disinformation, whether intentional or not, transcends national borders (i.e. global north and south), and languages, and gains cultural significance as it passes through different communities[1]. To fully grasp the impact of problematic information, it is necessary to understand the social and political contexts of different communities.

Understanding the human factors and how they contribute to the information disorder in context is essential to a meaningful response to the phenomenon. This includes the need to systematically understand the individual and group characteristics of those who are more (or less) susceptible to mis/dis/mal information, and what measures (such as the disseminating results of fact checking, media literacy programs, etc.) are most effective (or not) to which audiences are key parts of countering the phenomena.

1.2 Objectives of the study

The objective of the study was to understand the factors impacting an individual's likelihood of being susceptible to manipulated information, i.e.; what are the demographic, socioeconomic, ethnic, contextual and psychological factors that impact people's likelihood to believe or ability to be skeptical about information they encounter via various media sources and formats.

[1]https://misinforeview.hks.harvard.edu/wpcontent/uploads/2022/03/nguyen_diasporic_communities_research_beyond_anglocentrism_20220324.pdf

1.3 Study approach

The study consisted of two phases: a Formative qualitative phase followed by a quantitative phase.

The objective of the formative qualitative phase was to inform the design of the broader study methodology and tools; particularly the development of cue cards, by exploring the media consumption behaviors, attitudes toward misinformation, and verification practices of the target population. Using an inductive approach, this phase aimed to generate contextual insights into how different communities engage with traditional and digital media, their levels of trust in various sources, and their responses to misinformation.

The objective of the quantitative survey was to assess the human factors impacting an individual's likelihood of being susceptible to manipulated information, specifically, the demographic, socioeconomic, ethnic, contextual, and psychological factors that impact people's likelihood to believe or ability to be skeptical about information they encounter, with estimates to be made within a 95% confidence interval with a ±2.7% margin of error.

2. Formative Qualitative Research Phase

2.1 Target population

The target population for this phase consisted of Sinhala speaking and Tamil speaking adults in Sri Lanka, including both social media users and non-users. Participants represented a diverse range of age groups, geographies, and socio-economic backgrounds. The research intentionally included individuals with varying levels of digital literacy to capture how different segments of the population experience and navigate the evolving media environment.

2.2 Sample

The sample for the formative qualitative phase comprised of 60 individuals, who participated in 11 Focus Group Discussions (FGDs) including triads (FGDs with three participants each) and 11 In-depth Interviews (IDIs).

Discussions were conducted from May to October 2024, in Sinhala and Tamil, allowing for linguistic and cultural relevance. The sample reflected diversity in media use, digital access, and demographic characteristics, with particular attention to rural representation and older populations, who were found to be more reliant on traditional media and community networks for information. This sample size and composition were sufficient to surface recurring themes and insights relevant to the study's design.

2.3 Research instruments

All research instruments, including discussion guides, recruitment screeners, and consent forms were developed by LIRNEasia and localized into relevant languages: Sinhala and Tamil. FGD and IDI Instruments included modules to explore media consumption behaviors, attitudes toward misinformation, and verification practices of the target population.

2.4 Recruitment strategy

Participants were recruited using purposive sampling methods to ensure representation across social media usage status, ethnicity, age, gender, and urban-rural location. Local community contacts and field researchers familiar with the respective districts supported recruitment while ensuring geographic and demographic diversity in the

Western Province, Sri Lanka. Screening tools were used to identify participants who met eligibility criteria, including language proficiency and consent to participate. Special attention was given to including rural and estate sector voices, as these populations are often underrepresented in media research yet are crucial for understanding digital exclusion and misinformation vulnerability.

2.5 Data collection

All FGDs and IDIs were moderated in a face-to-face setting in local languages by LIRNEasia moderators. The moderators were trained on research protocols and research ethics. Semi-structured discussion guides served as the primary data collection tools, with a single guide used for both IDIs and FGDs. Interviews lasted between 60 to 90 minutes, and verbal consent was obtained from participants prior to the discussions. All discussions were audio-recorded with consent, and pseudonyms or respondent IDs were used to maintain confidentiality.

