

**National Survey Design Planning Report**  
**Skills Toward Employment and Productivity (STEP)**  
*Full Assessment*

**COLOMBIA**

**February 2012**



## INSTRUCTIONS TO COMPLETE THE NSDPR

Grey Boxes	<p>The grey boxes will be used for</p> <p>(A) Background information for the section            (B) Outline of the required information            (C) Agreed standards as required by the Terms Of Reference (TOR)</p> <p>NO changes required.            Occasionally, the survey firm will be required to fill basic information (such as the name of the country) within these boxes. These situations will be signaled by being written in blue between the ‘&lt;’ and ‘&gt;’ characters (e.g. &lt;Country&gt;)</p>
Green Boxes	<p>Information that needs to be filled out by the Task Team Leader for the World Bank country team.</p> <p>The empty spaces mean that details should be provided about certain sections. The responsible person should take as much space as needed in order to provide as complete information as possible.</p>
Yellow Boxes	<p>Information that needs to be filled out by the Project Manager on behalf of the survey firm / agency.</p> <p>The empty spaces mean that details should be provided about certain sections. The firm should take as much space as needed in order to provide as complete information as possible.</p> <p>If a yellow box contains both text and empty spaces, please refer to the document ‘<i>National Survey Design Planning Report [with examples]</i>’ in order to understand what information the countries are required to add.</p> <p>The person responsible for completing the NSDPR is responsible also for reading the entire document. If it seems like information is missing from a grey box, create a yellow box below the grey box, and add the missing information in the yellow box.</p> <p><b>FIRMS/AGENCIES SHOULD ADJUST THESE SECTIONS BASED ON THE COUNTRY’S EXPERIENCE.</b></p>
Orange Boxes	<p><b>TASK TEAM LEADERS and PROJECT MANAGERS</b> – Please fill in and read carefully the whole document. Being based on the TOR and technical proposals of each country, it is a binding document on implementation procedures. After ensuring that each box accurately reflects the implementation procedures, both Task Team Leaders and Project Managers are required to sign in the appropriate boxes under each section.</p>



## Table of Contents

<b>1.0</b>	<b>Introduction</b> .....	<b>1</b>
<b>2.0</b>	<b>Project Team</b> .....	<b>2</b>
2.1.1	Team and key staff.....	4
2.2	Project Structure.....	7
<b>3.0</b>	<b>Survey Objectives</b> .....	<b>11</b>
3.1	Major analytic objectives.....	11
3.2	Country-specific objectives .....	12
<b>4.0</b>	<b>Sample Design Factors</b> .....	<b>13</b>
4.1	Target Population.....	13
4.2	Method of Data Collection.....	17
4.3	Response Rate.....	20
4.4	Sample Frame .....	21
4.5	Sample size – Full Assessment.....	23
<b>5.0</b>	<b>Sample design</b> .....	<b>26</b>
5.1	SAMPLING METHODS AND SELECTION ALGORITHMS .....	27
5.2	Sample Allocation and Selection .....	29
5.2.1	Random Selection of Assessment Exercise Booklet .....	29
<b>6.0</b>	<b>Literacy Assessment</b> .....	<b>31</b>
<b>7.0</b>	<b>Household Questionnaire</b> .....	<b>36</b>
<b>8.0</b>	<b>Data Collection</b> .....	<b>38</b>
8.1	ASPECTS OF LOGISTICS.....	43
8.2	OPERATIVE SCHEMA .....	43
8.3	SUPERVISION OF THE FIELD WORK.....	45
8.4	Progress Reporting .....	47
<b>9.0</b>	<b>Data Processing</b> .....	<b>48</b>
9.1	Instrument Requirements to Facilitate Data Processing .....	48
9.2	Data Processing: Data Capture, Coding, Scoring, File Creation .....	49
9.2.1	Recruiting and Training Scorers .....	<b>Error! Bookmark not defined.</b>
9.2.2	Ensuring Inter-rater Agreement .....	<b>Error! Bookmark not defined.</b>
9.2.3	Documenting the Scoring Process .....	<b>Error! Bookmark not defined.</b>
9.2.4	Creation of International Data File .....	<b>Error! Bookmark not defined.</b>
9.2.5	Data Editing System .....	54
<b>10.0</b>	<b>Weighting</b> .....	<b>57</b>
10.1	Weighting Procedures.....	58
10.1.1	Benchmarking Variables.....	58
10.1.2	Source of Benchmark Variables .....	58
<b>11.0</b>	<b>Confidentiality</b> .....	<b>59</b>
<b>12.0</b>	<b>Quality Assurance</b> .....	<b>61</b>
<b>13.0</b>	<b>Schedule</b> .....	<b>63</b>
<b>14.0</b>	<b>Budget</b> .....	<b>65</b>

## 1.0 INTRODUCTION

### Rationale

The introduction should inform the STEP Consortium of the history of events leading to participation in the STEP. The intention is to provide an explanation of the nature of the literacy situation in the country. An overview of the literacy concerns and STEP expectations will provide insight into the needs of the country in this regard.

### Report Requirement

1. Provide a brief background of the country's involvement in the STEP survey.
  - a) Include, for example, a brief overview of the literacy situation in the country, reasons for participation, expected benefits for the country, etc...
  - b) Provide information on the country's involvement, if any, in previous adult literacy assessment surveys.

Skills are at the core of improving employment outcomes and increasing productivity and growth. Across countries, unemployment and low productivity employment can often be the result of workers not having the right skills to match the requirements in available job openings or having limited opportunities to access high quality pre-employment or skills upgrading training programs. In many countries education and training systems often lack quality and labor market relevance, leaving workers ill-prepared for the labor market.

Against this background, the World Bank launched a multi-country research program that finances country-level studies to determine how different skill sets affect individuals labor market opportunities. The studies are expected to fill critical knowledge gaps on the role and demand for different types of skills sets in the labor market and assist in the design of tailored education and training policies to boost employability and productivity.

## 2.0 PROJECT TEAM

### STEP Standard

*Each participating country will describe the qualifications and experience of the leading survey institute/project team responsible for the design and implementation of the STEP survey.*

### Rationale

The quality of the STEP survey depends on the practical abilities and experience of the survey institute responsible for the design and implementation of the survey. In order to ensure the survey quality, the leading institute must have expertise and qualifications specific to the design and implementation of large-scale international surveys. Also, in many participating countries there may not be a single institute that has all the qualifications needed to undertake the STEP survey on its own. Consequently, there may be a need for collaboration between different institutes and, at least the leading survey institute must have qualifications in collaborating with other national and international institutes so that expertise in the relevant STEP areas is available.

In general, each STEP national team should be made up of experienced, knowledgeable personnel with expertise in one of the following survey areas: survey management, probability sample design, data collection including interviewer training and non-response reduction, data processing including data capture, coding, and editing, survey weighting and estimation, or data analysis. Furthermore, expertise in coding levels of education and industry and occupation data to international standards is required. In addition, during the development phase of the project a language specialist is needed to provide expertise in the translation and adaptation of the survey instruments.

## 2.1 **Qualifications and Expertise of the Survey Institute.**

**Econometría and SEI** have an alliance of almost 20 years working in Temporary Unions like in this case to complement the methodological design of the projects, the sample design and the recollection and analysis of the information in a perfect way. They developed a mutual process that guarantees the high efficiency and the quality in all of the executed projects. A recent example of their work is the evaluation of the effects caused by the financial education program for the beneficiary families of Familias en Acción, and also the programs of reintegration, habitability and the National Plan of Territorial Consolidation.

Each consultant firm has the following national and international experience:

**Econometría S.A.** is a Colombian consultancy firm, which is a leader in the region producing different kind of analysis about economic and social issues. Since its foundation in 1977, the firm has been working to develop around of 505 projects. Econometría S.A. has a strong corporate management and it counts on a wide experience elaborating this kind of projects. In addition to the researches the company has performed about most of the Social Programs from

the Colombian Government, Econometría executed numerous studies about subjects as nutrition and health for the World Food Programme and the ICBF (Colombian Institute of Welfare Service) and has evaluated programs as RESA. The firm has carried out studies about education with the Colombian Ministry of Education and it has also evaluated programs aimed to increase the cognitive capacities of foundations as Terpel, Cerrejón, Luker, Cerromatoso y Promigas; additionally Econometría supports the displaced persons<sup>1</sup> with programs to improve their quality of life in areas as the Territorial Development and the Citizen Coexistence through the evaluation of Plan Pacífico, PNCT, MIDAS, PADEMER and the execution of plans designed to help the people returning to their homeland.

**Sistemas Especializados de Información, SEI s.a.**, is the most experienced Colombian Entity in design, execution and systematization of urban and rural field works aimed to obtain socioeconomic information from homes, individuals, economical establishments, agricultural production units or entities of population support. It is important to emphasize their wide experience applying educational competence tests, cognitive development tests and language, socio-economical and motor functions tests, as well as their execution of projects with the labor force as main part of the study. SEI s.a. developed 400 different projects since 1992, including census, survey sampling, and compilation and processing of data, information systems and designing of statistic investigations and dozens of projects to evaluate of the effect produced by the implementation of the government programs. Some years ago SEI s.a. was the pioneer in Colombia introducing the technology of Dispositivos Móviles de Captura<sup>2</sup> (DMC), which is digital equipment used to collect the information of high complexity formularies. During the development of the projects, the entity has worked all around the country. SEI s.a. has also a wide experience in collection, refinement and constitution of survey's data bases about the quality of health service or specialized aspects which requires clinical examination, anthropometrical measurements, the taking of blood samples to detect the anemia, surveys about the way the women are treated in the hospitals, oral health, dengue and malaria control, children's cardiopathies, youth drug addiction and HIV, an example of it was the National Health Survey. The characteristic quality of the entity service is certificated by the ICONTEC due to the development of the processes following the norm ISO 9001 (No. SC 3764-1).

---

<sup>1</sup> Refugees forced the live in cities, because of the countryside war.

<sup>2</sup> Mobile dispositives of data collection.

### 2.1.1 Team and key staff

Below we present the team and key staff, their description, their studies and their experience for the principal team.

- 1) The **Survey Project manager** is **Ms. Carolina Murgueitio**. Ms. Murgueitio is an economist graduated from the Javeriana University (1991). She obtained a Master's degree in Political Science from Los Andes University and an MBA with emphasis on Public Management. Her main experience is in impact evaluation and household surveys. She has also participated in several projects related to social and economic evaluation of educational projects, project structuring and potential markets. Expert in design and application of tools for the evaluation of public policies. She was director of the impact evaluation of the educational programs of Fundación Promigas; Consultant in a research project on citizens perception on justice, truth and reparation in rural and urban zones of the country, as well as in the consultancy to measure the perception of justice, truth, reparation and reconciliation in the Department of Nariño; Project Manager of the impact evaluations of the subprograms "Families in Action" Familias en Acción carried out by the Investment Fund for Peace, from the Presidency of the Republic.
- 2) **Field work expert Ms. Angela Yamiled Palacio Basto** is Assistant manager at SEI and psychologist with master degree in population studies, certification in quality management systems- ISO 9000 and technical formation in pre-primary teaching. Her main strength is the deep knowledge in methodological design; she has developed the methodology and implementation of field works in census and surveys, and she also has directed the data entry and the processing information. She has wide experience in management of national census and surveys. Her main experience has been in experimental census in San Andres and Villavicencio (Census 1993 – DANE, national department of statistics), national survey about coffee, pre-census in Bogotá and Soacha (Census 1993 – DANE), national inventory of potable water and basic sanitation in all country. She was the director in different projects like the inventory of resources and services of health care attention – first level, assignments for the ICBF – Colombian institute of family welfare, attention and prevention model for children and families who live in the street in 16 cities of Colombia, census of the program "Madres comunitarias", evaluation to the program of Breakfast in kids, perception survey about corporation image and services for DIAN – national direction of taxes and customs. She participated in the field work of SISBEN (this project collected 600.000 interviews), in the CASEN survey and in the evaluation of governmental programs like "Familias en Acción", "Jóvenes en Acción" and "Empleo en Acción". She has been counselor in the technical design of surveys in Central America for OPS (Pan American Health Organization). She always has been interested in learning different issues in order to give a technical guide in every project that she directs. Like operational chief at SEI, she is responsible for organizing and coordinating every project and she also accompanies them.

- 3) The leader of field teams will be **Mr. Antonio Rey**. Mr. Rey is Cadastral engineer and geodetic with specialization in demographic analysis methods. Now he is doing the master program in population Studies. Since 1993, he has worked with SEI in different projects and with different positions, now he is Data Management Director in many projects. He was field work coordinator in different zones of Colombia in the National survey about coffee, and he also has coordinated and directed the field work for the impact evaluation of governmental programs like “Empleo en acción” and “Familias en acción”. He also has directed evaluation projects about agriculture like “Alianzas Productivas”, “Pademer”, National census of palm oil, “Agro Ingreso Seguro”, USAID interventions. He also has participated in educational projects like “Computadores Para Educar”, National survey about consumption of psychoactive substances in young people between 12 and 17 years old and in the national project about transportation system in the schools. In the financial field, he has worked in the impact evaluation of bank opportunities. In the official and private sector he has worked like coordinator, supervisor and engineer in the cadastral actualization.
- 4) The **data management specialist with experience in capture and information processing** will be **Ms. Sonia Roa**. She has a specialization degree in database development; she is business administrator with emphasis in finance and insurance and she has a technical degree in Analysis and Design of Computer Systems. She has twenty years of experience in evaluation, documentation, programming, and treatment of statistical information and commercial information of large databases. She belongs to SEI s.a. staff and she has been Processing Coordinator in databases management in huge projects like assessment of quality education, national health survey and evaluation projects of governmental programs: “Familias en acción” and “Empleo en acción”. She was administrator of SISBEN survey database, this survey was made by SEI s.a. in 300.000 households. She has processed the first Educational census in Pereira. She has been the administrator of different databases in minor complex projects but oriented to special populations. When she was chief in the Information Services Department, she participated in projects like the design and implementation of data banks and information service of SEI s.a., she also worked in different studies like localization of bank offices and spatial analysis of points of sale. She has knowledge in ISO system (International Organization for Standardization), especially in quality management systems ISO 9001-2008 and she is chief auditor in quality management system at SEI s.a. She has participated in processing projects with ECONOMETRIA S. A., projects like Surveillance and Control Inspection of SSSS, FEDEMOL and incomes and expenses surveys. She was information analyst in the System Information Department of FICITEC – Foundation for the Promotion and Development of Scientific and Technological Research and she participated in the design and implementation of urban information system with georeferenced information.
- 5) The **Chief Scorer** is **Ms. Olga Romero**. Ms. Romero is bilingual, Economist with a master in Economics from the Universidad de los Andes (2006) and a Master in Public Administration and International Development from the Harvard Kennedy School (2010). Experience working in development and social programs for Colombia, Latin

America and India. She worked in Econometria from 2006-2008 and recently rejoined the firm after completing her studies and work experience abroad. As a consultant in Econometria she has searched relevant bibliography, designed instruments (e.g. surveys), carry out field work, conduct statistical and econometric analysis of the data, and elaborate reports. She also worked at the WORLD BANK in Washington, D.C .Unit of Poverty and Gender for Latin America and the Caribbean Region. Her tasks were: Support analytical and advisory assistance for countries in the region. Interact with countries, conduct econometric analysis and elaborate reports.

- 6) Additionally to the team and key staff, SEI s.a. will assign to the project one survey expert. **Hanner Sanchez** will be our counselor in survey selection. . Mr. Sanchez is expert in statistics and Survey methodology at SEI s.a. He was research assistant in math area in the project on “Leveling for Excellence” developed by National University of Colombia trough “RED” program for District Education Secretary. Mr. Sanchez worked like consultant for Sample design in the National Survey about commerce, Industry and Service at DANE (National Department of Statistics), he also was consultant for Colombian institute of family welfare (ICBF) and Profamilia in the statistical analysis of different components in the National Survey of Health and Demography 2005, in the National survey about Nutritional state in Colombia 2005 and Mr. Sanchez also was counselor for the National Institute for Health in Colombia for many years. Mr. Sanchez has participated elaborating final reports and analysis for citizens and school population survey (9th, 10th and 11th grades) in the project “Baseline measurement of citizen culture” for the Observatory of Citizen culture. Mr. Sanchez selected the sample for VI and VII stages in Social situation survey made for FEDESARROLLO, he also proposed a methodology for the variables measurement of this study in displaced population concentrated in some zones in Bogota. Mr. Sanchez is co-author of the methodology about anthropometric measures for the impact evaluation of subprogram “Familias en acción”. He supported the sample design for the Impact Evaluation for the subprogram “Empleo en acción”. Mr. Sanchez was counselor in Case Studies in order to evaluate the food necessities in displaced population in 41 municipalities in Colombia for the World Food Program (WFP). Mr. Sanchez made the sample design for impact evaluation of coal mining in Cesar and he also has participated in the last nine years in more than 50 projects making many National sample designs.
- 7) Additionally to the minimum key staff, SEI s.a. assign to the project Angélica Becerra, statistician who works at SEI s.a. (Sistemas especializados de información s.a.). Angelica has experience in survey methodology and national surveys; she belongs to the European Survey Research Association (ESRA). Angelica was research assistant in the first and in the second national survey to displaced population; she supported the coordination of the field work and made the reports to the sample in field work, she also collaborated in the definition of the survey methodology. She made the survey design in the national project about the incomes to health system, trough lotteries tax, and in this research she also interacted with the field work team during the field work. On the other hand, Angelica participated in the baseline indicators to life quality in

senior population and in the production of judicial statistics.

Functions of the Language Specialist will be covered by Carolina Murgueitio, Olga Romero and Angelica Becerra, people from the firm who are bilingual. They are going to be in charge of the process of adaptation.

**I agree with the above,**

PABLO A. ACOSTA

**Signature:**

CAROLINA MURGUEITIO

**Signature:**

## 2.2 Project Structure

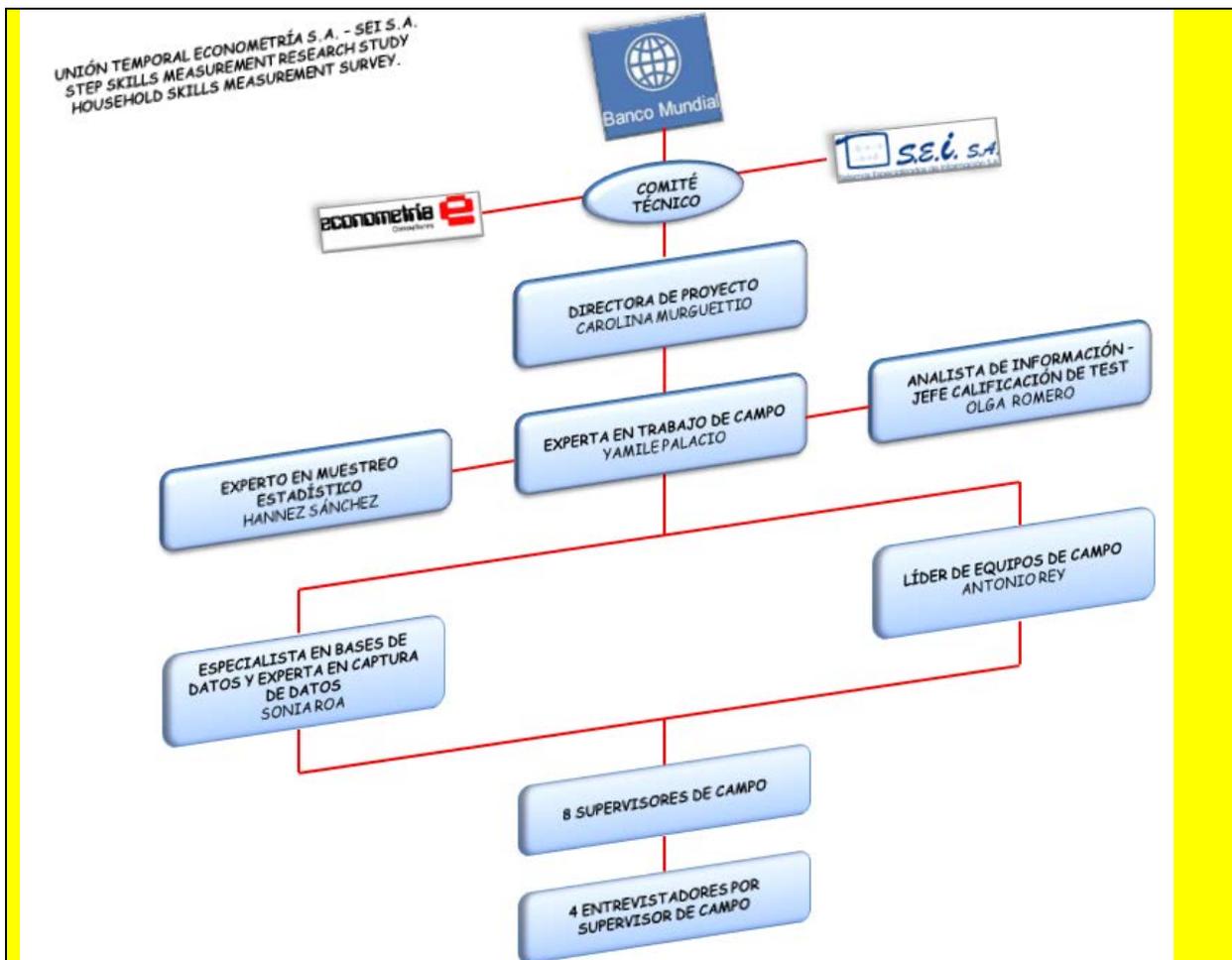
### Report Requirement

1. Identify the sponsoring organization

In accordance with the technical focus and the operative structure explained in the proposal, SEI s. a. offers to develop the whole project and the activities described, using organization and team work. The project shall be realized on the model of a functional organization schematic which is shown in the organization chart on the following page.

We propose a **technical committee**, lead by the World Bank and with the participation of the UT, which would serve a channel of coordination and complete functions such as lend technical support, help facilitate the creation of contact with authorities and different persons interested in the project. Also to facilitate the transfer of informative communications, letters of presentation explaining the realization of the project, approve the final content of the forms, and continue the development of the project through periodic meetings.

The management of **contractual issues** will be the responsibility of the UT representative. Using the institutional support from the physical infrastructure, they shall ensure; communications and data processing, assigning personnel to execute the necessary administrative and logistic processes, the contracting of personnel and of transport services, the administration of the adequate remuneration system, the acquisition of supplies, which the goal to ensure that the project is realized at the highest level of quality.



- The **Survey Project manager** is Ms. **Carolina Murgueitio**. Ms. Murgueitio is an economist graduated from the Javeriana University (1991). She obtained a Masters degree in Political Science from Los Andes University and an MBA with emphasis on Public Management. Ms. Murgueitio is going to be in charge of technical direction and she will guarantee the application of the instructions from the World Bank. Ms. Murgueitio will be responsible for monitoring and directing all methodological issues of the project and quality aspects, and for elaborating the final report and supervising its edition. She will have ongoing communication with the technical committee.
- **Field work expert** Ms. **Angela Yamiled Palacio Basto** is psychologist with master degree in population Studies. Ms. Palacio is going to be responsible of methodological aspects. She will have ongoing communication with technical committee. Ms. Palacio will advise all activities related with operative design and logistics, like organization of questionnaire, manuals elaboration, pilot test and training sessions. Ms. Palacio is going to be responsible for monitoring the calendar activities and quality of the project, coordinating elaboration of the partial and final reports. Ms. Palacio is going to be chief of the leader of field teams, the data management specialist with experience in capture and information processing and the survey expert.

- The leader of field teams will be Mr. **Antonio Rey**. Mr. Rey is Cadastral engineer and geodetic with specialization in demographic analysis methods. Now he is doing the master program in population Studies. Mr. Rey will be responsible of field plan and response rates results, quality and fulfillment of commitments, he also is going to be in charge of national schedule of field work. He also will coordinate pilot test and interviewers training. Mr. Rey will supervise organization of questionnaire, manuals elaboration, data cleaning instructions and he is going to be responsible for elaborating the report about conclusions and recommendations of pilot test and executive reports about field work. Mr. Rey will have ongoing communication with supervisors and **field work expert**, implementing suggestions from World Bank.
  - The **data management specialist with experience in capture and information processing** will be Ms. Sonia Roa. Ms. Roa is going to be in charge of the data entry and data cleaning process based on consistency instructions made with team leaders, field expert and World Bank specifications. Ms. Roa also will participate in technical committee meetings and meetings with the World Bank in order to discuss the requirements related with data presentation and final results of the project.
  - Field supervisors or area supervisors will be leaders of interviewers, they will be in charge of work assignment, controlling quality and non-response rate applying the instructions for supervising mentioned in this proposal; they also will be responsible for directing the selection of households, main respondents and test applications in metropolitan areas. Their time involvement will be approximately for four months.
  - Field Interviewers will be responsible for selecting the household member (main respondent) to be interviewed, obtaining information through direct interviews and applying the questionnaire with the manual instructions.
  - The Chief Scorer will be responsible for training of scorers, monitoring the scoring process, consulting with the literacy and numeracy scoring experts following the instructions of the World Bank, the chief scorer also will provide status reports to the World Bank. The chief coder will have a group of qualified scorers in order to facilitate arbitration of scorer disagreement during the rescoring of items
  - Additionally to the team and key staff, SEI s.a. will assign to the project one survey expert. **Hanner Sanchez** will be our counselor in survey selection. Mr. Sanchez will prepare the survey material and he will control the field work related with the survey. He will analyze weekly the information and the indicators of the system, after that, Mr. Sanchez will elaborate reports in order to improve the rate response and solve problems that will be documented.
- 8) **Angélica Becerra** is **statistician**, she is going to support the sample design, the training sessions for field work, and all things related with the household survey and the booklets applications. She also will collaborate with the analysis of the information

and the indicators of the system, and with the reports of field work in order to improve the rate response and solve problems.

<b>I agree with the above,</b>	
<b>PABLO A. ACOSTA</b>	<b>Signature:</b>
<b>CAROLINA MURGUEITIO</b>	<b>Signature:</b>

### 3.0 SURVEY OBJECTIVES

#### **STEP Standard**

*The STEP research program has two international objectives:*

- 1. Develop and apply harmonized survey instruments to: (i) assess the distribution of literacy, non-cognitive, and technical skills in the labor force of middle-and low-income countries and the demand for these skills by employers, (ii) assess the impact of different types of skills on labor market outcomes, and (iii) analyze the extent to which there are skills mismatches in participating countries; and*
- 2. Support country research teams to adapt and implement the surveys in several countries, analyze the results, and identify policy interventions that may be useful to step up the supply of skills sets needed to improve employability and productivity. The application of harmonized surveys in a broad range of country contexts will provide an opportunity to validate findings across countries and distill lessons that may be applicable beyond the countries under review.*

*Each participating country will design and implement its STEP survey to support the above international objectives. In addition, if applicable, each participating country will specify any country-specific objectives that differ from the international objectives.*

#### **Rationale**

A description of the main objectives of the study is required to assure uniformity and consistency in the design and analysis of the STEP across participating countries. It is important that the participating countries share a common set of assessment objectives to facilitate comparisons of the results between countries.

In addition to the main survey objectives, each participating country may define country-specific analytic requirements for the STEP. In this case, the country must ensure that adequate sample sizes will be obtained to allow analyses with acceptable precision to meet these needs. For example, countries may wish to produce survey estimates for special subgroups of the population, in which case additional sample may be required to yield survey estimates with a desired precision.

### 3.1 Major analytic objectives

#### **Report Requirement**

1. Include a list of the STEP international objectives.

Skills are at the core of improving employment outcomes and increasing productivity and growth. Across countries, unemployment and low productivity employment can often be the result of workers not having the right skills to match the requirements in available job openings or having limited opportunities to access high quality pre-employment or skills upgrading training programs. In many countries education and training systems often lack quality and labor market relevance, leaving workers ill-prepared for the labor market.

Against this background, the World Bank launched a multi-country research program that

finances country-level studies to determine how different skill sets affect individuals' labor market opportunities. The studies are expected to fill critical knowledge gaps on the role and demand for different types of skills sets in the labor market and assist in the design of tailored education and training policies to boost employability and productivity.

The research program has the following two objectives:

1. Develop and apply harmonized survey instruments to: (i) assess the distribution of literacy, non-cognitive, and technical skills in the labor force of middle-and low-income countries and the demand for these skills by employers, (ii) assess the impact of different types of skills on labor market outcomes, and (iii) analyze the extent to which there are skills mismatches in participating countries; and
2. Support country research teams to adapt and implement the surveys in several countries, analyze the results, and identify policy interventions that may be useful to step up the supply of skills sets needed to improve employability and productivity. The application of harmonized surveys in a broad range of country contexts will provide an opportunity to validate findings across countries and distill lessons that may be applicable beyond the countries under review.