2.6 Data processing and analysis

All discussions (FGDs, and IDIs) were transcribed verbatim in the local langue and translated to English as needed. LIRNEasia research team applied inductive thematic analysis using a codebook developed by the team. Preliminary themes were discussed mid-way through data collection to guide refinement of the analytical framework.

2.7 Ethical considerations

All research activities adhered to established ethical standards for qualitative research. Prior to participation, respondents were fully informed about the purpose of the study, the voluntary nature of their involvement, and confidentiality of their responses. Participants were also informed to their right to withdraw at any time without penalty.

No identifying information was linked to the data, and pseudonyms or respondent IDs were used throughout transcription and analysis to protect anonymity. Audio recordings were made only with explicit consent. All data, including transcripts and recordings, were stored securely and accessible only to authorized personnel. Upon conclusion of the study, all personal identifiers will be destroyed in accordance with data protection protocols.

3. Quantitative Research Phase

3.1 Survey approach

Data was collected through face-to-face surveys conducted using Computer Assisted Personal Interviewing (CAPI) and according to the methodology outlined in this document. The fieldwork was conducted by Survey Research Lanka (pvt.) Limited (SRL), a market research company procured through a competitive bidding process.

The questionnaire was developed by LIRNEasia. SRL was responsible for the fieldwork set-up, including questionnaire scripting, translation, and pilot testing, as well as enumerator training and dataset delivery. LIRNEasia, on the other hand, supervised field training in most cases and monitored the fieldwork both remotely and by going on field visits. SRL team visited over 3,000 households across 150 Grama Niladhari Divisions (GNDs)[2] for data collection after undergoing a rigorous one-day training.

Throughout the data collection process, SRL and LIRNEasia adopted several quality control mechanisms that helped to provide timely feedback to data collectors, allowing them to improve their work (refer section 3.xx). The subsequent sections of this document provide more information on the research design, coverage, field problems, remedial actions, quality control mechanisms, and other related topics.

3.2 Survey target group

The target group comprised of Sri Lankan individuals aged 18 years and older with the ability to read and visually process information.

3.3 Sampling

3.3.1 Sample size determination

The desired level of accuracy was set at a confidence level of 95% and an absolute precision (relative margin of error) of 5%. The population proportion (p) was set conservatively to 0.5, which yields the largest sample size. The minimum sample size (n) was determined by the following equation.

[2] Grama Niladhari Division is the smallest administrative unit in Sri Lanka. These divisions are subdivisions of larger Divisional Secretary's Divisions and are headed by a Grama Niladhari (village officer) responsible for administrative and community-level tasks.

$$n = \left(\frac{Z_a \sqrt{p(1-p)}}{C_p} \right)^2$$

Were,
n = Minimum sample size
Za= Z-value for 0.05 level of significance
Cp = Confidence level
p = Population proportion

Inserting the parameters yields the minimum sample size for simple random sampling; therefore, for our sample design (stratified with multiple levels in some cases) the minimum sample size was multiplied by the design effect variable. In the absence of empirical data from previous surveys that would have suggested a different value, a value of 2 was used as the design effect to consider the additional levels of selection (province groups and districts) in the sample design. This yields then, a minimum sample size of 768 individuals.

The actual sample size was increased beyond the minimum requirement to compensate for clustering effects, and to allow for urban, rural and estate disaggregation of data, as well as gender-based disaggregation.

3.3.2 Sample frame and PSU definition

The targeted sample size was 3,000 individuals. The primary sampling unit (PSU) was the Grama Niladhari division (GND). The method was developed using population and housing data at GN level from the Department of Census and Statistics, Sri Lanka. Random sampling was performed in three steps resulting in the selection of 150 PSUs across Sri Lanka.

3.3.3 Overview of the sampling procedure

The study methodology ensured national representation through random selection of all sampling units (GND, household, individuals, etc.), enabled by household listing at the PSU level.

The methodology involves the following steps:

1. The GND level sample frame was separated into urban, rural, and estate PSUs.

2. The required number of GNDs were sampled from each stratum[3] using probability proportionate to size (PPS).
3. Listing maximum of 260 households in the selected PSU. The lists served as the sample frame for random selections of households.
4. Systematic random sampling selection of the required number of households (20 households) from each selected PSU.
5. Listing of all household members who are aged 18+ years, and selecting a member using random sampling. The name, year of birth, age, gender, and relationship to household head was recorded for each member who was 18+ years old.