### 3.2 Country-specific objectives

#### **Report Requirement**

One country-specific objective is to have information sources, that allows to the government quantifies the labor demand and create strategies of education for work.

<b>I agree with the above,</b>	
<b>PABLO A. ACOSTA</b>	<b>Signature:</b>
<b>CAROLINA MURGUEITIO</b>	<b>Signature:</b>

## 4.0 SAMPLE DESIGN FACTORS

### 4.1 Target Population

#### STEP Standard

*The STEP target population is defined as all non-institutionalized persons 15 to 64 years of age (inclusive) living in private dwellings in the urban areas of the country at the time of data collection. This includes all residents except foreign diplomats and non-nationals working for international organizations.*

*There may be exclusions from the target population for practical operational reasons but such exclusions should not exceed 5% of the country's urban population aged 15 to 64 years of age.*

*A country may include other subpopulations in its target population provided that its sample design includes any necessary augmentation of the sample size to accommodate the analysis requirements for these additional subpopulations.*

#### Operational Definitions

- *A Private Dwelling Unit is defined as a room or a group of rooms used, or intended to be used, for living purposes. A dwelling unit must be capable of permanent human habitation and must have a private entrance either outside or from a common hall, lobby, vestibule or stairway inside the building. A private entrance is one that can be used without passing through the living quarters of someone else.*
- *A Household Member is a person who*
  - 1) *considers the dwelling to be their usual place of residence, or who has no usual residence elsewhere;*
  - 2) *makes some common provision for food and other essentials of living;*
  - 3) *spent most of their daily rest at the dwelling for at least nine (9) of the past twelve (12) months; the exception to this rule are persons who have recently joined the household, have no usual residence elsewhere, and intend to spend most of their daily rest at the dwelling.*

#### Rationale

A clear definition of the target population identifies the population of interest for the STEP. This definition is necessary in order to assure that adequate steps are taken to correctly cover the population of interest in the sampling process, and to assure that appropriate and accurate statistical inferences are made using the survey data. Limited exclusions from the target population are not unusual, but should be specified to assure that the survey population is clearly defined and to assure that no extensive biases are introduced due to the coverage of the target population. In essence, the definition of the target population specifies the population from which the sample is to be selected and, consequently, the population to which the sample results may be generalized.

## **Report Requirement**

1. Definition of the target population.
  - a) Specify any exclusion from the STEP target population.
  - b) Include the definitions of concepts related to the survey unit, e.g., dwelling, household, usual place of residence.

The following are considered “institutionalized” and therefore excluded from the STEP survey:

- Residents of Institutions (prisons, hospitals, etc.)
- Residents of Senior Homes and Hospices
- Residents of other group dwellings such as college dormitories, halfway homes, workers’ quarters, etc.

Other exclusions from the target population that are acceptable are:

- 9) Persons living outside the country at the time of data collection, e.g., students at foreign universities.
- 10) Members of the population who are unable to complete the STEP assessment due to a physical or mental condition, e.g., visual impairment or paralysis.

## **Country’s Target Population Description**

The STEP target population is defined as all non-institutionalized persons 15 to 64 years of age (inclusive) living in private dwellings in 13 metropolitan areas of the country at the time of data collection. This includes all residents except foreign diplomats and non-nationals working for international organizations.

On the other hand, one of the most important surveys in Colombia is the household survey “Gran encuesta integrada de hogares”<sup>3</sup> which main results are showed for the 13 principal cities and their metropolitan areas<sup>4</sup>. These areas represent in a good way many aspects of the economy, industry, education and other main issues of the country. In order to compare, next table shows results from household survey for entire country and 13 metropolitan areas.

<b>Indicators</b>	<b>National estimates</b>	<b>13 metropolitan areas estimates</b>
% Working-age population (10 years and over)	78,9	80,9
% Economically active population (working-age population who are either employed or looking for a job)	51,2	54,5
Employment rate	58,5	60,3
Unemployment rate	9,8	10,4
rate of inadequate employment according to skills	6,7	7,4
rate of inadequate employment according to incomes	9,9	10,1

<sup>3</sup>[http://www.dane.gov.co/files/investigaciones/fichas/Gran\\_encuesta\\_integrada\\_hogares.pdf](http://www.dane.gov.co/files/investigaciones/fichas/Gran_encuesta_integrada_hogares.pdf)

<sup>4</sup> [http://www.dane.gov.co/index.php?option=com\\_content&view=article&id=423&Itemid=67](http://www.dane.gov.co/index.php?option=com_content&view=article&id=423&Itemid=67)

So, because for restrictions in the budget and because there are not huge differences between National non-metropolitan areas and the 13 metropolitan areas, the Universe of Study for STEP skills measurement is formed by the 13 main metropolitan areas of Colombia<sup>4</sup>. These areas are composed by 33 municipalities or cities that hold the most of the economical development of the country and the 44% (20.5 million) of the population, 64.6% of them are from 15 to 64 years old. At the end we can make some comparisons for employment.

We also excluded people in institutions such as hospitals or prisons, collective dwellings or group quarters and those who are unable to complete the STEP assessment due to a physical or mental condition, e.g., visual impairment or paralysis. In that sense, based on Census 2005, urban population from metropolitan areas is 58% (18.615.366) of the entire urban population and 12.025.526 million are between 15 and 64 years old. The population distribution is:

		Total		Age between 15 and 65	
		Census - 2005	Projections of population 2012	Census – 2005	Projections of population 2012
Urban population	in Metropolitan Areas	18.615.366	20.475.071	12.025.526	13.226.896
	in Non Metropolitan Areas	13.273.933	14.902.067	8.574.961	9.626.735
<b>Entire urban population</b>		31.889.299	35.377.138	20.600.487	22.853.631
<b>Rural population</b>		10.999.293	11.204.685	7.105.543	7.238.227
<b>Entire population</b>		42.888.592	46.581.823	27.706.030	30.091.858

Source: Census 2005 and projections- DANE

For the survey, according to the standards, a Private Dwelling Unit is defined as a room or a group of rooms used, or intended to be used, for living purposes. A dwelling unit must be capable of permanent human habitation and must have a private entrance either outside or from a common hall, lobby, vestibule or stairway inside the building. A private entrance is one that can be used without passing through the living quarters of someone else.

Additionally, a Household Member is a person who:

- 1) considers the dwelling to be their usual place of residence, or who has no usual residence elsewhere;
- 2) makes some common provision for food and other essentials of living;
- 3) spent most of their daily rest at the dwelling for at least nine (9) of the past twelve (12) months; the exception to this rule are persons who have recently joined the household, have no usual residence elsewhere, and intend to spend most of their daily rest at the dwelling.

2. Specify any special additions to the target population.
  - a) Include the relevant background and rationale for additions.

3. Provide counts (or estimated counts) of the target population by sample design variables.

The Table 1 is an illustration of the population structure per metropolitan area.

**TABLE 1. Universe and population structure**

Metropolitan Area	Quantity of cities	Dwellings	Households (B)	Urban Population (A)	A/B	Population between 15 and 64 years old	Sections	Blocks
BOGOTA	1	1.860.651	2.036.472	6.824.507	3,4	4.408.632	2.803	37.096
A.M MEDELLIN	10	896.729	896.113	3.100.826	3,5	2.003.134	1.348	15.366
A.M CALI	2	538.058	579.692	2.164.046	3,7	1.397.974	948	12.770
A.M B/QUILLA	2	325.802	359.863	1.603.196	4,5	1.035.665	848	11.820
A.M B/MANGA	4	244.740	266.150	983.742	3,7	635.497	525	6.560
A.M CUCUTA	4	177.497	184.401	710.709	3,9	459.118	535	7.056
CARTAGENA	1	194.171	206.634	842.632	4,1	544.340	459	6.083
PASTO	1	92.343	96.364	312.480	3,2	201.862	153	2.318
IBAGUE	1	129.251	136.441	468.378	3,4	302.572	251	3.682
A.M PEREIRA	3	186.865	190.296	571.952	3,0	369.481	379	5.107
A.M MANIZALES	2	116.349	114.686	389.806	3,4	251.815	267	3.174
MONTERIA	1	81.119	84.578	286.631	3,4	185.164	195	3.152
VILLAVICENCIO	1	96.011	102.795	356.461	3,5	230.274	262	3.690
<b>13 Areas</b>	<b>33</b>	<b>4.939.586</b>	<b>5.254.485</b>	<b>18.615.366</b>	<b>3,5</b>	<b>12.025.526</b>	<b>8.973</b>	<b>117.874</b>

SOURCE: SEI s.a based on population and dwelling census 2005 - DANE, basic household questionnaire.

4. Explain any expected deviations from the STEP Technical Standards.

While ideally the team would have liked to survey all urban areas, budget restrictions and

excessive logistic costs to make the survey urban representative, the analysis is restricted to the main metropolitan areas of the country.

## 4.2 Method of Data Collection

### **STEP Standard**

*STEP is a household survey in which the key goal is to conduct an in-depth interview and literacy assessment with one person per selected household. Each component of the STEP Survey will be carried out by a personal visit using a Paper And Pencil Interview (PAPI) method.*

*All components of the literacy assessment must be administered in the same visit (i.e., General Booklet and applicable Exercise Booklet).*

### **Rationale**

The data collection method must be the same for all participants in order to avoid any potential bias that might be introduced, e.g., the data collection method might affect the quality of respondents' answers.

### **Report Requirement**

1. State the method of collection to be used for each survey component, i.e., Filter Module, Household Questionnaire, and Literacy Assessment.
2. Explain any expected deviations from the STEP Technical Standards.

The information shall be obtained through direct interviews done by the interviewer with people who are randomly selected. We discussed with the World Bank the use of DMC or PDA instead of the printed forms and the World Bank approved, but the assessment booklets, i.e., the General and the Exercise Booklet, will be administered using a Paper And Pencil Interview (PAPI) technique. The PDA will be use for the scoring process after the assessment of the general booklet administered using a Paper And Pencil Interview (PAPI) technique, and based on the score the PDA will select on exercise booklet, and again the exercise booklet selected will be administered using a Paper And Pencil Interview (PAPI) technique, and all components of the literacy assessment will be administered in the same visit.

It is important to mention that SEI s.a. has been using DMC to obtain data in the field in the hundred thousands of interviews that they have had the opportunity to do, which ensures that we are dealing with the best of the country in this field and this guaranties a collection process that is not limited to digitalizing the information contained in the forms. It means that we will succeed in ensuring that this information is consistent and that the operation of data collection will gain the support of these tools to control aspects such as the assignation, the coverage and the creation of reports. This strategy allows the relay of partial data during the period of data collection.

SEI s. a. elaborated the application for the intelligent collection of data from the questionnaire

using a robust software which will allow to ensure the respect of the requirements of the study. The program that will be installed in the DMC will have various components.

1. A sample administration menu which contains the data related with the location of the houses and information about the operative control, which makes possible a link with the collection module of the form, the production of reports about unresolved surveys and the creation of backups.
2. A module to collect the data which permits to register the information of every specific survey.
3. A module of data extraction.

Initially, the form defined as the BM will be diagramed to be ready for eventual paper prints, it shall be developed in the same way the instructions of collection defining a friendly interface which shows the questions (one at a time) in the device screen with the aim of controlling the data quality. The previously designed form with its instructions shall be validated during the Pilot Test. Once the final design is approved it is expected to have just marginal failures and the device instructions shall be modified.

The technical staff of SEI s.a. will develop validation mechanisms as documents that explain the logical relation between the different chapters and the variables of the forms to elaborate a sequence of questions and valid codes that makes possible the identification of inconsistencies and mistakes contained in the data. These validation mechanisms shall work as guides to organize the future validation of the digital format.

To execute the project using the proposed methodology, the supervisors will have laptops and the survey takers will have the DMCs with the software developed in CSProX which is specially designed to enter the data and validate it directly during the survey, which offers much more benefits than the physical form because of the following reasons:

- The systematized application of the sub-sampling method, to select the correspondent observation units during each sample design. It is more accurate and avoids possible influences caused by the survey takers lack of objectiveness.
- The use of the DMCs permit to avoid processes as the form printing, the delivery, distribution reception and file of the filled forms and the review, typing and correction of mistakes in the filled forms.
- The review of the data is performed by the system, which has at its disposal the whole of the information provided on the forms, it makes the system able to send an alert to ask directly to the informer and correct the mistaken data. In fact, the program does not allow to the survey taker to continue with the interview if the mistake is not corrected, It also prevent the informer from answering non applicable questions, depending on their specific profile. Therefore the information will be saved in the system without mistakes directly from the field work.
- Nevertheless, SEI s. a. has a policy of data consistence verification before sending it to the costumers, using a softeware which includes some of the variables non-susceptible to be included during the survey application, which guarantees to the costumers the delivery of depurated information.

- Permits to focus the supervision in substantial matters more than in formal matters as the collection of data and the improvement of the survey takers techniques of collection, etc.
- The confidentiality is guaranteed in an effective way because the collected information are no visible for third parties, not even the survey taker is able to access to the data after saving it on the computers. The experience of SEI s. a. indicates that the informers give more attention and consider more important the surveys taken with this kind of technology.

### 4.3 Response Rate

#### **STEP Standard**

*A minimum response rate of 70% is the goal.*

*The method for calculating the STEP response rate will be consistent across participating countries. The overall survey response rate is defined as “the result of dividing the total number of complete interviews by the total number of ‘unduplicated, in-scope’ sampled individuals”.*

#### **Rationale**

The first sampling priority is to obtain the prescribed minimum number of STEP interviews per STEP reporting language. Secondly, an overall response rate of 70% must be achieved. Generally, in surveys that employ a personal-interview data collection method, one might expect to achieve a response rate in excess of 80%. However, it is realized that the response rate for STEP may indeed be lower than other surveys due to the fairly lengthy interview and psychometric assessment that some respondents may find intimidating. Nevertheless, in previous international surveys with a literacy assessment component, a low response rate was identified as a data quality concern and an area for improvement. Achieving a response rate of 70% may be a challenge for some participating countries, but should not be considered impossible to attain. As well, the credibility and quality of the survey results requires a reasonably good survey response rate.

#### **Report Requirement**

1. State the expected response rate for STEP.
  - a) Explain the rationale for the expected response rate.
2. Explain any expected deviations from the STEP Technical Standards.