3.3.4 PSU selection

The selection of PSUs was conducted by LIRNEasia and the list of selected PSUs was provided to SRL for fieldwork execution.

Based on the population data available at the National Census Data Tables[4] compiled by the Department of Census and Statistics of Sri Lanka, the population was split into 5 province groups Western, Northern, Eastern, Southern and Rest of Sri Lanka and within each province group, again split into urban, rural and estate.

A pre-defined number of PSUs (GNDs) were selected using PPS sampling techniques. The proportional allocation at district into urbanity level (urban / rural / estate) was used to maintain the district-wise coverage of the sample.

Six initially selected Grama Niladhari Divisions (GNDs) were replaced due to various challenges, including opposition from religious leaders, community reluctance to participate in surveys close to the election, and lack of approval from Divisional Secretaries who had been instructed to restrict survey activities during the pre-election period. Additionally, one GND was partially replaced after the Grama Niladhari officer raised objections midway through data collection[5].

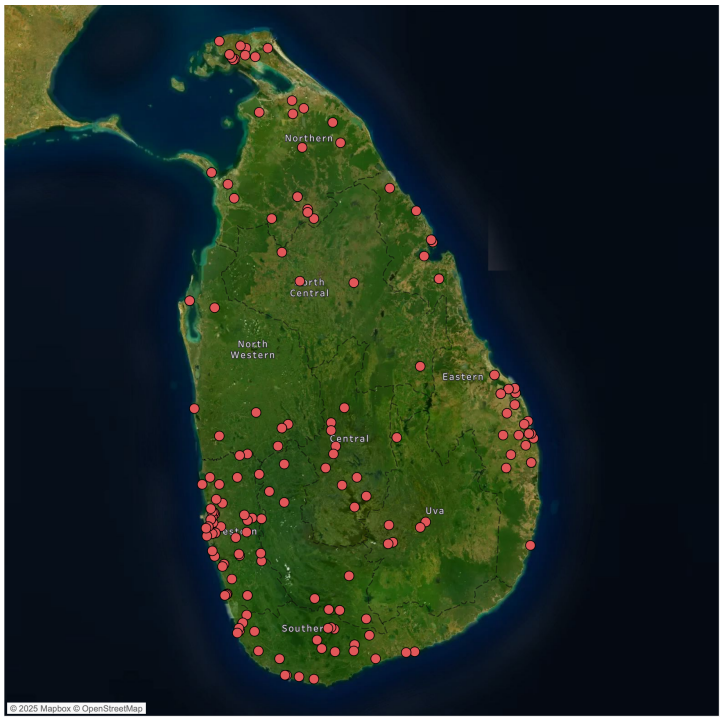
[3] There were 13 stratums used in this study. The 5 province groups considered in the study was divided into urban, rural and estate PSUs depending on the availability of those relevant type of GNDs.
[4] Census of Population and Housing 2012
[5] The National Survey was conducted from August to September 2024, prior to Sri Lanka's presidential election on September 21, to comply with restrictions on research and public activities during election periods. Given the study's sensitive content—political affiliation and ethno-political misinformation —LIRNEasia and its partner, SRL, obtained approvals from local authorities. Despite this, enumerators faced access issues in areas like Batticaloa, Ampara, and Gampaha, with some being briefly detained by police. Fieldwork was halted in affected areas, permissions were renegotiated, and alternate locations were used where necessary. Though delayed, the survey was completed ahead of the election as planned.

Selected PSUs are depicted in Figure 1. Please refer Annex 1 for sample sizes and margin of errors.

Table 1:Number of PSUs by each stratum

Province group	Number of PSUs			
	Urban	Rural	Estate	Total
Western	17	19	2	38
Southern	5	18	2	25
Northern	6	19	-	25
Eastern	10	15	-	25
Rest of Sri Lanka	9	20	8	37
Sri Lanka	**47**	**91**	**12**	**150**

Figure 1: The 150 PSU locations used in the survey



3.3.5 Mapping, listing, and conducting main surveys within sample PSUs

Since there was no readily available GND-level household and population data in the form of a list, mapping and listing of all households within the PSU was necessary to draw random samples at the household level. The objective of mapping and listing was to ensure the coverage of the entire population residing in the PSU. This facilitated the equal probability random selection of households for the survey, enabling all eligible households in the PSU to have an equal chance of being selected.

The listing process was essential for creating a comprehensive list of occupied residential households in the selected GND and provided a complete range of households for the survey. By employing systematic and equal probability random selection of households, the listing process helped overcome the problem of biased household selection for the survey.

The listing of households and conducting the main surveys among households selected using systematic random sampling took place concurrently. To facilitate this approach, the survey script included both the listing and the main survey components.

3.3.6 Mapping and identifying the GND boundaries

Administrative boundaries of each selected GND were identified and validated, and GN boundary maps were created in the form of Google Maps and provided to each enumerator. The map was able to show the real-time location of the enumerator in relation to the GND boundary to ensure that the survey is done within the GND.

3.3.7 Starting points (SP)

A randomly generated starting point (geo-coordinates) was assigned to each selected PSU. The field team used the starting point to start the listing process. When GPS points were located in large uninhabited areas such as lakes, forests, and lagoons, or if there were no households within a 100m radius of the starting point, another random starting point was given to the field team to start the listing.

3.3.8 Determining the interval "k" for systematic random sampling

A total of 20 interviews (pre-determined) were conducted from each selected PSU. Hence, it was decided that in PSUs with more than 260 households, a maximum of 260 households will be listed whereas in PSUs with less than 260 households, all households will be listed.

The number of households in each GND was verified with the Grama Niladhari Officer, the respective Divisional Secretariat, or an equivalent authority. Once confirmed, systematic random sampling was implemented during household listing. To determine the sampling interval (k), the total number of households was divided by 20; the target sample size surveys per PSU. In PSUs with more than 260 households, a fixed interval of 13 was used (260/20). In PSUs with fewer than 260 households, the interval (k) was calculated by dividing the total household count by 20; which was the expected sample size per PSU.

3.3.9 Listing of households and selection of households for the survey

The listing exercise was conducted using CAPI devices. The SRL team listed all the structures (or a maximum of 260 households) within the PSU to ensure no households were missed out. Starting from the closest household to the SP, right hand-rule was used to list the households in a PSU.

The following details regarding the household was gathered in the listing process:
a) Whether the structure is a residential or a non-residential structure
b) Address of the structure
c) GPS coordinates of the structure
If a residential structure:
d) Contact information (mobile or landline number of the household member spoken to, if available)

First household to be enumerated was selected based on a random number (r) between 1 and k of a particular PSU, generated by the CAPI device. While conducting the listing process, starting from the first randomly selected household, in every k^{th} household, a main survey was conducted. Accordingly, listing was carried out for the effective number of households of the GND while concurrently administering the main survey for every k^{th} household until the expected number of listings and main surveys for the PSU were completed.

3.3.10 Selection of household member for the main survey

In a selected household, the eligible members (members aged 18 years and above) were listed, and one member was selected for the survey using simple random sampling. The selected household member's ability to see and read was assessed by presenting a short, written passage and asking them to read it aloud. Individuals who were able to do so were included in the main survey. Those who could not were screened out, and the household was replaced The screened out interviews were retained and was used in weighting the data.

In the case that the randomly selected respondent from the household was not available during the first visit, a minimum of three (3) attempts (first visit + two follow-ups) were made (at different times of day or based on an appointment) before considering replacement. If the respondent could still not be reached, the entire household was replaced by moving onto the next household. Individuals were not replaced by another individual within the same household.

3.3.11 Replacement procedures

Strict replacement procedures were followed to ensure the randomness of the sample at all levels was preserved. The circumstances under which replacements were allowed are detailed below.

- The selected individual refused to participate in the survey
- The selected individual was unavailable after three attempts to contact him/her
- The selected individual was unable to be survey due to severe disability
- The selected individual was not able to see and/ or read his/her preferred language of communication. (Sinhala, English or Tamil)

If any of these situations arose, the household was replaced with the next household (without affecting the selection of the k^{th} household in the systematic random sampling process).

Replacement of individuals within the same household was not permitted.