The response rate for the STEP Survey in Colombia is expected to be 70%. This overall response rate takes into account an expected assessment CORE exercise ‘Pass’ rate of 95%.

This expected response rate are based in similar household surveys made by the firm where response rate was around 70% - 80%. Expected Core pass rate was formulated based on the literacy rate of 94.1% in 2009 (<http://www.mineducacion.gov.co/observatorio/1722/article-226309.html>)

## 4.4 Sample Frame

### **STEP Standard**

*The sampling frame should provide coverage of the target population so that the number of unique, in-scope survey units on the sampling frame comprises at least 95% of the target population.*

*Upon receipt of the sample frame and agreement of its suitability for STEP sampling, the STEP Consortium will select the sample of PSUs, including a reserve sample of PSUs for use when it is not possible to conduct any interviews in an entire initially-selected PSU.*

*If there is no recent Census available (i.e. less than 3 years old), the firm will carry out a 'door-to-door listing of households' in each selected PSU (see OM for details).*

### **Rationale**

In essence, the sampling frame is the list of the population from which the STEP sample will be selected. The sampling frame defines the coverage of the target population and provides access to the selected sample. The frame can be a source of non-sampling errors, such as error due to under-coverage or over-coverage of the target population, or errors due to duplication of population members on the frame. Since the frame provides the means to identify and locate selected population members, the quality of the information on the sampling frame directly affects the quality of the selected sample as well as the data collection operation. Therefore, the information contained on the survey frame must provide acceptable coverage of the target population and be complete, accurate, and up-to-date.

It is essential to ensure that the sampling frame provides acceptable coverage of the target population, and satisfactorily meets the requirements for sampling, locating selected population members, and for estimation purposes.

### **Report Requirement**

1. Description of the sample frame
  - a) Frame type (e.g., population register, household list, list of geographic units, etc.)
  - b) Source of frame (e.g., 2006 Census, Labour Force Survey, etc.)
  - c) Definition of survey units on the frame for each stage of sampling.
  - d) Data items on the frame for each stage of sampling (e.g., name, address, age, gender, education, etc.)
  - e) Identify the variables to be used for stratification if applicable.
  - f) Provide survey frame counts by stratum and type of survey unit (e.g., PSUs, dwellings) as applicable to the sample design.
  - g) Quality assurance procedures (i.e., assessment of quality of frame information)
    - i) Provide any information regarding known frame issues, e.g., under-coverage of target population, inclusion of out-of-scope units, up-to-date, duplication, etc.

- ii) Explain any steps to ensure that the frame is complete and up-to-date.
- h) Explain any expected deviations from the STEP Technical Standards.

### 1. Sample Frame

The firm will select the sample frame of PSUs

The firm proposes to use the frame of areas from DANE, which contains the information of the Population and households per block based on the Census performed in 2005 and the geo-statistical information for the urban level. Also, based on Census 2005, the frame has the identification of sections and blocks for each metropolitan area. To select households, a dwelling frame in each block will be made; the staff will check and recount (using the format of identification and location of the dwellings), so in each block selected we are going to do a visual listing operation per block in order to select 8 households per block.

### **DWELLING AND HOUSEHOLD LISTING OPERATION**

There are some cartographic information for identifying and locating big areas and blocks in the urban area, but we cannot have the list or the information for edifications, dwellings and households in that areas or blocks. So, we need to construct the sample frame of dwellings and households in every selected block after visual listing operation of edifications, dwellings and households in dwellings. The procedure for the sample frame construction is to make a listing in every block selected to identify dwellings. That is mean we are going to do a visual listing of dwellings in selected blocks, and after we select dwellings we are going to do a door-to door listing of households in each dwelling selected.

The listing format is going to do by de interviewers but validated in order to verify the right application of the randomized selection algorithm implemented for the selection of dwellings and households, for the supervisor. This means that the validation of this process is one important issue in field supervising because if the interviewer has differences (more than 25%) between listing and the information in the census of 2005, the listing operation has to be validated for the supervisor. After the validation of the listing operation, the supervisor is going to select the dwellings trough an application in the laptop.

#### 4.5 Sample size – Full Assessment

##### **STEP Standard**

*The sample size requirement for each STEP reporting language population is as follows:*

- *A minimum of 2,400 complete STEP interviews for each STEP reporting language target population are required.*
- *In addition, the final sample must include for each STEP reporting language target population at least 600 completed cases for each of the four exercise booklets.*

*A case is considered complete for inclusion in the required sample size if it satisfies the following conditions:*

- 1) *All modules in the Household Questionnaire have been administered.*
- 2) *All items in the General Booklet were attempted.*
- 3) *All items in the assigned Exercise Booklet were attempted.*

*Each participating country will develop and implement procedures to regularly monitor the sample returns during data collection to ensure that the sample size goals are achieved.*

##### **Rationale**

The standard sample size is the minimum required to ensure the stabilization of the theoretical model that is used to produce the estimates of plausible literacy levels. The STEP minimum sample size requirements must be met to ensure that the estimates produced from STEP can be generalized to the population from which the sample is selected, and that these estimates have an acceptable level of precision while meeting a minimum response level criterion.

##### **Report Requirement**

1. STEP target sample sizes
  - a) Provide the country's final sample size goal by sample design variables, e.g., by strata, by PSU, etc..
  - b) Provide the country's overall initial sample size, including the size of the reserve sample, by sample design variables, e.g., by strata, by PSU, etc..
  - c) Describe the basis for the size of the reserve sample, e.g., non-response expectation, design effect.
2. Provide the rationale for additions to the sample size to satisfy country-specific data analysis objectives.
  - a) What are the data analysis objectives? For example, identify the important data breakdowns or survey estimates to be derived from the survey data.
  - b) What are the precision goals for the survey estimates?
3. Sample monitoring procedure.

- a) Describe the planned strategy for monitoring the sample returns to ensure that the sample size goal is achieved.

## 1. STEP Target Sample Sizes

### Sample Size

According to the terms of reference, the firm proposes a sample size of 2.400 households from, selecting a person in each home. The size of the sample is distributed in a proportional way to the population size in the formed stratum, therefore there shall be approximately 243 cartographical sections of the population census performed by DANE in 2005 and in each one of them, 2 blocks in average shall be selected, to give a final selection of 481 blocks, as it appears in the following:

Cuadro 3. Sample units distribution

Metropolitan Area	Quantity of households and people to be selected	Quantity of Blocks to be selected	Quantity of Sections to be selected
1. A.M Bogotá	878	176	88
2. A.M Medellín	405	81	41
3. A.M Cali	265	53	27
4. A.M Barranquilla	193	39	20
5. A.M Bucaramanga	123	25	13
Probabilistic stratum (we are going to select 4 metropolitan areas from 8)	536	107	54
<b>Total</b>	<b>2.400</b>	<b>481</b>	<b>243</b>

According with the proposal adjustments after contract negotiation with World Bank, we want to clarify some aspects about sample design:

- 11) 9 metropolitan areas (5 are certainty PSU and selecting 4 from 8 )
- 12) 240 sections in 9 metropolitan areas selected.
- 13) 480 blocks in 240 sections selected.
- 14) In average we are going to select 7.1 dwellings in each block because we are expecting a non-response rate around 30%. So the sample size of entire people is 3429 but because the non-response rate that we're expecting, at the end we will expect to have 2400 people with the core booklet applied.
- 15) 1 household in each dwelling
- 16) 1 person between 15 and 64 years old in each household

The sampling goal is to obtain 600 complete cases for each of the four Exercise Booklets.

A case is considered complete for inclusion in the required sample size if it satisfies the

following conditions:

- 1) All modules (1 to 8) in the Household Questionnaire have been administered.
- 2) All items in the General Booklet were attempted.
- 3) All items in the assigned Exercise Booklet were attempted.

The firm will regularly monitor the sample returns during data collection to ensure that the sample size goals are achieved.

**I agree with the above,**

**PABLO A. ACOSTA**

**Signature:**

**CAROLINA MURGUEITIO**

**Signature:**

## 5.0 SAMPLE DESIGN

### **STEP Standard**

- 1) *A probability sample design whereby each person in the survey population has a known (i.e., calculable), non-zero chance of being included in the sample must be used. As well, the sample selection process must be objective (i.e. a random selection method must be used) at all stages of sample selection.*
  - a) *The ‘preferred’ sample design is a multi-stage design that employs sampling with probability proportional to size (PPS) for as many stages as practically possible.*
  - b) *The selection of households (15 original and 15 reserves) within selected PSUs will follow STEP Consortium guidelines.*
  - c) *All countries must use the same procedure for selecting a household within a multi-household dwelling, if applicable. The procedure will be provided by the STEP Consortium.*
  - d) *All countries must use the same procedure for selecting a person within a household. The procedure will be provided by the STEP Consortium.*

### **Rationale**

The development of a sample design should consider the STEP objectives as well as methods of data collection and the relative cost of the data collection. An appropriate sample design should be driven by the desire to obtain the best precision possible for the stated sample size balanced against the need to establish a highly efficient data collection. Only probability sample designs are based on recognized sampling distribution theory, permitting the estimates derived from the survey sample to be legitimately generalized to the population from which the sample is selected. Also, only with a probability sample design can the sample data be used to produce estimates of measures of precision of the survey estimates, such as the coefficient of variation, the standard error, or the margin of error.

As well, the credibility and quality of the survey results requires a reasonably good survey response rate.

### **Report Requirement**

1. Description of the probability design to be used, including any stratification and multi-stage sample design considerations
  - a. Number of sampling stages.
  - b. Describe sampling unit at each stage of selection.
    - i. Provide counts of sampling units for each stage.
  - c. Describe the procedure for sample selection within a household if applicable.

## 5.1 SAMPLING METHODS AND SELECTION ALGORITHMS

Initially, during the construction of any sampling plan, it's important to access information highly related with the main variables of study to think about the designs and the diverse aspects to estimate and determinate the way to develop the most efficient proposal.

Consequentially, the sample design proposal is a probabilistic kind, stratified, multistage and with elements. So, sampling stages are:

- First stage: Metropolitan areas.
- Second stage: Sections
- Third stage: Blocks
- Fourth stage: Dwellings
- Fifth stage: Households.
- Sixth stage: Person

We proposed this kind of sample design in order to optimize budget, and based on that we could not include all metropolitan areas, so we proposed a stratification method (Geometric-Method by Gunning and Horgan - 2004), and for the experience in other projects we have proved its efficiency selecting a sample of areas. With results of different simulations, we proposed two strata, one of them are certainty PSU (we mean PSU with inclusion probability = 1) and this stratum includes the main five metropolitan areas (Bogotá, Medellín, Cali, Barranquilla and Bucaramanga) and the second stratum is probabilistic with 8 PSU and we are going to select 4 areas from 8.

In general in Colombia, we use a geostatistical frame for this kind of researches and we also include selection of sections and blocks. And for technical and logistic reasons, we include listing and selection of dwellings, households and people because we don't have this kind of information (frame of elements) from Census.

Our proposal considers the stratification method in order to obtain an efficient design according with the budget. And although the 13 metropolitan areas are 33 municipalities, in fact each metropolitan area is a unit actually a sample unit; the 33 municipalities can't be selected separately because when we make a selection, we can have municipalities selected in each metropolitan area so that can increase the budget .

The firm will use the stratification technique, which is a mechanism that permits to adjust the samples designs to the conditions of the formed stratum and the requirements of the research staff. To define the corresponding stratum, the firm used the geometrical algorithm of stratification developed by Gunning and Horgan (2004) and used as a free software R by Marcin Kozak<sup>5</sup>. This method defines the stratum in a way the variation coefficient (CV) is less than a determinate value in each stratum. Specifically, the variation coefficient should be equivalent to less than 3% for each one of the stratum (look at the table 2).

---

<sup>5</sup> Kozak M y Verma M. Geometric versus optimization approach to Stratification: A comparison of efficiency. Survey Methodology, Vol 32 No. 2 pp 157-163. 2006

**TABLE 2. Stratification method results**

Stratum	Stratum boundaries proposed by the method	Anticipated means in each stratum	Anticipated variances in each stratum	Nh (Number of units in each stratum)	nh (number of units to sample in each stratum)
Stratum 1	995.133	541.068	3.3 e10	8	4
Stratum 2	2.597.248	1.583.181	22.1 e10	3	3
Stratum 3	6.778.692	5.047.525	299.7 e10	2	2
Total sample size in the first stage:		9 metropolitan areas			
Anticipated population mean::		1.474.857			
<b>Anticipated CV</b>		<b>2,70%</b>			

**Source of information:** R, package “stratification” by Marcin Kozak.

In the second phase of sampling, after having selected the 9 metropolitan areas (5 of forced inclusion and 4 of probabilistic inclusion), a selection of the cartographic sections of the population census applied by Dane in 2005 shall be performed using the proportional method of the quantity of population between 15 and 64 years old (PPT Design). Subsequently, during the third phase using the PPT method, once again cartographic blocks shall be selected according to the same age sample of population. We are going to do a listing in every selected block, and using a simple random sampling without replacement (SSR design) to select 8 houses in each block. In Colombia, most household dwellings contain a single household, but in case of a multi-household dwelling, the selected one will be done randomly. Each home shall be identified through the registration of the name of the head house in order to select a home in every house based in a SSR design. At the end, after the identification of all of the home members, the survey takers will select randomly (through a programmed random algorithm in the PDA) select a person between 15 and 64 years old, who are mentally capable, to take the survey and the correspondent cognitive and abilities test. This sampling design was made based on the budget and in order to facilitate the implementation in field work.

### **STRATEGIES TO MANAGE THE “NO ANSWER”**

The adjustment because of “No coverage” or “No Answer”, is valid when the conditions are given during the sampling phases. This kind of adjustment permits to improve the validity of the estimations, because it allows the staff to correct the defects of the sampling framework and the lack of answer due to diverse kind of reasons as: reject, the suitable informer is busy, absent or other reasons. The adjustment is possible if it is sustainable the hypothesis of obtaining an average of similar results from the non-surveyed people and the surveyed people inside of the same sample unit (UPM, USM o UFM).

For the consecution of the 2.400 surveys, the firm proposes to estimate an over-sample that may be enough, bearing in mind the following recommendations:

- ✓ The estimation of the over-sample has to be different depending on the socioeconomical level of the people. Households in PSUs with high socio-economical status have a high

rate of non-response, so we have to select a larger sample in these PSUs to account for higher non-response of high socio economic status.