3.4 Research Instrument Development, translation, and scripting

A structured questionnaire in English language was developed by LIRNEasia to capture data. The questionnaire was translated into Sinhala and Tamil by SRL. Translations were extensively checked by the LIRNEasia research team.

The questionnaire included the following modules:

Module 1:	Introduction of survey and consent of respondent
Module 2:	Household member details, random selection of household member to be interviewed for the main survey (including screening for ability to see and read) and consent from selected respondent
Module 3:	Mobile and Social media usage
Module 4:	Classification of cue cards (refer section 3.4.1 for more details of cue cards)
Module 5:	Sources of information, their awareness on misinformation and their behaviour in relation to receiving of misinformation
Module 6:	Psychological factors that could impact the identification of misinformaiton
Module 7:	Political affiliation
Module 8:	Demographics

The questionnaire was pilot tested extensively by the LIRNEasia research team in both Sinhala and Tamil languages prior to finalising the content. The finalised questionnaire was then shared with SRL team for scripting.

The scripting of the questionnaire was done by SRL. The SRL project teams tested the script for all logical and consistency checks before planning for pilot-test. The bilingual tools and script with login details were also shared with LIRNEasia for comments and feedback before pilot-test. Modification of the script was done based on comments from LIRNEasia as well as SRL internal testing.

The modified script was then pilot tested among xx respondents. (please refer section xx for more details on the pilot testing)

3.4.1 Cue cards

As stated above, the questionnaire included a module containing 40 cue cards which contained information across three thematic areas: climate change, political and economy, and political–ethno-religious issues, and included a mix of true, mostly true, mostly false, and false news items. Three such sets of cue cards were developed, one in each language (Sinhala, English, and Tamil) by Watchdog/Appendix under the supervision of LIRNEasia.

The cards were pilot tested in a few qualitative groups by the LIRNEasia research team and was modified based on the learnings. These cards were then submitted to the Ethics Review Committee for Social Sciences and Humanities (ERCSSH), Faculty of Arts, University of Colombo, and clearance was obtained prior to their use in the survey.

3.4.2 Scoring of cue cards

In order to assess the ability of respondents to correctly identify information as true, mostly true, mostly false, and false, the following scoring method was utilised. Each cue card was given a maximum score of 3 and a minimum score of 0.

Table 2: Scoring matrix for evaluating respondent classifications against the actual truthfulness of cue cards

		Correct classification of the cue card			
		True	Mostly true	Mostly false	False
Classification by respondent	True	3	2	1	0
	Mostly true	2	3	2	1
	Mostly false	1	2	3	2
	False	0	1	1	3

3.4.3 Pilot testing

Pilot tests were done by LIRNEasia prior to scripting the questionnaire. These were conducted in non-sample locations, and among a range of respondents to ensure representation across different languages, urban and rural settings, socio-economic groups, and genders. The pilot tests aimed to estimate the interview length, assess the clarity of the questionnaire, and test CRT-2 questions which were adopted to suit the Sri Lankan context.

Post scripting the survey questionnaire, 4 pilot interviews were conducted by SRL in Colombo, Gampaha, Kegalle and Kalutara.
The goal of the pilot survey included estimating a realistic median length of interview (LOI) and testing the following:

- Skipping and routing
- Language (phrasing)
- Translation
- Understanding of the questions, cognitive difficulties and question sensitivities
- Tablet Functionality

Data collection needed to be accelerated due to the need of completing the fieldwork for the survey prior to the Presidential Election scheduled for 21 September 2024. Therefore, given the 4 weeks time frame available starting from 17th August, only a material pre-test was conducted post-scripting the survey.

3.5 Enumerator training

Enumerator training was conducted prior to rolling out the fieldwork. Three enumerator training sessions were conducted, one in Sinhala language in Colombo (with the participation of all enumerators working in all areas except Northern and Eastern provinces) and two training sessions in Tamil in Jaffna and Ampara.

3.6 Ethical considerations

Since the respondents of the study were exposed to cue cards containing information, of which, some were not entirely true, there was a risk that the study could introduce misinformation, potentially causing individual, communal, or societal harm if some respondents believed the fake stories and spread them further. To mitigate this risk, respondents were provided with a debriefing once they completed the survey. The

purpose of this debriefing was to prevent the perpetuation or reinforcement of any false information that may have been presented during the survey.