- ✓ The over-sample has to be modified according to the performance of the survey takers during the field work.
- ✓ The first estimation of the general over-sample has to be around the 25%-30%, according to the results of similar socioeconomical studies with similar application length.

Our proposal is orientated in obtaining a limited and random amount of “non response” answers to control the subjectivity of the sample. The sample shall therefore be greater in number (3429 people) to ensure that the survey will be composed of a minimum of 3000 people while taking into account the maximum percentage of “non response” answer of 30%.

On the other hand, the reports of field work will be the referent in order to decide what is affecting the response rate. Specifically for previous experiences, the reports establish the responses rates depending on the socio-economical level: in the high level we have to select more blocks or households in order to guarantee a minimal coverage, which in the Colombian case is characteristic of high socio-economic level population.

## **5.2 Sample Allocation and Selection**

We will make a stratification of the thirteen metropolitan areas, 5 metropolitan areas are going to be with certainty. The sample allocation is proportional to x-total, that is mean people between 15 and 64 years old.

With probability proportional to size (PPS); in this case proportional to the population between 15 and 64 years old; we are going to select 4 metropolitan areas from 8. In the second stage we are going to select 243 cartographical sections from the geostatistical framework in the last census with PPS and from those sections, we are going to select 481 cartographical blocks again with PPS, and we are going to use the cumulative total method for PPS selection. Using a Simple Random Sampling without replacement, in every block are going to be selected probabilistically dwellings, households, and each person in selected households. Dwellings will be selected systematically; households and person in selected households will be selected through negatively coordinated

### **5.2.1 Random Selection of Assessment Exercise Booklet**

The Firm will randomly select one of the four Exercise Booklets to be administered to a respondent through the PDA in order to have at the end the same quantity of each exercise booklet.

**I agree with the above,**

<b>PABLO A. ACOSTA</b>	<b>Signature:</b>
<b>CAROLINA MURGUEITIO</b>	<b>Signature:</b>

## 6.0 LITERACY ASSESSMENT

### **STEP Standard**

*A participating country will implement the full literacy assessment design as prescribed by the STEP Consortium.*

*The General Booklet comprised of the Reading Components and the Core Literacy Items will be administered to each selected person.*

*Each interviewer will use a stop watch to time the Reading Components exercises in the General Booklet.*

*Subsequently, for those respondents with 3 or more correct answers for the Core items, one of the four Exercise Booklets will be randomly selected to be administered to the respondent.*

*The assessment component of the survey ends with the administration of the General Booklet for respondents with less than 3 correct answers for the Core items.*

### **Rationale**

In order to ensure that the STEP assessment results are comparable across participating countries it is essential that the assessment be consistently administered in all participating countries.

For countries that are implementing a full literacy assessment, respondents that pass the Core (Part B of the General Booklet) will be routed to the second part of the assessment, which consists of four exercise booklets. However, each respondent must complete only one of the four booklets, and that booklet must be randomly assigned to the respondent.

The STEP Consortium will be responsible for the sampling of first stage sample units for each participating country and will provide a sample file that will indicate the Exercise Booklet to be assigned for each of the sampled cases in the initial sample of <3,000-4,000> target population. When there is a non-response case that is replaced by a reserve sample unit, the supervisor responsible for the release of the reserve sample unit will ensure that the Exercise Booklet # assigned to the reserve sample unit is the same as the Exercise Booklet # assigned to the initial sample unit that is being replaced.

For example, consider the case of a PSU with 15 households initially sampled. If a selected household that was assigned Exercise Booklet #3 is determined to be a non-response case after the appropriate follow-up attempts then the supervisor will release a reserve sample to replace this non-response case. In this situation, the supervisor will ensure that the replacement reserve sample unit is assigned Exercise Booklet #3, the same booklet # that was assigned to the originally sampled household.

### **Report Requirement**

1. Description of the STEP assessment to be implemented.
2. Languages in which assessment will be administered.
3. Explain any expected deviations from the STEP Technical Standards.

Colombia will implement a full literacy assessment in Spanish.

The Interviewers will be trained to administer the assessment according to the instructions provided by the STEP Consortium.

Interviewers will not use a stop-watch, since the PDA will be programmed to have a very simple application (“start”/“stop”) to time the Reading Components exercises in the General Booklet.

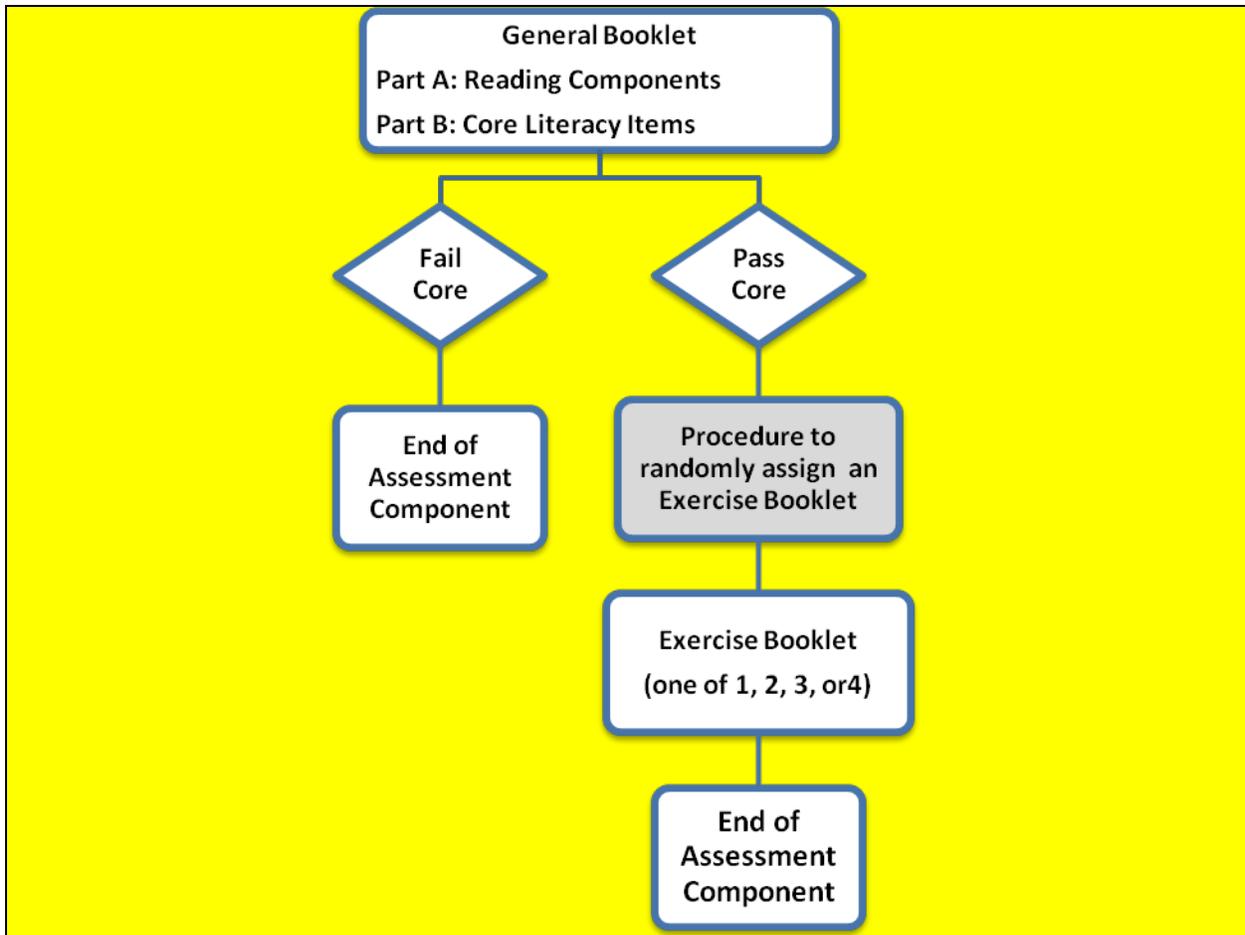
The assessment portion of the STEP Survey will include the administration of a General Booklet to each respondent. The General Booklet is comprised of two sections,

- Part A is an assessment of reading skills.
- Part B is a set of CORE literacy items.

The Interviewer will score the Core items during the interview. If a respondent gets 3 or more correct answers to the Core items then one of the four Exercise Booklets will be randomly chosen to be administered to the respondent. If a respondent gets less than 3 correct answers to the Core items the assessment component is terminated, i.e., no Exercise Booklet will be administered to the respondent.

The workflow for the full assessment component of the interview, as described above, is shown in Figure 1 below.

**Figure 1. Proposed Workflow – Full Assessment**



## 6.1 Translation and Adaptation of Literacy Materials

### **STEP Standard**

*Participating countries are responsible for the translation of the assessment instruments and their adaptation to national circumstances.*

*The translation and cultural adaptation of the assessment items will be carried out according to the translation and adaptation guidelines prepared by ETS.*

*The recommended procedure for developing the national versions is double translation by two independent translators, followed by reconciliation by a third translator. It is also recommended that translations be reviewed by a national panel of domain and/or survey experts.*

*Each participating country will prepare STEP national assessment booklets modeled after the master assessment booklets provided by the STEP Consortium.*

*Each country will submit its translated literacy booklets in each of its STEP reporting languages to the STEP Consortium for verification and approval.*

### **Rationale**

In order to ensure that the STEP assessment results are comparable across participating countries it is essential that the same skill assessment information is collected across countries. If valid comparisons of assessment results are to be made across countries, the equivalence of different language versions of the assessment instruments is essential. For literacy test instruments, “equivalence” refers to semantic equivalence (content), as well as equivalence in terms of register, style, readability and other characteristics likely to affect literacy performance.

### **Report Requirement**

1. Description of the Translation process for the literacy materials.
2. Languages in which the assessment materials will be translated.
3. Explain any expected deviations from the STEP Technical Standards.

Each participating country will prepare STEP national assessment booklets modeled after the master assessment booklets provided by the STEP Consortium. In other words, the number of pages, the numbering and order of pages, the layout of stimulus material and directives, the graphics, the response format, the text format, and the print quality will all be the same as in the master assessment booklets provided by the STEP Consortium.

The literacy exercises are provided to the Survey firm/agency directly in a Word document. These should be translated respecting the text size and formatting. The recommended method of translation is the following: independent translation of all text by two translators who have no contact with each other, and one reconciler will reconcile the two translations, keeping the best parts from each of them. The reconciler will be responsible for creating a single form of the Verification Follow-up Form (VFF) that includes comments from either or both translators as well as his or her own comments. The final VFF form together with the translated materials will be then submitted to the World Bank, who will then send the materials to ETS, the organization that is providing the literacy exercises for the STEP program. The materials will then be sent for independent verification with feedback provided to countries about the translated versions of their materials.

We did not follow the recommended method of translation. The literacy exercises were provided by the World Bank to the Survey firm directly in a Word document in Spanish. These were adapted respecting the text size and formatting. The firm used the Verification Follow-up Form (VFF) that included comments from either or both translators as well as his or her own comments.

Questionnaires and interviewer’s manuals were translated for the Bolivia survey and adapted by the firm to adjust the Spanish to the Colombian case.

## 6.2 Printing of Literacy Assessment

### **STEP Standard**

*Each country will ensure that the approved assessment booklets from ETS are printed according to the printing instructions provided in the document 'STEP\_Paper\_Booklet\_Printing\_Specifications.doc'.*

*The assessment booklets must be printed (NOT photocopied) as saddle-stitched booklets so that each page of the final printed booklets is either a letter-size page or an A4-size page.*

### **Rationale**

It is critically important that each country print the assessment booklets in the same manner using the same paper size.

### **Report Requirement**

1. Description of the printing criteria that will be followed.
2. Include information on who will print the booklets.

Each participating country will ensure that the booklets are printed according to the ETS criteria.

The number of pages, the numbering and order of pages, the layout of stimulus material and directives, the graphics, the response format, the text format, and the print quality will all be the same as in the master assessment booklets provided by the STEP Consortium.

The booklets will be printed in the firm's headquarters with a special printer machine with the requirements and the standards from ETS (as saddle-stitched booklets so that each page of the final printed booklets is either a letter-size page or an A4-size page).

<b>I agree with the above,</b>	
<b>PABLO A. ACOSTA</b>	<b>Signature:</b>
<b>CAROLINA MURGUEITIO</b>	<b>Signature:</b>

## 7.0 HOUSEHOLD QUESTIONNAIRE

### **STEP Standard**

*Each participating country will implement the Household Questionnaire prescribed by the STEP Consortium.*

### **Country-specific Questions**

- 1) *Each country is permitted to insert up to 5 country-specific questions in the Household Questionnaire. The questions and their placement must be confirmed with the STEP Consortium.*
  - a) *The rationale for the inclusion of country-specific questions must be provided.*
  - b) *The wording of the proposed questions must be included in the NSDPR.*
  - c) *The placement of the questions must be discussed with the STEP Consortium.*
  - d) *An outline of the pre-test strategy for the questions must be included in the NSDPR.*

### **Rationale**

The household questions must have the same meaning for respondents in all participating countries despite differences in language and culture. A core set of questions with standard concepts and definitions related to the survey objectives is necessary to allow comparability of the survey results between participating countries. Since there are many participating countries, each with its own language and culture, a standard translation procedure is also critical to ensuring that the household questions do indeed have the same meaning for respondents, survey researchers and data users.

### **Report Requirement**

The basis for the Household Questionnaire applied to the Colombia case is the one prescribed by the STEP Consortium.

Some questions were included in order to make some comparisons with national studies:

- a) Module 1-B: How do you dispose residuals?
- b) Module 2: What is the Maximum Degree Received?
- c) Module 2: What was the duration of the last two courses you took?
- d) Module 2: In which year did you finalize these two courses?

## **PILOT TEST**

The Pilot Test is an event in which is designed to develop the processes used in the application of surveys in the field and with a goal to validate the instruments and the evaluation used, as well as, adjust any changes needed for the actual field work. Before of pilot test, the pre-test strategy it is to apply some questionnaires to some staff of the firm. Previous to that, the team revises all questionnaire and questionnaires from national surveys in order to adjust.

To avoid subjectivity in the evaluation of the Pilot Test, the conditions in which it is set must replicate the conditions in the actual field, specifically regarding the forms and the manuals that are used. They must also replicate the process of sampling and the verification of the surveys to help determine their limitations and establish the diversity of situations that may come up.