3.7 Consent

Prior to the commencement of an interview, the respondent was informed of the following:

- The objectives of the research
- That his/her participation was voluntary
- That he/she could choose to end the interview at any point

Once informed of these, the respondent's written or verbal consent was obtained from the respondent to (1) proceed with the interview; (2) record the interview; (3) be photographed [if applicable]; and (4) be contacted again for further research. Respondents who refused to be recorded and photographed could still be interviewed if they agreed to participate.

3.8 Photography

Two-three teams captured photographs during the listing / mapping exercise (structure number on walls, listing household, KI interview or permission) using CAPI devices. Proper consent of the was taken before taking photograph.

4. Quality Control (QC) Mechanisms

To ensure the integrity, reliability, and ethical standards of the study, a comprehensive set of quality control measures were implemented across all phases of the research by both.

4.1 Quality control mechanisms used during the Formative Qualitative research

a. Research instrument design: All instruments (discussion guides, recruitment screeners, consent forms) were developed and discussed collectively by the LIRNEasia research team.
b. Recruitment and screening: Respondents were recruited according to the strict eligibility criteria, including demographic characteristics. Quotas were maintained to ensure adequate representation across districts.
c. Moderator and field staff training : All moderators and field staff were trained in discussion facilitation, ethical conduct and privacy protocols. Moderators were required to demonstrate proficiency before deployment into the field.
d. Informed consent and ethical compliance: Informed consent was obtained verbally prior to participation. Participants were informed of their rights, study objectives, confidentiality terms, and the option to withdraw at any time. Recordings were taken only with participants consent.
e. Supervision: Field teams were supervised by senior field staff, who ensured protocol adherence and addressed any issues as they arose. LIRNEasia received weekly progress updates from SRL.
f. Transcription and data management: Transcripts were reviewed for completeness and accuracy. Files containing respondent information were stored and accessible only to authorized personnel.
g. Analytical rigor: Triangulation was employed to validate findings across data sources (FGDs and IDIs).

4.2 Quality control mechanisms used during the Quantitative survey

1. Map checks (MP) - 100% GPS location checks were carried out for all listings to ensure the use of correct stating points and that the enumerators have followed the right-hand rule accurately.

2. Accompaniments checks (AC) - 178 partial and 74 full accompaniments were conducted across 126 enumerators. Accompaniments were carried out by SRL field supervisors, field executives, researchers as well as LIRNEasia research team.
3. Telephone back checks (TBC) – 1450 TBCs were done by the supervisors
4. Physical back checks (PBC) – 880 PBC were conducted randomly in 17 GN divisions to confirm interview validity based on TBC feedback.
5. Voice checks (VC) – 2,846 Different sections of audio files were examined.
6. Timestamps – Automated checks were conducted to ensure that interview was conducted during work hours.
7. Length of interview (LOI) checks – LOI checks were carried out to assess the interview length of each enumerator. Further, LOI checks were also carried out for some sections of the questionnaire to assess that adequate time has been spent on the particular sections.
8. CAPI script ensured the right filters for the questionnaire.
9. Enumerators were trained on general Dos and Don'ts in the field, on concurrent conduct of listing and main interviews, on the questionnaire content, and to administer different questions and scales. Further, mock interviews were conducted by enumerators priior to be sent on field for data collection.
10. Debriefing sessions were done for the enumerators who made mistakes during the interview process.

All research process and practices followed by SRL are in accordance with the ESOMAR research guidelines.

Annex 1 - Sample sizes and margin of errors

Table 3: Sample size, number of PSUs and margin of error at province level after data cleaning process

Province	Sample size				Margin of Error			
	Total	Urban	Rural	Estate	Total	Urban	Rural	Estate
Western	713	309	365	39	5.19%	7.88%	7.25%	22.19%
Southern	470	89	344	37	6.4%	14.7%	7.5%	22.8%
Northern	377	86	291	0	7.1%	14.9%	8.1%	-
Eastern	375	149	226	0	7.2%	11.4%	9.2%	-
Rest of Sri Lanka	685	175	375	135	5.3%	10.5%	7.2%	11.9%
Total	2620	808	1601	211	2.7%	4.9%	3.5%	9.5%