In this context, one disposed of the following conditions to guaranty the validity of the results:

- Prepare the instruments and processes used for the Pilot Test in the same manner that the real test would be prepared. This is in reference to the creation of the diagram of the form, and the norms and processes contained in the instructions.
- Define a methodology for the Pilot Test. More specifically, design a technical schematic for its execution which established the objective parameters and allows for a quantification of the obtained results. In this case, it is necessary to determine what shall be evaluated and how, as well as, the analysis that shall be done on the information obtained.

**The sample for the Pilot Test.** The homes and people with which shall be applied the Pilot Test cannot be part of the sample. This is done to avoid future conditioning of the answers.

The Temporary Union is considered vital for the project and for the support and participation of the World Bank in the different phases such as the formation, execution and evaluation of the test.

### **1. Translation of the Household Questionnaire**

The required method of all questionnaire translations is the following: independent translation of all text by two translators, who have no contact with each other. A third translator will reconcile the two translations. (Before the reconciliation of the questionnaires, the questionnaires will be put into their proper questionnaire format in Excel by the World Bank). This reconciliation by a third translator will ideally be done in collaboration with a senior member of the survey team who is strong in English and who is involved in the survey piloting and fieldwork, as well as a representative of the World Bank if possible.

**The World Bank regional team has asked the firm to include the following additional questions:**

- Dwelling Section:

- 1) How do you dispose residuals?
- 2) To which socioeconomic strata this household belongs?

- Education Section:

- 3) What is the maximum degree received? (for secondary and tertiary education)
- 4) What was the duration of the course? (for training courses)

**I agree with the above,**

**PABLO A. ACOSTA**

**Signature:**

**CAROLINA MURGUEITIO**

**Signature:**

## 8.0 DATA COLLECTION

### STEP Standard

*Each country will develop a data collection strategy that incorporates a survey promotion strategy, a contact strategy, a response rate strategy, an interviewer hiring and training plan, interviewer supervision procedures, and field quality control procedures.*

*Key elements of the data collection strategy are the following:*

- 1) *A minimum response rate of 70% is the goal.*
  - a. *The method for calculating the STEP response rate will be consistent across participating countries. The overall survey response rate is defined as “the result of dividing the total number of complete interviews by the total number of ‘unduplicated, in-scope’ sampled individuals”.*
  - b. *In cases of non-contact and temporary absence, at least three follow-up attempts are required before classifying a case as a non-response.*
  - c. *A supervisor must attempt to convert refusal cases.*
- 2) *The Interviewer training should last a minimum of 10 full training days, including field practice where each trainee will interview at least two households and two selected individuals.*
- 3) *A field supervisor will revisit each household in the following situations:*
  - a. *A household refuses or does not begin the interview because of special circumstances (result codes 1 or 2).*
  - b. *A household stops before finishing the Household Module, Module 1.*
  - c. *A household where the selected individual is not able to begin the questionnaire – for refusal, for special circumstance, absence, other reasons.*
  - d. *A household where the individual stops without finishing the individual*

*modules 2-7.*

- e. A household where the individual stops without finishing the assessment exercises, Module 9.*
- 4) *A verification of an interviewer's visit will be carried out by a revisit to 30% of the households in each interviewer assignment, consisting of a personal revisit to 15% of the finalized cases and a telephone follow-up to the remaining 15% sample of finalized cases.*
  - a. The households involved in the verification process will be randomly selected within each PSU.*

### **Rationale**

The collection of data from respondents should be as consistent as possible so that potential bias may be minimized. There is a need to ensure that the interviewers have the necessary material for selecting a respondent within a household and the survey instruments are administered uniformly by all countries.

The participating countries consider literacy to be an important topic. The survey results will reflect on the image of the participating countries. Raising public awareness of literacy and the STEP survey through a public promotion campaign should result in a more informed population that will hopefully be more cooperative in participating in a burdensome data collection effort. In any survey, respondents are usually more cooperative when they are provided information pertaining to the survey purpose, the survey sponsor, the use of the data, etc.

A well-formulated contact strategy is important to ensure that interviewers make every effort to reach selected individuals. Such a strategy is essential to maximize response rates and thus lead to quality data.

Whenever there is any non-response to a survey there is a possibility that non-response bias may exist in the survey results. Non-response bias occurs when the non-respondents differ from the survey respondents with respect to important characteristics. If this is the case, the survey researchers and data users should not assume that the respondents' data is necessarily representative of the target population. Although such non-response bias can occur whenever there is any non-response, the risk of such an occurrence increases as the response rate decreases, i.e., as the number of non-respondents increases. Therefore, the success of the STEP requires that each country develop a strategy to minimize non-response.

A key ingredient in the success of STEP is the interviewing staff, which has a direct bearing on the quality of the data collected. Each interviewer must be given an assignment that is large enough to make it financially worthwhile but at the same time is not so large that it is difficult to complete on time. Interviewers should also be fairly paid for the number of hours that they actually work rather than being remunerated on a piece-meal basis according to the number of completed interviews achieved. If paid on a piece-meal basis there is an increased risk that the quality of an interviewer's work may suffer, e.g. an interviewer might consciously or sub-consciously rush to complete interviews without due regard to the quality of the data collected

from respondents. In addition, the interviewer supervision is required to ensure that the interviewer work is of acceptable quality, to uncover potential problems that may have an impact on the survey data, and to provide opportunities to receive and provide interviewer feedback.

### **Report Requirement**

1. Survey promotion strategy
  - a) If applicable, briefly describe the planned activities for public awareness.
2. Contact strategy
  - a) Survey promotion and advance materials.
  - b) Do you plan to conduct initial household/respondent contact in-person, via telephone, or both?
3. Response rate strategy
  - a) Briefly describe the methods to be used to minimize non-response.
  - b) Respondent incentive.
4. Interviewer hiring plan
  - a) Describe the desired interviewer characteristics (e.g. number of years of survey experience, familiarity with computers, etc.).
  - b) No. of interviewers
  - c) Method of payment
5. Interviewer training plan
  - a) Describe your proposed training approach for train-the-trainers, supervisor training, and interviewer training. For each, provide the following information:
    - i) Training dates (given as number of weeks prior to data collection);
    - ii) Location of training (site and city);
    - iii) Number of hours of classroom training, home-study; and
    - iv) Whether all trainees will be trained in one session or in multiple sessions (such as in various locations around the country).
  - b) Number of Trainers.
    - i) Trainer background/experience.
  - c) Training evaluation.
6. Interviewer supervision procedures
  - a) No. of supervisors (senior interviewers)
  - b) Responsibilities
  - c) Indicate the methods of staff communication (i.e., scheduled weekly telephone calls, e-mail, newsletters, etc.) proposed for data collection.
  - d) Interview validation – percentage of cases.
    - i) Techniques to monitor interviewer performance.
7. For each item above, explain any expected deviations from the STEP Technical Standards.

The listing process is going to be a survey promotion strategy, because it is a great opportunity to become sensitive the population to avoid non-response and improve the collaboration with the project, and specially, in order to make special management in blocks with residential complex with restricted entry.

The field work groups will be formed with 4 interviewers for each supervisor. In total, there will be 8 groups and 32 interviewers. All shall depend directly on the Project director who shall centrally control the field work insuring a schematic of telephonic verification. This verification will be of the collection of data as well as the verification of various municipalities. The Director of data will have the support of an area supervisor who shall lead the groups and serve as a link between them and the central control.

**Cuadro 5. Distribution of field work groups**

Metropolitan Area	Quantity of households and people	Field work groups
1. A.M Bogotá	850	1,2 y 3
2. A.M Medellín	410	4
3. A.M Cali	270	5
4. A.M Barranquilla	200	6
5. A.M Bucaramanga	130	
Probabilistic stratum	540	7 y 8
<b>Total</b>	<b>2.400</b>	<b>8</b>

The project shall be composed of the directive section of permanent staff members from the temporary union and also of the coordination of the collection work, as well as, part-time personnel in charge of supervising or of taking the polls.

At first, the minimum requirements for the survey takers is to have the appropriate bachelor, although it is desirable and therefore a preference shall be given to those who have experiences in home surveys. Also preferred are people which technical, technological or professional studies. In any case, the candidate must be able to provide the complete time allocated to the project and be ready to travel in the event that this would be required. Also, the candidate must demonstrate, once chosen, the adequate use of the information collection program installed in the DMC.

For the supervision of the field work, preference shall be given to those with experience in field work supervision and oversight of home surveys.

The **recruitment** is the process of identification of interested candidates suitable for the project. The work starts by identify the principal sources from where shall be obtain the necessary personnel. To optimize the resources at the disposition of the project, only the personnel that fit the adequate profile shall be contacted. In any case, there are only two

sources that shall be used: internal (using the *Bitácora Digital* (Digital Blog) which is an Information System developed by SEI s.a. to store the data of anyone that has worked for the company, including their experience certifications and their academic studies) and external, for example, through ads on the internet or in universities.

Also considered is the possibility to recruit survey takers that are residents in different cities of the country with the motive of reducing the costs, although the training shall be conducted at a central level to avoid the loss of the unit of criteria and therefore, a higher possibility of error. In the case of the supervision personnel, the majority shall be recruited in Bogota because the firm prefers to count on trusted personnel that have experience in the domain.

The **official announcement** is the invitation sent to the candidates so that they may send their C.V.s to the SEI s. a. offices with the adequate support. This process is complete using the “Sistema de Hojas de Vida” (A System to receive the C.V. of applicants to work in the company) program which is a WEB designed system to which one can gain Access through the company network (intranet) or using the link “Trabaje con SEI” (Work with SEI ) of the company’s webpage [www.sei-consultores.com.co](http://www.sei-consultores.com.co).

As well as the Curriculum Vitae analysis and the demanded requirements, the candidates shall be submitted to evaluation tests during the training. These tests are considered to be an important part of the selection process

The **training** of the supervisors will include additional sessions on their specific functions and the field work procedures, the quality control and the coverage of the surveys, as well as the use of the DMC, the computers and the control applications for the survey.

In the SGC of the SEI s. a., the proficiency of the training shall be tested through the interviewing of the personnel that is being trained. This way, if doubts remain or there is need of strengthening of the training, it can be provided in the appropriate manner.

The selection of survey takers will be done using the field work practices and tests applied during the entire training time. This will be support by visual aids and exercises prepared by the firm. The objectives will be:

- To know the context of the study
- To meet expectations in the interpretation of the variables used in the form.
- To understand and able to correctly use the form.
- To have a satisfactory use of the digital formats and of the revision of information using these digital formats.
- Reduce the observation errors during the interviews.
- Reach a level of commitment of the personnel to the project through the motivation and the comprehension of its objectives and of its importance to the BM.

A professional team of the project shall be in charge of the training of the personnel. They will have the support of other professionals of the firm, as well as, of the supervisors (trained previously) regarding the preparation in class and in the field. Also expected is the participation of BM.

The Interviewer training will be of 7 full training days, including field practice where each trainee will interview at least two households and two selected individuals.

### **8.1 ASPECTS OF LOGISTICS**

1. Each group will have the following material for the completion of the field work: DMC and rechargeable batteries for the survey taker, a charger, a case to protect the DMC, screen protectors (to avoid the deterioration of the screen, the protector must be change monthly), a laptop, extra batteries, a charger, and USBs and CDs to save the information.
2. In the event where it is impossible to use the equipment (loses or damages), the survey taker shall have at his disposition a reserve of forms which shall then be revised by the supervisor and digitalized subsequently (in the field).
3. Each group shall also receive a census of the blocks which are assigned to them. It will serve them as a planning tool so that they will be able to correctly locate the houses.
4. Each group will be in charge of a metropolitan area, while the rest of the area while be assigned to the group to first succeed in completing their task. For the survey taker's assignment, each zone shall be grouped in routes according to the communication lines and their proximity. This way it shall represent equal divisions of work.
5. The supervisor shall have a donated computer in which he shall enter the necessary information for the coverage control, and also shall install a program of information administration for the information provided by the survey takers using the DMC. The system shall assign to each home a Number of Identification and shall register the appropriate survey taker's code.
6. The work done shall be saved daily into the supervisor computer.
7. The Temporary Union shall provide the technical support in whichever systems is necessary for the duration of the project. This shall be done with the help of system engineers of the SEI s. a.
8. Strategies to deal with public difficulties and other possible issues that have an impact on the project shall be created later.

### **8.2 OPERATIVE SCHEMA**

Which all these aspects and the context of the study taken into account, the completion of the study of each route shall generally be conducted in this manner:

1. The supervisor responsible of the group shall receive a diagram of the study. In their computer shall be the according showing. They shall also receive an urban diagram of the zone with the address of the homes to visit.
2. Once arrived in the neighborhood, the supervisor shall present the project information to the authorities to ask for their collaboration. In the event of a home located in a zone considered to be dangerous, it is needed to ask for information about the security, as well as, confirm the address and the method of access to the area. This is especially important for the organization of actualized data.

3. It is important to present the letter of presentation of the project signed by the World Bank or by an entity assigned by them which provides sufficient support to the field work personnel.
4. The supervisor shall be in charge of organizing their groups work by assigning to each survey taker the blocks that they shall be required to visit. The supervisor shall do this while taking into account the location of each home to facilitate transportation.
5. When arriving to the assigned block, the survey takers shall proceed to recount the homes trying to identify and select the most appropriate informers, according to the research objectives.
6. Once the homes selected, each survey taker shall visit those assigned to him. First, he needs to contact the home owner, his spouse or an adult in charge to allow, during the interview, the completion of part one of the form, concerning general information on the home. This shall be done in accord to the instructive norms of information recollection.
7. Once the selection program has chosen the principal informant, one must proceed by interacting with him through a face-to-face interview, which will allow the completion of the 3 remaining parts of the form. These include the reading-writing tests. If the principal informant is no present, one will have to investigate to the date and hour in which they can be located and organize an interview.
8. For the issue of the main test, the information shall be written down on paper or immediately saved in the DMC so that the system applies the logarithm of selection which shall determine which of the for possible tests to apply
9. Each supervisor shall apply the programmed techniques to verify the quality of the information collected by their group; and shall also routinely control the coverage for which there will be a tracking on the state of the survey in each unit of observation in reference to the number of visits realized and the result of these.
10. The coverage shall be control at the level of the home and of the person that has been selected as principal informant; also shall be examined the coverage of the application of the tests.
11. The cases in which it results impossible to obtain a complete survey, it shall be codified according to the cause (incomplete interview, refusal, absence, impossibility to find the home) so that there may be frequencies generate from null answers and the anticipate indicators in the table of data will explain the situation at a later date.
12. In cases of non-contact and temporary absence, at least three follow-up attempts are required before classifying a case as a non-response.
13. With the information obtain in each unit of observation; the system shall form a single archive per neighborhood.
14. Periodically, the supervisor shall send the available information to Bogotá using a magnetic device. Also, they shall send the actualized materials of the sample and the format of control that has been carefully used in each neighborhood and a full report of these results. They shall have a strict control over the information that is associated with the database that shall be constructed from the sample information and the interview result.
15. This report shall be created by a quality control software which will produce a chart of the results of each interview of those polled (answers will be limited according to the cause), and indicators on the levels of application of the different supervision techniques used.
16. In Bogota, an archive shall be created that guaranties the secure storage and convenience of the data charts.

### 8.3 SUPERVISION OF THE FIELD WORK

Processes such as the study and review of the forms and the direct supervision of the survey takers work in the field are usually realized for all surveys.

The two first activities will be replaced, in the present case, by the exhaustive validation realized in all interviews through the use of the collection programs installed in the DMC of which the advantages have already been discussed.

Liberating the supervisors of such routinely activities such a revision, their work can instead be concentrated on the use of the mechanism which are used for the continued search of quality of the survey. This quality depends of three major conditions which the information collected must contain:

- The **comparability** is guaranteed by the use of standardize criteria for every person involved in the project and by the correct application of the data collection instruction which must be reviewed by the supervisor through his permanent support to his survey takers. An essential condition for this project is the standardization of the anthropometric measurements and cognitive proof.
- The **integrity** of the information implies the control of the coverage of the simple and the revision of the forms to ensure that the information is complete. It is also necessary to complete reports on the results of the interviews to know the range of answers obtained and their impact of the results.
- The **veracity** of the information requires the honesty of the survey takers and of their informants, as well as, a pledge from the field personnel to obtain real data. In this aspect, the quality of the measurement equipment used is essential.

To obtain these optimum conditions, the supervisors can use different techniques of supervision which will be described in the following paragraphs. The forms shall include, in the operative control section, the register corresponding to the technique applied to allow for a proof of the quality of the supervision. The norms of revision shall also be incorporated in the instruction of supervision.

#### 1. Verification of the information:

Since it is not practical to verify in each home the information of the forms, we shall resort to a sample in which the data and coverage of every module shall be verified. This verification will be done with the information obtain by the survey taker. If there is are any important differences, the supervisor must clarify the cause of this difference with the informant and decide if the survey taker will be asked to continue working on the project.

Also, the central level will complete **telephonic** verifications of certain homes to confirm the visit of the survey taker, his behavior, as well as, check the veracity of the information

provided from certain variables.

## **2. Accompaniment**

Other ways to supervise and control the quality consists in accompanying the survey taker in the surveying of certain homes. The accompaniment replaces the verification when it is done directly in the home and allows for the verification of the veracity of the information and the interviewing of every person of the home (this is especially applicable in homes that are found in less accessible zones).

The function of the supervisor during the accompaniment is the participation as an “active observer” in the interview; this can give the supervisor the opportunity to observe the technique of the survey taker and his ability, as well as, correct his faults.

So in that sense, field supervisors will revisit each household if a household refuses or does not begin the interview because of special circumstances (result codes 1 or 2), or if a household where the selected individual is not able to begin the questionnaire – for refusal, for special circumstance, absence, other reasons.

When a household where the individual stops without finishing the individual modules 2-7, or where the individual stops without finishing the assessment exercises, Module 9, the supervisors are going to verify the 10% and 5% through accompaniment.

A verification of an interviewer's visit will be carried out by a revisit to 25% of the households in each interviewer assignment, consisting of a personal revisit to 10% of the finalized cases, a telephone follow-up to the remaining 10% sample of finalized cases, and 5% of accompaniment of the supervisors to interviewers. The households involved in the verification process will be randomly selected.

## **3. Consultant to the personnel**

Part of the supervisor’s functions is to act as a permanent consultant for the survey takers. This activity must always be used to allow a improving of the quality and of the productivity of the work. During the whole duration of the project, the supervisor shall support the survey takers in this aspect and complete the training through additional instructions. This type of activity is especially important in the initial steps of the project.

## **4. Revision of the forms:**

Through the use of this activity, we wish to ensure that the norms about filling in the forms, contained in the instructive attached to the protocol, have been properly taken and that the information is complete and consistent. The revision of the information about the homes will be done through a program installed in every computer which includes the orientated validations to ensure the internal consistency of the information.

#### **8.4 Progress Reporting**

Each week during the survey period, the project manager will submit to the STEP Consortium a data file containing all the entered survey data to date.

<b>I agree with the above,</b>	
<b>PABLO A. ACOSTA</b>	<b>Signature:</b>
<b>CAROLINA MURGUEITIO</b>	<b>Signature:</b>

## 9.0 DATA PROCESSING

### 9.1 Instrument Requirements to Facilitate Data Processing

#### **STEP Standard**

- 1) *A field for recording the respondent Sample Identification Number is required on all survey instruments (i.e., Household Roster, Household Questionnaire, General Booklet, and Exercise Booklet) and any pertinent supplementary material.*
- 2) *Countries must assign a unique booklet ID (serial number) to each prepared assessment instrument (i.e., to each General Booklet and Exercise Booklet).*
- 3) *Fields are required on the Household Questionnaire (or other Interviewer document, such as a case folder) for recording the final completion status of the Household Questionnaire, the General Booklet, and any applicable Exercise Booklet.*

#### **Rationale**

The survey instruments are the primary source of information for creating the international data file. The instruments serve as a vehicle for recording respondent answers to questions as well as administrative information that is needed for case control purposes or that could be used for non-response analysis.

It is essential that allowance has been made on the survey instruments for recording critical information (i.e., Sample Identification Number, final status of the instruments and the sampled case) for linking all survey instruments and related materials for a respondent as well as other administrative and analytical information.

An important consideration is that each country keep track of all assessment booklets, both the used booklets as well as the unused booklets. Countries must assign a unique booklet ID (serial number) to each assessment instrument prepared. This is required to verify that instruments distributed to interviewers have been used for the respondents, or returned and eventually destroyed so that all instruments are accounted for.

A standard set of disposition codes (i.e., final status codes) must be used by all participating countries to ensure that the status of each sampled case is consistently classified.

#### **Report Requirement**

1. Indicate that a unique sample ID will be on all documents pertaining to a sampled case.
2. Outline the procedure for keeping track of the assessment instruments, including a unique booklet ID on all printed assessment documents.
3. Outline the procedure for recording the final completion status of each sampled case.

##### 1. Case Identification

Colombia will ensure the identification of each sampled case by including a unique sample ID on all documents pertaining to the sampled case.

2. Assessment Booklet Identification

Colombia will print a sequential booklet ID on each printed assessment booklet.

3. Case Final Status Code

The interviewer is required to record the final completion status of each case in his/her assignment. The accuracy of the recorded status code will be verified by the field supervisor.

<b>I agree with the above,</b>	
<b>&lt;INSERT TTL NAME&gt;</b>	<b>Signature:</b>
<b>&lt;INSERT PROJECT MANAGER NAME&gt;</b>	<b>Signature:</b>

**9.2 Data Processing: Data Capture, Coding, Scoring, File Creation**

**STEP Standard**

**1) Data Capture**

- a) *For each selected PSU, the data entry must be carried out no later than 5 (five) days after the finalization of the PSU cases.*
- b) *The responses from the Household Questionnaire and the Assessment Scoring Sheets will be manually keyed.*
  - i) *The data capture of the Assessment Scoring Sheets should be carried out by using the Data Entry Program (DEP) provided by the STEP Consortium.*
  - ii) *The Household Questionnaire should be captured using a Consortium approved DEP that incorporates the list of edit checks provided by the Consortium.*
- c) *Each country must key-enter the ‘write-in’ entries from the response category “Other. Please specify” for all questions where this category has been selected.*
  - i) *Furthermore, each country is responsible for coding these ‘write-in’ responses and providing the code set to the STEP Consortium. Each country must ensure that these coded responses are included in the editing of the data.*
- d) *The responses from the Household Questionnaire and the Assessment Scoring Sheets will be 100% verified. In other words, there must be double data entry of these instruments by different key entry operators.*

**2) Coding**

- a) *The Household Questionnaire data and assessment data will be coded as specified by*

*the STEP Consortium.*

*i) The coding of all education variables will be done using the country-specific classification of education. The country will be responsible for providing the mapping of its classification to ISCED 1997 codes. Based on that mapping, the STEP Consortium will be responsible for constructing the internationally comparable education variable.*

*ii) The following codebooks will be used to code occupation and industry information from the Household Questionnaire.*

*01 'International Standard Classification of Occupations (ISCO 08)' will be used to code the occupation variable. The level of disaggregation will be 3-digit.*

*02 'International Standard Industrial Classification of All Economic Activities, Fourth Revision' will be used to code the industry variable. The level of disaggregation will be 3-digit.*

*b) The verification of the coding of Household Questionnaire data and assessment data will be performed according to the specifications of the STEP Consortium.*

*i) Data that has been manually coded will be 100% verified by another coder. The average error rate for manually coded data must not exceed 6%.*

### **3) Scoring**

*a) Each country requires a Chief Scorer who is fluent in English and the country's STEP reporting language(s) and at least one other scorer who is fluent in English and the country's STEP reporting language(s).*

*b) The assessment booklets will be scored according to the scoring rules and procedures provided by the STEP Consortium.*

*c) Each country is required to carry out the quality control procedures for the scoring of the assessment booklets. The quality control procedures will be provided by the STEP Consortium.*

### **4) Data Editing**

*a) Each country will perform an edit of its STEP data file in order to identify and resolve errors in the data. Each country is responsible for ensuring that its final data file submitted to the STEP Consortium is error-free'.*

### **5) Data File Creation**

*a) Each country's STEP data file will be created according to the International Record Layout (IRL) as specified by the STEP Consortium.*

### **Rationale**

The processing of data from the STEP survey must be done using uniform methods to ensure that the captured data is as free of capture errors as possible. As well, the data capture system must be fully tested prior to the commencement of data capture. In addition to a fully-tested data capture system, sound quality control procedures such as 100% verification of the data

capture (i.e., data capture by two different data entry staff) will ensure that the STEP dataset is free of data capture errors.

### **Report Requirement**

1. Data capture and verification plan
2. Plan for coding and verification of data
3. Plan for scoring of the task booklets
4. Description of database creation and record layout
5. Description of the editing system
6. Explain any expected deviations from the STEP Technical Standards.

#### 1) Data Capture

- a) For each selected PSU, the data entry will be carried out in the moment of the interview through PDA.
- b) The responses from the Household Questionnaire and the Assessment Scoring Sheets will be manually keyed.
  - i) The data capture of the Assessment Scoring Sheets will be carried out by using the Data Entry Program (DEP) provided by the STEP Consortium.
  - ii) The Household Questionnaire will be captured through PDA, but with same structure and same checks of the Consortium approved DEP.
- c) With PDA, we are going to record the ‘write-in’ entries from the response category “Other. Please specify” for all questions where this category has been selected.
  - i) We are going to coding these ‘write-in’ responses and providing the code set to the STEP Consortium. We ensure that these coded responses are included in the editing of the data.
- d) For the use of PDA, we are going to have checks directly in the interviews. For that reason we are not going to do double data entry of household questionnaire. Although, for the data entry for the score sheets from booklets we are going to have double data entry by different key entry operators.

#### 2) Coding

- a) The Household Questionnaire data and assessment data will be coded as specified by the STEP Consortium.
  - i) The coding of all education variables will be done using the classification for Colombia of education.
  - ii) The following codebooks will be used to code occupation and industry information from the Household Questionnaire.
    - 01 ‘International Standard Classification of Occupations (ISCO 08)’ will be used to code the occupation variable. The level of disaggregation will be 3-digit.

02 'International Standard Industrial Classification of All Economic Activities, Fourth Revision' will be used to code the industry variable. The level of disaggregation will be 3-digit.

b) A 100% verification of the coding of Household Questionnaire data and assessment data will be made by coordinator and data manager based on the information recorded through the PDA.

### 3) Scoring

No data will be manually coded.

#### **Recruiting and Training Scorers**

The scorer's team will be coordinated by Olga Romero and Carolina Murgueitio, both bilingual. The rest of the team will be by three people who will be recruited and trained by the chief scorers. The team will work in the same location (Econometría's offices). The scorers will be students of last year of university or recent professionals. At least six scorers will be recruited and three will be selected after a test of scoring abilities. Also the STEP test will be provided and the best scores will be selected. During the training sessions scorer's guides provided by ETS will be provided and in the training there will be also a participation of Angelica Becerra who will receive ETS training in Washington DC. The scoring training materials will be used during training sessions as manuals, examples and exercises.

#### **Ensuring Inter-rater Agreement**

Each scorer will have a standard scoring sheets constructed under ETS guidelines and protocol. 1/3 of the assessment booklets will be re scored by a second scorer. A monitoring of the process will be permanent in order to identify possible mistakes and how to avoid them. Although, it is important to clarify that the chief scorers only are going to supervise, that is means that they are not going to score booklets, only they are going to score the English booklets send by ETS, not booklets from field work in Comlombia.

#### **Documenting the Scoring Process**

The scoring process will be documented under ETS guidelines and standardized report for all countries. There will be a digital notebook of the scoring process where the atypical cases will be pointed and obstacles and ways to overcome them.

### 4) Data Editing

a) We will perform an edit of its STEP data file in order to identify and resolve errors in the data. And we ensure that its final data file submitted to the STEP Consortium is error-free'.

### 5) Data File Creation

a) Although the data collection is going to do in PDA we are going to follow the same structure of data collection give by the World Bank. So, the data file will be created according to the International Record Layout (IRL) as specified by the STEP

Consortium.

## **9.2.1 Data Editing System**

There are two major types of data to be edited for the STEP survey. These are the data resulting from administration of the Household Questionnaire and the data from the literacy assessment instrument. A computerized editing system will be provided to each Survey firm/agency to facilitate the cleaning of the data.

### **9.2.1.1 Editing Household Questionnaire Data**

The edit of the STEP international data file will include the following minimum checks for the Household Questionnaire data. For each of these edits, if errors are discovered they will be resolved, i.e., the original erroneous value will be replaced with a corrected value.

#### 1) ID check

The record identification numbers on the STEP data file will be checked for uniqueness and integrity to ensure that there is only one record per respondent on the file, and to ensure that the record identification number is unique and in the specified format.

#### 2) Range checks

A range check will be carried out for all those variables that can only take on specific values.

#### 3) Logic checks, i.e., question flows

The STEP data file will be edited to check the flow of respondents through the various sections of the Household Questionnaire. The objective of this edit is to ensure that the responses for respondents who should have skipped a given set of questions have been properly coded as a 'valid skip', and that there are appropriately coded responses for respondents who should have completed a given set of questions.

#### 4) Consistency checks

An edit of the STEP data file will be performed to identify inconsistencies that may have arisen as a result of response errors, coding errors, and data capture errors.

#### 5) Outlier check

An edit will be performed to identify possible outliers, i.e., extreme quantitative data values. All identified outliers will be reviewed for legitimacy and to assess the potential effect on the survey estimates.

Imputation methods will not be used to treat missing Household Questionnaire data, i.e., item non-response and complete non-response.

### **9.2.1.2 Editing Assessment Data**

The edit of the STEP international data file will include the following minimum checks for the

assessment data. For each of these edits, if errors are discovered they will be resolved, i.e., the original erroneous value will be replaced with a corrected value.

1) ID check

The editing of the assessment instrument will consist of confirming that the correct booklet was administered to each respondent and to confirm that the completed assessment booklet was received and labeled with the appropriate case information.

The record identification numbers on the data file will be checked for uniqueness and integrity to ensure that there is only one record per respondent on the file, and to ensure that the record identification number is unique and in the specified format.

2) Range checks

A range check will be carried out for all score variables that can only take on specific values.

Scored literacy responses will be checked to ensure that they conform to the specified structure of the IRL.

Data Editing Strategy

The edit of the STEP international data file will include the following minimum checks for the Household Questionnaire data. For each of these edits, if errors are discovered the PDA is not going to pass to the next question until it will be resolved, i.e., the original erroneous value will be replaced with a corrected value.

1) ID check

The record identification numbers on the STEP data file will be checked for uniqueness and integrity to ensure that there is only one record per respondent on the file, and to ensure that the record identification number is unique and in the specified format.

The editing of the assessment instrument will consist of confirming that the correct booklet was administered to each respondent and to confirm that the completed assessment booklet was received and labeled with the appropriate case information.

The record identification numbers on the data file will be checked for uniqueness and integrity to ensure that there is only one record per respondent on the file, and to ensure that the record identification number is unique and in the specified format.

2) Range checks

A range check will be carried out for all those variables that can only take on specific

values.

A range check will be carried out for all score variables that can only take on specific values.

Scored literacy responses will be checked to ensure that they conform to the specified structure of the IRL.

3) Logic checks, i.e., question flows

The STEP data file will be edited to check the flow of respondents through the various sections of the Household Questionnaire. The objective of this edit is to ensure that the responses for respondents who should have skipped a given set of questions have been properly coded as a 'valid skip', and that there are appropriately coded responses for respondents who should have completed a given set of questions.

4) Consistency checks

An edit of the STEP data file will be performed to identify inconsistencies that may have arisen as a result of response errors, coding errors, and data capture errors.

5) Outlier check

An edit will be performed to identify possible outliers, i.e., extreme quantitative data values. All identified outliers will be reviewed for legitimacy and to assess the potential effect on the survey estimates.

Imputation methods will not be used to treat missing Household Questionnaire data, i.e., item non-response and complete non-response.

**I agree with the above,**

**PABLO A. ACOSTA**

**Signature:**

**CAROLINA MURQUEITIO**

**Signature:**

## 10.0 WEIGHTING

### **STEP Standard**

*The weighting of each country's clean data file will be carried out by the STEP Consortium. The weights to be added to each countries clean data file include the theoretical or sample design weight, a non-response adjusted weight, a benchmark weight, and a set of jackknife weights.*

*Each country will be required to provide a recent dataset containing the most recent target population counts of the key benchmark variables age, gender, education, and urban-rural indicator. This dataset may be derived from the most recent census file or similar national file, and it will be used to create the benchmark weights.*

### **Rationale**

In order to generalize the sample findings to the survey population, the estimate of a population characteristic as well as the estimate of the associated sampling error should be based on the appropriate survey weights. Each record on the STEP data file should include a single overall weight for use in producing survey estimates. It is strongly advised to include weight component factors, e.g., non-response adjustment factor, that contribute to the weight calculations. Replicate weights will also be required for use in estimating the standard error of the survey estimates.

As well, benchmarking the sample weights to agree with external population counts involves making adjustments to the sampling weights, so that when the resulting weights are summed across a particular population subgroup, the resulting total agrees with an external known population count (e.g. census counts) of the size of that subgroup. Benchmarking increases the precision of the survey estimates and also reduces the bias due to problems of survey coverage such as non-response, deficiencies in the sampling frame or data collection operations, etc..

### **Report Requirement**

1. Description of weighting procedures including a list of the weights which will comprise part of the final survey data file
  - a) Include a description of the post-stratification strategy.
    - i) Specify the variables to be used for 'benchmarking'.
    - ii) Specify the source of the file to be used to create benchmark weights.

## 10.1 Weighting Procedures

The weighting of each country's clean data file will be carried out by the STEP Consortium. The weighting of the respondent records will be consistent with the Colombia probability sample design.

Survey weights will be calculated from the clean sample file (i.e., the file resulting from the editing process). The survey weights will be appended to each respondent record on the clean survey file. The following weights will comprise part of each respondent record:

- 1) Theoretical or sample design weight - the inverse of the probability of selection at the sample selection stage.
- 2) Non-response adjusted sample weight - based on the sample design weight and adjusted for non-response.
- 3) Benchmark weight - the weight resulting from the adjustment of the survey results to known population totals.
- 4) Jackknife replicate weights - there will be 30 of these weights which are used to calculate the standard error of the survey estimates

### 10.1.1 Benchmarking Variables

The variables to be used for benchmarking are age, gender, and region.

### 10.1.2 Source of Benchmark Variables

Colombia will provide the most recent counts of the benchmark variables to the STEP Consortium. These counts are the most recent known population totals for the variables, age, gender, and education. The counts will be from a reliable current source of data such as a recent Census or other recent national survey. If the most current known totals of the benchmark variables are from the sample frame used to select the STEP sample then the benchmark weights will be created using the data from the sample frame.

The benchmark weights will be created based on the known population totals for age, gender, and education using data from the [National department of statistics in Colombia – DANE with projections based on census of 2005](#).

<b>I agree with the above,</b>	
<b>PABLO A. ACOSTA</b>	<b>Signature:</b>
<b>CAROLINA MURGUEITIO</b>	<b>Signature:</b>

## 11.0 CONFIDENTIALITY

### **STEP Standard**

*Each participating country will advise the STEP Consortium of its confidentiality rules regarding collection and handling of respondent data.*

*Each country will submit the international data file to the World Bank STEP team (regional team and core team). Data release beyond these World Bank teams may not be done until 6 months after the submission of the final data files.*

### **Rationale**

The STEP Consortium needs to ensure that each country is permitted to share the collected data with the Consortium. The Consortium must therefore be informed of the need to ensure that each country's rules on confidentiality regarding the handling of respondent information are respected.

### **Report Requirement**

- 1) Outline the country's data confidentiality requirements.
- 2) Outline the steps to ensure data confidentiality.

The SGC de SEI s. a. guarantees the property, the confidentiality, and the security of the information of the client according to the norms ISO (7.5.4 and 7.5.5.). These norms are applicable to the information supplied by the project (for example; the sample, the methodology, the forms and the manuals), as well as, for the information obtain through interviews. To allow this, there is a process of maintenance and support system which are in charge of giving a direct service to the user in the information technology area. Also, there are various protective dispositions installed against intrusions or viral software (protective network firewall, antivirus, and firewall for the client's equipment) which impedes the access of none authorized persons and the damage of various archives of the server.

The information that is product of the census and surveys will be administered through a server assigned specifically for this. Automatically and of a permanent form, copies will be made of all archives which will be sent to a back-up server that is outside the firm's network.

The respect of information property is a policy of the SEI s. a. which is materialized in different legal commitments through contracts with the clients and the financial backers of the firm. It is also materialized in labor contracts and service provision contracts signed with the part-time personnel used in the completion of projects. These contracts have specific clauses regarding the respect and the obligation of confidentiality of the information obtained during interviews. Failure to follow this procedure will lead to the termination of the contract.

The team will submit the international data file to the World Bank STEP team (regional team and core team), which will in turn share the data with the Government of Colombia as requested. Data release beyond these World Bank teams may not be done until 6 months after the submission of the final data files.

<b>I agree with the above,</b>	
<b>PABLO A. ACOSTA</b>	<b>Signature:</b>
<b>CAROLINA MURGUEITIO</b>	<b>Signature:</b>

## 12.0 QUALITY ASSURANCE

### **STEP Standard**

*Each country will outline the procedures put in place to assure the overall quality of the STEP data.*

### **Rationale**

The success of the STEP depends on the steps taken to assure that the study is designed and implemented according to common goals and sound methodology and operational practices so that the survey results are reliable. The quality assurance procedures throughout the survey process will help to ensure that the sources of survey variability may be kept to a minimum and the comparison of survey results across participating countries is both feasible and credible.

### **Report Requirement**

1. Outline the country's quality assurance plan.
  - a) List all the procedures/activities that the country will implement in order to assure the quality of the survey results.

The collection process shall be realized using the “Sistema de Gestión de Calidad de SEI s. a.” (Quality Control System) which is certified by the ICONTEC (No. SC 3764-1); which covers 24 processes, including those of the provision of services which are directly associated with the collecting and processing of the information; and which also covers the other processes which meets the implicit requirements defined by the firm as necessary for the provision of a quality service. Copies of the certificate are included in the annex X. Each process relies on quality monitoring indicators which allow the measuring of the performance of the project and the corrective or preventive actions to use if necessary.

As a result of the quality control, there will be feedback reunions with the data collection personnel as a preventive measure to avoid further errors. It will be required that the survey taker returns to the field and corrects the errors that cannot be rectified in office. The project management director shall exercise a quality control of the supervision, ensuring the consistency of the information and the organization of the group; verifying the coverage of the survey and the completion of the time tables planned in the chronogram.

In the next table we present the quality control plan.

	<b>QUALITY PLAN</b>		COD:SGC-08
	APROVED BY: QUALITY COMMITTEE	DATE: May 24th, 2010	VERSION: 02
	PROJECT: STEP Skills Measurement		CODE OF THE PROJECT: ARD - 397

DATE OF PRESENTATION: Day: |\_|\_|\_|\_| Month: |\_|\_|\_|\_| Year: |\_|\_|\_|\_|\_|\_|\_|\_|\_|\_|

<b>RESPONSIBLES:</b>			
DIRECTION:	<u>YAMILE PALACIO</u>	TECHNICAL ADVISOR:	<u>YEZID BOTIVA</u>
COORDINATOR OF THE PROJECT:	<u>ANTONIO REY</u>	DATA CLEANING:	<u>SONIA ROA</u>
SURVEY DESIGN AND SELECTION:	<u>HANNER SANCHEZ - ANGELICA BECERRA</u>	DATA PROCESS:	<u>ALEJANDRO GARNICA - GABRIEL RAMIREZ</u>
DATA ENTRY:	<u></u>	ADMINISTRATIVE ISSUES:	<u>PATRICIA ARENAS - ENITH MALDONADO</u>

Process	Description	Products	Indicators
PS1 PLANIFICATION OF THE PRODUCT	Organize y coordinate the resources in order to guarantee all quality and oportunity requirements.	Quality plan	
PS 2 TECHNICAL DEVELOPMENT OF THE PROPOSAL	Develop the methodology: adopt the basic definitions and concepts to define the manuals and questionnaires.		
PS 3 WORK OPERATIVE AND LOGISTIC OF THE PROPOSAL	Create zones and routes of work, assign resources, planning work, define the system of supervision and following the field work.		
PS 5 SURVEY DESIGN AND SELECTION	Stablish the parameters of the survey design of the project.		
PS 6 OPERATIVE OF FIELD WORK	Get the information require by the cliente through the use of digital questionnaires	Parcial reports	PS6-PC1 Response rate
PS 6-DMC DMC's PROGRAMMING	Develop capture programs with CSProX to make the collection of the information.		
PS 7 DATA CAPTURE	Incorporate in digital medium the information require by the client, through automatic process.		
PS 8 DATA PROCESS	Make the project information outputs, based on the analysis plan acorded with th client.		
PS 9 SUBMMIT OF THE RESULTS	Elaboration of final reportand backup preparation.	Final report Clean data base	PS9-PC4 Execution time accomplished
AT 2 TRAINING AND HIRING OF THE TEMPORARY STAFF (SUPERVISORS AND INTERVIEWERS)	With training sessions, guarantee the best staff with right skills to develop the activities of data collection, data editing, supervision and data capture.	Training agenda List of asistency to training and control of test and quizzes.	AT2-1 Skills of temporary staff AT2-2 Training quality
AT 3 SUPERVISION OF FIELD WORK	Verify, validate and follow the field work.		AT3-1 Percentage of verification AT3-2 Percentage of revisit AT3-3 Percentage of accompaniment
AT 4 INTERNAL AUDITORY AND COMPLEMENTARY ACTIVITIES	Evaluate the quality control system to define if the requirements of the system are appropriate		

<b>I agree with the above,</b>	
<b>PABLO A. ACOSTA</b>	<b>Signature:</b>
<b>CAROLINA MURGUEITIO</b>	<b>Signature:</b>



<b>I agree with the above,</b>	
<b>PABLO A. ACOSTA</b>	<b>Signature:</b>
<b>CAROLINA MURGUEITIO</b>	<b>Signature:</b>

## 14.0 BUDGET

### STEP Standard

*Each country will provide a budget for their STEP implementation activities.*

### Rationale

A budget of expenditures is a key planning tool.

### Report Requirement

1. Provide the budget for the major activities for the STEP implementation.

The estimated expenditures for the STEP in Colombia is summarized in the following table:

Budget Estimate – STEP - EDITED -	
Item	Estimated Expenditures
1) Project Team Salaries, main Survey Data Processing	
2) Travel (e.g., international meetings)	
3) Pilot Survey Preparation, collection, data processing and evaluation	
4) Main Survey Preparation, main Survey Evaluation	
5) Main Survey Collection	
TOTAL Estimated Expenditures	

I agree with the above,

PABLO A. ACOSTA

Signature:

CAROLINA MURGUEITIO

Signature: