

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

CHINA-Yunnan

23 December 2011



<Taozhen Bai>
STEP Survey

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 ‘<’ and ‘>’ characters (e.g. <Country>)</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>FIRMS/AGENCIES SHOULD ADJUST THESE SECTIONS BASED ON THE COUNTRY’S EXPERIENCE.</p>
Orange Boxes	<p>TASK TEAM LEADERS and PROJECT MANAGERS – 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

1.0	Introduction.....	1
2.0	Project Team	2
2.2	Project Structure.....	6
3.0	Survey Objectives.....	9
3.1	Major analytic objectives.....	9
3.2	Country-specific objectives	10
4.0	Sample Design Factors	11
4.1	Target Population.....	11
4.2	Method of Data Collection.....	13
4.3	Response Rate.....	14
4.4	Sample Frame	15
4.5	Sample size – Partial Assessment.....	17
5.0	Sample design	20
5.1	Sample Allocation and Selection	23
6.0	Literacy Assessment.....	24
7.0	Household Questionnaire	28
8.0	Data Collection	31
8.1	Survey promotion Strategy	34
8.1.1	Community Communicator System.....	34
8.1.2	Advance Survey Information.....	34
8.2	Contact Strategy.....	35
8.3	Response Rate Strategy to Minimize Non-response.....	35
8.4	Interviewer Hiring Plan.....	36
8.4.1	Number of Interviewers	37
8.4.2	Method of Payment.....	37
8.5	Interviewer Training Plan	37
8.6	Interviewer Supervision Procedures	38
8.6.1	Number of Supervisors	38
8.6.2	Supervisor Responsibilities.....	39
8.7	Progress Reporting.....	40
9.0	Data Processing	41
9.1	Instrument Requirements to Facilitate Data Processing.....	41
9.2	Data Processing: Data Capture, Coding, Scoring, File Creation	43
9.2.1	Data Capture	45
9.2.2	Data Capture System Test.....	46
9.2.3	Data Capture Verification	46
9.2.4	Coding.....	46
9.2.5	Scoring Task Booklets	47
9.2.6	Creation of International Data File	48
9.2.7	Data Editing System	49
10.0	Weighting.....	51
10.1	Weighting Procedures.....	52
10.1.1	Benchmarking Variables.....	52
10.1.2	Source of Benchmark Variables	52

11.0	Confidentiality	53
12.0	Quality Assurance	55
13.0	Schedule	58
14.0	Budget	60

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.

Yunnan is a multi-national province located in south-west of China and is adjacent with Myanmar, Laos and Vietnam. Even in China, Yunnan is a developing province with relatively lower social and economic development. Kunming, the capital city of Yunnan, is the sole super-city in this province, which owns more than 6.4 million resident population and about 3.28 million in its main urban areas.

As far as we have known by now, there are no any formal reports on the adult literacy assessment in Yunnan. According to the statistical data from local government, there are about 250 thousand illiterate persons in Kunming in 2010, which is accounting for about 4% of total population of Kunming. Here the "illiterate" is classified based on the educational level, that is, person who has not received any formal education is regarded as a "illiterate" person. This can be regarded as a reflection of the literacy situation in Kunming.

Participating in STEP survey will provide a light for us to learn more about the literacy situation in the urban areas of Yunnan Province, as well as the skills conditions of the working-age population in the urban areas of Yunnan, to find the skills gap lying between the owned and the requirements. This insight will help us to develop more efficient policy interventions to improve the skills level and employment of working-age population in Yunnan.

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.

Report Requirement

1. Identify the leading survey institute.

Modern Statistical Application Research Centre of Yunnan University of Finance and Economics is the leading institute for STEP survey in Yunnan(China) and is responsible for the design, management and implementation of the survey.

The Centre is a key institute of YNUFE, which was established in 2003. Now it has 12 full and part-time researchers and 5 of them possess doctor's degree in statistics or mathematical statistics. It relies on the discipline superiority of Statistics and Mathematics College to carry out broad cooperation with governments and enterprises. Recent years, we have carried out many projects with Yunnan Statistics Bureau, Yunnan Science and Technology Bureau, Yunnan Federation of Trade Unions, Yunnan Provincial Industry and Information Technology Commission, etc.

Statistical analysis, social surveys and marketing researches are the superiorities of the Centre.

With the close cooperation with field work institutes, we have successfully conducted several provincial surveys such as Human Resource Survey of Private Enterprises in Yunnan, Sports Industry Survey of Yunnan, Employee's Hospitalization Costs of Yunnan, WBN Usage Survey for Yunnan Telecom, etc. The sample sizes of these surveys are all more than 2000. In all these surveys, we are responsible for survey management, sampling design, implementation direction, data processing and analysis, including sample weighting and estimation.

2. Provide an overview of the qualifications and expertise of the key project team members.
 - a) Indicate whether or not the key team members will work full-time on STEP.
 - b) Include names and types of surveys conducted.
 - c) Indicate the size(s) of survey(s) (i.e., sample size) undertaken.

We have established a project team to carry out the STEP in <Yunnan(China)>. This section identifies the project team members, outlines their qualifications and experience, and summarizes the responsibilities of each team member. The main team members, including Project manager, Senior Survey Methodologist, Data Collection Manager and Data Processing Manager, will be working full-time on this project.

1. National Project Manager

Mr. <Bai Taozhen> has been appointed as the National Project Manager. Mr. <Bai Taozhen> graduated from <Yunnan University> with <Master degree in Mathematical Statistics>. He is currently <the vice director of Modern Statistical Application Research Centre of YNUFE>. Mr. Bai has working as a researcher in marketing survey companies for more than 10 years. And now he holds technical consultant for the two leading consultant companies in Yunnan. He has also successfully managed and directed a variety of large-scale surveys including for example the <Human Resource Survey of Private Enterprises in Yunnan>, the <Sports Industry Survey of Yunnan>, and the <WBN Usage Survey for Yunnan Telecom>. These surveys ranged in size from a sample of 2450 for the < WBN Usage Survey for Yunnan Telecom> to a sample of 4000 for the <Human Resource Survey of Private Enterprises in Yunnan >.

<Mr. Bai> will be responsible for the whole management of this project, including designing the survey plan, assigning the tasks, training other managers, supervising data collection and processing and communicating with task team from World Bank. He will be working full-time on this project except the class time he needs to spend as a lecturer in YNUFE. Supporting <Mr. Bai> is an experienced group of individuals having expertise in 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, and data analysis. The team will consist of <Mr. Li Xingxu, Senior Survey Methodologist>, <Ms. Xia Weiping, Data Collection Manager> and <Mr. Chen Fei, Data Processing Manager>. These individuals have worked directly with previous large-scale surveys as well as with the design and development of paper-and-pencil surveys. While each individual is assigned an area of primary responsibility, they will work so that each has a complete understanding of the overall project objectives and requirements. Within these areas of primary responsibility an individual will be required to monitor and oversee progress on all

relevant tasks, and to report regularly to < Mr. Bai >.

In addition, <Mr. Bai Taozhen> has been engaged to be directly responsible for the translation and adaptation of all STEP materials such as the Background Questionnaire, the Assessment Items, Interviewer Manuals, etc.

2. Senior Survey Methodologist

<Mr. Li Xingxu> is a Senior Survey Consultant with <Yunnan Statistics Bureau and other governmental institutes>. He received in 2003 a Doctor of Science (Major: Mathematical Statistics) from < Yunnan University>. Before this, he had been working in Statistics Bureau for more than 15 years. And now, he has been working as a professor in Yunnan University of Finance and Economics.

<Mr. Li > has approximately twenty-five years experiences in the design and implementation of surveys, including such activities as sample design, questionnaire design, interviewer hiring and training, data processing, weighting, estimation, and analysis of survey results. He also has extensive experience in the development and delivery of training programs pertaining to survey design and implementation, and data processing. When he worked in Statistics Bureau, he had operated the population sampling survey in size more than 60000. Recent years, he worked as senior survey consultant or project manager in such surveys as <Employee's Hospitalization Costs of Yunnan>,<Human Resource Survey of Private Enterprises in Yunnan>, <Sports Industry Survey of Yunnan>.

<Mr. Li> is the methodology representative on the survey team. He is responsible for all aspects of the survey design, weighting, and estimation. In addition, He provides advice to the other team members regarding the methodological considerations for the data collection and data processing. He will be working full-time on this project except the time he need to spent as a professor of YNUFE.

3. Data Collection Manager

<Ms. Xia Weiping> is a data collection manager with < Kunming Feeling Marketing Research Co., Ltd.>. She graduated in 2010 from the Yunnan University with a <Bachelor> in <Economics>. But before this she has been working in the field of survey implementation for more than 8 years and has continued to add to her knowledge by taking courses in Project Management, Survey Methodology, etc.. <Ms. Xia> has been the data collection manager for a variety of large-scale surveys and now she is the invited lecturer of MSARC.

As the data collection manager in our team, <Ms. Xia Weiping> will be responsible for the implementation of the survey. This includes responsibility for the hiring, training and supervision of Interviewers and Interviewer Supervisors, preparation of the data collection schedule, monitoring data collection progress and costs, etc.. She will be working full-time on this project during the period of data collection.

4. Data Processing Manager

<Mr. Chen Fei> will be working as the Data Processing Manager in our team. He received in 2008 the Doctor of Statistics from the Chinese University of Hong Kong. He is good at statistical theory, data analysis and computer programming and has accumulated several years of experience in survey development and data processing.

<Mr. Chen Fei> will be responsible for the data entry management, data cleaning and preparing the final database. He will be working full-time on this project during the period of data collection and processing.

5. Language Specialist

<Ms. He Yu> has been invited as the Language Specialist in our team. She is the Master in English and Statistics in Yunnan University of Finance and Economics. She has been concentrated in English study for many years and is fluent in English and Chinese. Especially she has an extraordinary understanding about the difference in English and Chinese culture.

<Ms. He> will give her efficient help to the work of translating and adaptation of all STEP materials. She will be responsible for examining and verifying the initial result translated under the direction of Mr. Bai. In addition, she will be responsible for administrating the scoring work, which includes training the scorers, allocating the scoring work and entering the assessment sheets.

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

I agree with the above,	
XIAOYAN LIANG	Signature:
<TAOZHEN BAI>	Signature:

2.2 Project Structure

Report Requirement

1. Identify the sponsoring organization

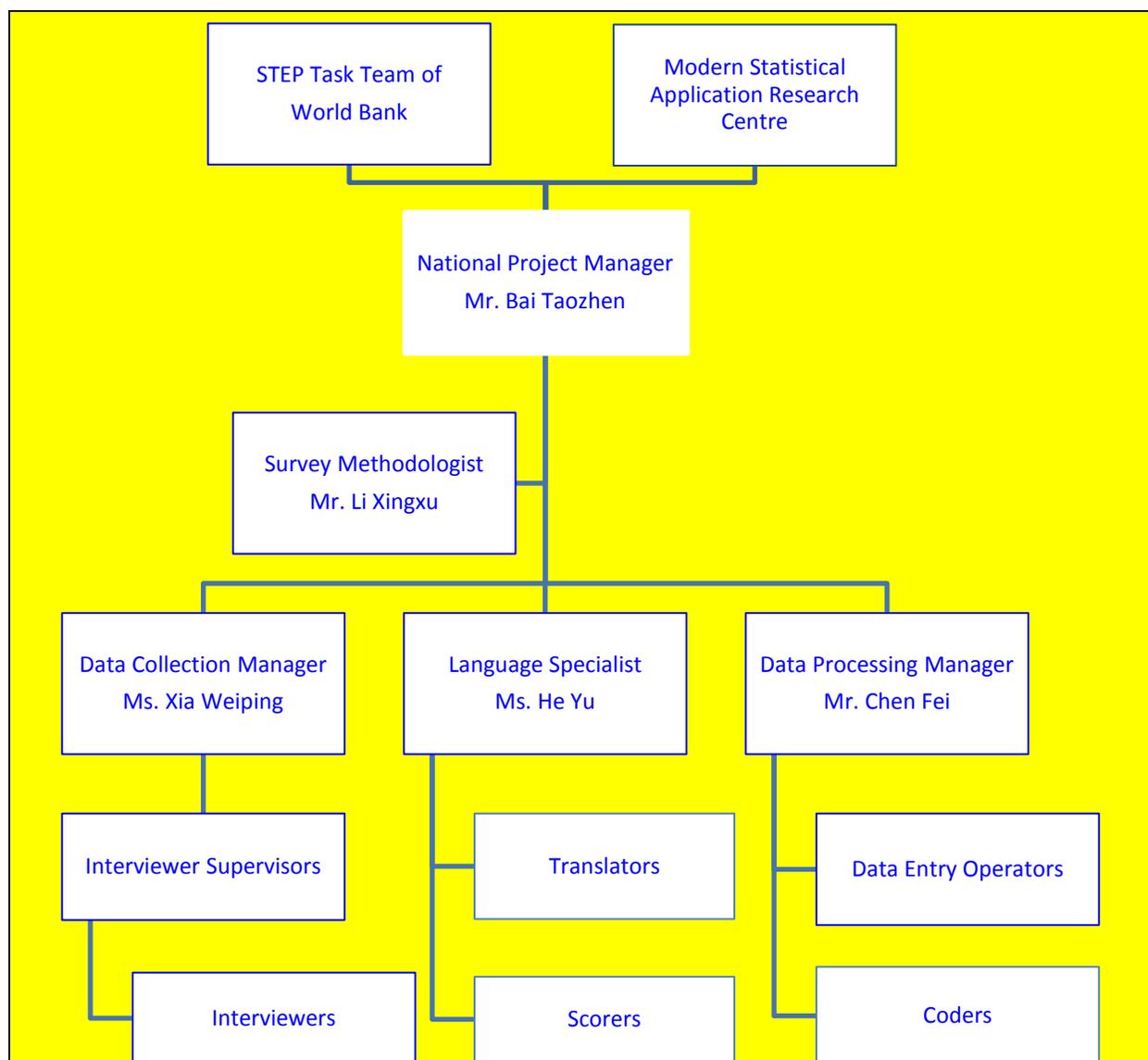
STEP is jointly sponsored by the <World Bank China Team and Modern Statistical Application Research Centre>. The World Bank China Team will provide supervision and support to the project and the Modern Statistical Application Research Centre will be responsible for the organization and implementation of the project. A Steering Committee has been established to provide direction to the project. The committee is comprised of experts from the Yunnan University of Finance and Economics, and experts from the Yunnan Statistical Bureau.

2. Provide an overview of the management reporting structure. (Include the project organization chart)

The STEP will be carried out by the project team headed by <Mr. Bai Taozhen> who will keep close communication with the task team from the World Bank. <Mr. Bai Taozhen> will also receive the direction from Mr. Li Xingxu, the senior survey methodologist of Yunnan project team.

The organization chart for STEP is as follows:

Figure 1: <Yunnan(China)> STEP Project Organization Chart



<Mr. Bai Taozhen>, the National Project Manager, will have responsibility for:

- overall direction and management of the project,
- liaising with the members of the STEP Consortium,
- the development and validation of all instrumentation for STEP,
- developing the project plan and ensuring its implementation,
- ensuring that the project requirements are delivered on time and within budget

<Mr. Li Xingxu>, the STEP methodologist, will be responsible for

- sampling design and implementation,
- weighting and estimation,
- directing the project implementation ,
- liaising with the outside governmental department

<Ms. Xia Weiping>, the STEP Data Collection Manager, will have responsibility for:

- hiring, training, monitoring and control of data collection staff, such as interviewers and interviewer-supervisors,
- development of interviewer materials
- development and implementation of data collection procedures
- control of both item non-response and complete non-response

<Mr. Chen Fei>, the STEP Data Processing Manager, is responsible for:

- data capture,
- coding,
- editing,
- file creation
- tabulation of survey results and regular data analysis.

<Ms. He Yu>, the STEP Language Specialist, will be responsible for:

- translating and adapting of all STEP materials
- administrating the scoring works

I agree with the above,	
XIAOYAN LIANG	Signature:
<TAOZHEN BAL>	Signature:

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

1. Provide a list of the 'country-specific' objectives

For the STEP survey carried out in Kunming, Yunnan Province, there are no additional country-specific objectives with it.

2. Provide any relevant background and supporting rationale for the 'country-specific' objectives

I agree with the above,	
XIAOYAN LIANG	Signature:
<TAOZHEN BAI>	Signature:

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:

- 1) Persons living outside the country at the time of data collection, e.g., students at foreign universities.
- 2) 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.

China’s Target Population Description

The STEP target population of Yunnan (China) project comprises all non-institutionalized persons aged 15 to 64 years (inclusive) living in the private dwellings in the urban areas of Kunming at the time of data collection.

Hereafter the urban areas of Kunming refer to the urban scope of the four administrative districts, which are Wuhua, Panlong, Guandu and Xishan. The exact geographic border of this scope is hard to be declared because of its irregularity. But the scope can be finely defined by the administrative units lying in this scope. In fact total 38 street offices with 299 resident committees construct the urban area of Kunming and it is the same manner used in Census.

The definitions of other terms, including “private dwelling unit”, “household member”, “institutionalized” etc., follow the rules specified by Technical Standards, which are given as above. Specifically, the usual place of residence means a place where a person spends most of their daily rest for at least 9 months over the past 12 months.

The exclusions from the target population include:

- 1) People living in rural areas;
- 2) People in institutions, which are specified as above
- 3) People living outside the scope for the duration of data collection
- 4) Members of the population who are unable to complete the STEP assessment due to a physical or mental conditions
- 5) Foreign diplomats and non-nationals working for international organizations

There is only one language, Chinese, used in Kunming. So the literacy assessment conducted

in this survey will be administered in Chinese.

2. Specify any special additions to the target population.

a) Include the relevant background and rationale for additions.

There are no special additions to the STEP survey in <Yunnan(China)>.

3. Provide counts (or estimated counts) of the target population by sample design variables, e.g, by strata, by PSU.

According to the up-to-date census conducted in 2010, there are about 1,067,256 households with 2,539,034 individuals living in the urban areas of Kunming. And 76.13% of these urban populations belong to working-age population aged 15-64. This means that there are about 1,933,000 individuals in the target population.

In the Census, there are total 299 CEAs in the urban areas of Kunming. These CEAs completely cover the target population. But they vary in size from a minimum of 196 household dwellings to a maximum of 20968. The median number of household dwellings in a CEA is 2696 and the average is 3570.

The following table shows the information about the CEAs.

HHLD number	Number of CEAs	Percentage	Total HHLDs included	Percentage
>=10000	11	3.7%	154114	14.4%
7000-9999	22	7.4%	184330	17.3%
4000-6999	61	20.4%	322083	30.2%
1000-3999	165	55.2%	382699	35.9%
700-999	14	4.7%	11523	1.1%
400-699	20	6.7%	10614	1.0%
100-399	6	2.0%	1893	0.2%
Total	299	100.0%	1067256	100.0%

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

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).

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, Background Questionnaire, and Literacy Assessment.
2. Explain any expected deviations from the STEP Technical Standards.

Each component of the STEP Survey will be carried out by a personal visit using a Paper And Pencil Interview (PAPI) method.

Personal visit using PAPI method will be the only and basic method used in STEP survey in Yunnan, China.

The household questionnaire (household module and individual module) will be finished by face-to-face visit, i.e., the interviewers read the questions to the respondents and record the answers in the questionnaires. Considering the actual situation, it is allowed to finish the whole questionnaire by several visits.

All components of the literacy assessment will be administered in the same visit (i.e., General Booklet and applicable Exercise Booklet). And this module will be finished by self-writing method, i.e., the respondents will answer the questions by themselves.

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 <Yunnan (China)> is expected to be <70-80%>. This overall response rate takes into account an expected assessment CORE exercise 'Pass' rate of <100%> because STEP survey in Yunnan does not involve in Exercise Booklet.

This expected response rate and expected Core pass rate are based on the following reasons:

1. A community communicator will be employed for contacting the respondent and helping the interviewer when refused. Because the communicator is familiar with the respondents, they can provide efficient help in the implementation.
2. A fully-trained interviewer will be employed to conduct the survey. The interviewers will be instructed in techniques to obtain the cooperation of individuals who initially refuse or are reluctant to participate in the survey.
3. We will emphasize that this is a multi-country survey and more than ten thousand people are involved in it and the respondent will be the representative of China. We believe that national pride will be an important stimulus to make the respondents happily accept the interview.
4. If approved and funded, the respondents will firstly be given a small gift (e.g., a bottle of shampoo) as a token of appreciation for their participation in this task. It should help obtain a higher screener response rate, and provide an upfront incentive/inducement to participate for households with eligible study participants. In addition, respondents who complete the survey will receive a material incentive for their cooperation.

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 and agreement of the suitability of the sample frame, 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, the firm will carry out a door-to-door listing of the

households in each selected PSU.

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 type of the sample frame will be area frame correspondence to administrative unit, which comes from the National Population Census just conducted in November, 2010. The frame is a complete list of first stage sampling units in the urban areas of Kunming, which includes the PSU ID # and household dwelling numbers of each PSU.

For the first-stage sampling, the sampling unit (i.e. PSU) is the CEA Block. Just as mentioned above, a CEA in Kunming may be too large for the sampling. So we divide a CEA into different numbers of Block. A Block consists of about 180 to 220 household dwellings. This

division will be convenient for the sampling and implantation. Blocks cover all the territory of survey scope and are stratified by CEAs.

The frame counts by PSU are include in the accompanying EXCEL file, which includes CEA ID, Block ID and PSU ID # and household dwelling numbers for each PSU of total 5,564 PSUs in the urban areas of Kunming. Note that the households here only include family households and the collective households have been excluded out.

Because the Census was conducted just in Nov. last year and population mobility is not so frequent, we do think that the sample frame based on the 2010 census is up-to-date and provides complete and accurate coverage of the target population with the exception of the exclusions previously mentioned. Of course the PSU counts of dwellings will be revised for the selected PSUs at the time of implementation.

If the Census listing is not available, we will carry out a door-to-door listing of the household in each selected PSU. This listing work will be done by the supervisor responsible for this PSU with the help of the community communicator in this PSU. The contents to be listed will follow the listing table designed by STEP Consortium, which will include address of dwelling, type of the dwelling, name of the head of the household, etc.

4.5 Sample size – Partial Assessment

STEP Standard

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

- *A minimum of 2,000 complete STEP interviews for each STEP reporting language target population are required.*

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.*

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 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..
 - b) 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

Yunnan(China) will be conducting the STEP Survey with a partial literacy assessment administered only in Chinese in urban areas of Kunming. STEP consortium requires a minimum sample of 2000 completed interviews in such case. So our target sample size will satisfy this requirement without any hesitation, that is, 2000 complete cases. Here a complete case means: 1) all modules in the Household Questionnaire have been administered; 2) all items in the General Booklet were attempted.

According to the recommendation by Technical Standards, the overall initial sample size will be 4050 with 2025 reserve sample included. The following table provides the sample composition.

	Target sample	Reserve sample	Total
In each selected PSU	15	15	30
Number of sampled PSU	135		
Total	2025*	2025	4050

* the final sample size is 2000. Extra 25 sample will be used as reserved sample.

Based on an efficient and tactful operation of this survey, we expect the response rate to be greater than 70%. Although based this expectation, it is not necessary to reserve a sample of 2050 cases, we still prefer the reserve sample at such a size for dealing with any unexpected situations. And the reserve sample will be selected at the same time as the target sample with 15 households in each selected PSU.

2. Special additions to the STEP sample size.

There are no special additions to the STEP sample size.

3. Sample Monitoring procedure

Yunnan data collection team will monitor the number of complete cases on a daily basis. At the end of each day:

(1) each field supervisor will use the Excel file “Supervisor Daily Processing Table” to summarize the number of finalized cases, including “complete cases” and finalized non-response cases.

(2) each supervisor’s summary report will be transmitted to Data Collection Manager and copied to Project Manager at the centre office.

(3) Data Collection Manager will incorporate the supervisor reports into the Excel file “Yunnan (CHINA) Daily Processing Table” , which will indicate everyday and accumulated number of complete cases.

Every Monday, during the data collection period, the central office will send to the STEP Consortium a one week summary of the finalized cases.

I agree with the above,	
XIAOYAN LIANG	Signature:
<TAOZHEN BAI>	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.

Yunnan Team will implement a four stage sample design in the urban regions of Kunming.

First Stage Sampling:

The first sampling units are CEA Blocks. As demonstrated above there are total 5564 PSUs with a median size of 190 household dwellings in the urban area of Kunming In the first stage.

At this stage, 135 PSUs will be selected with PPS methods. The measure size of will be the number of dwellings.

The STEP Consortium will carry out the sampling of units at this stage. Based on the sampling frame provided by us, the Consortium will give back to us the selected PSUs with a random sequence of dwelling numbers to be used as the initial target sample and a random sequence of dwelling numbers to be used as the reserve sample.

Second Stage Sampling:

The second sampling units are household dwellings. The sample frame for the selection of dwellings in each selected PSU will be the updated list of all dwellings in the PSU. At the second stage, a sample of dwellings will be randomly selected from each selected PSU. In addition, a reserve sample of dwellings will be selected for use in cases of non-response. The reserve sample will be selected at the same time as the original sample.

As mentioned above, there are about 190 SSUs in each PSU. According to the Technical Standards, 15 dwellings will be selected in each selected PSU by systematic equal probability sampling. At the same time, a reserve sample of 15 dwellings will be selected for use as needed to ensure that the target sample size is achieved.

The detailed procedure for selecting SSUs in each PSU is as follows:

- (1) Firstly for each selected PSU, we will create a fresh listing of the dwellings with only eligible household dwellings included. Record the total number of dwellings as A.
- (2) Letting A divided by 30, we get the sampling interval B (reserving one digit for decimal). Then we use the integer part of the result of “B*random number +1” as random start point. Here the random number is given by STEP Consortium and provided with the selected PSU.
- (3) From the random start point, we sample the dwellings every B. Once total 30 dwellings are sampled, we arrange a sequential number from 1 to 30 to each sampled dwelling following the order they are selected.
- (4) The dwellings with the same sequential numbers as the initial sample numbers given by Consortium with selected PSUs are regarded as the initial sample, and the rest will serve as the reserve sample.

Third Stage Sampling:

The third sampling units are households in selected dwellings. At this stage, one household will be selected within each selected dwelling with equal probability. This will be only used in

such cases in which there is more than one household within a selected dwelling.

The general description of the household selection procedure is as follows:

(1) The interviewer will list on a separate sheet the names of the household heads for each household residing in the selected dwelling, and will also record the birth date (at least year and month) of each household head;

(2) The interviewer will then assign a sequential number to each household head, with the youngest household head being assigned number 1, and the oldest being the last on the list;

(3) The largest sequential number (that of the oldest household head), say C, indicates the number of households within the selected dwelling.

(4) Each Household Questionnaire will include a table of random numbers to be used in selecting a household to interview.

(5) The Interviewer refers to the table of random numbers and, reading across the table row(s), chooses the first number in the table, say D, that is less than or equal to C.

(6) The Interviewer then refers to the list of household heads and selects the household corresponding to the number D. The Interviewer must then proceed with the interview of the selected household.

Fourth Stage Sampling:

The fourth stage sampling units are eligible members in each selected household. The main respondent will be randomly selected with equal probability in each selected household from among all household members aged 15 to 64 years.

The selection method for the main respondent is described as follows and also indicated in the Household Questionnaire.

(1) The Interviewer should first list all household members in the household roster part of the Household Questionnaire and then assign a sequential number to each household member that is 15 to 64 years of age.

(2) Each Household Questionnaire will include a table of random numbers to be used in selecting a person to interview.

(3) A question in the Household Questionnaire guides the Interviewer to determine the number of eligible household members, say A.

(4) The Interviewer refers to the table of random numbers and, reading across the table row(s), chooses the first number in the table, say B, that is less than or equal to A.

(5) The Interviewer then refers to the household roster and selects the eligible household member numbered B. **This is the eligible person to be interviewed without any**

replacement allowed.

The Interviewer must then proceed to interview the selected person if the person is at home. Otherwise, the Interviewer would arrange a return visit to interview the selected person.

We will be responsible for getting the updated list of all dwellings in the PSU for the second stage sampling described above. If this list is not available from administrative sources, then we will carry out a dwelling listing exercise. Based on this full list, the assigned supervisor will carry out the sampling of the households in its own PSU.

On completion of the household listing operation, we can show it to supervisor from the country team. But according to the provisions by government, we cannot deliver it to the World Bank. Of course a listing of each selected household can be delivered freely.

5.1 Sample Allocation and Selection

As mentioned above, a total sample size of 4050 households is required with half of it as reserve sample in case of non-response.

135 PSUs of total 5564 PSUs will be sampled at the first stage by PPS method. The total sample will be allocated to 135 PSUs with equal numbers in each PSU, that is, 30 households in each selected PSU. 15 of them consist of an initial sample and others as a reserve sample.

In each selected PSU, the households will be sampled by systematic random sampling method. Note that the initial sample and reserve sample will be selected at the same time and assigned into two parts according to the rules from STEP Consortium.

If and only if a selected household is regarded as a non-response case, the corresponding reserve sample should then be used. Furthermore the reserve sample should be used according to the sequence in which it is given.

In each selected households, there will be only one eligible member to be selected as main respondent to finish the individual modules and literacy assessment in the questionnaire. The selection method is based on the Kish-style within-household random sampling and is described thoroughly as above.

I agree with the above,	
XIAOYAN LIANG	Signature:
<TAOZHEN BAI>	Signature:

6.0 LITERACY ASSESSMENT

STEP Standard

A participating country will implement the partial 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.

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.

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.

<Yunnan(CHINA)> will implement a partial literacy assessment as prescribed by the STEP management.

This partial literacy assessment only depends on a General Booklet, which is comprised of two sections. Section A is an assessment of reading skills, which is comprised of three parts; Section B is a set of CORE literacy items. The assessment will be administered in Chinese only.

The work related with literacy assessment portion of the STEP Survey will include:

- (1) translation and adaptation of the literacy materials
- (2) administration of the General Booklet to each respondent
- (3) capturing the response of Section A and scoring the section B
- (4) data entry of this module

The administration of the assessment will strictly follow the instructions provided by the STEP Consortium and ETS.

For the Reading Component, each interviewer will use a stop watch to time this exercises in the General Booklet. The stop watch will be bought by the team and give it to the interviewer as a gift after the project finished.

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.

The methods and processes of translation and adaptation will follow the requirements from STEP Consortium and ETS, which include:

- (1) These materials will be translated into Chinese respecting the text size and formatting.
- (2) There will be an independent translation of all texts by two translators who have no contact with each other. Language specialist will serve as reconciler to reconcile the two translations and to create the VFFs.
- (3) The initial translating results will be sent to the World Bank and will be verified by international verifiers. Then if necessary, we will revise the materials based on the feedback.

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.

We will print the literacy assessment booklets after receiving the final verified Chinese version. As prescribed by the STEP Consortium, 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 original English version.

The approved assessment booklets from ETS are all printed (not photocopied) in A4 size pages according to the printing requirements. After printed, the booklets will be bound as saddle-stitched booklets.



I agree with the above,	
XIAOYAN LIANG	Signature:
<TAOZHEN BA>	Signature:

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

1. Rationale for the inclusion and placement of country-specific questions on the Household Questionnaire.

There are no country –specific questions on the household questionnaire.

2. Description of the pretest strategy

Yunnan (China) will implement the Household Questionnaire prescribed by the STEP Consortium. It will be pre-tested as required by the STEP Consortium and as specified below.

A qualitative pre-test

We will conduct a qualitative survey of a set of 30-40 key questions taken from the household and individual questionnaire that the World Bank has developed. The qualitative survey takes the form of open (in-depth) interviews. We will produce a report that summarizes the key results of the surveys.

We will keep active collaboration with the WB in the process of definition of activities, both in

implementation and analysis.

(a) The sample of interviewees should follow these general rules: half of the participants will be from low-income households, and half from medium/high income and the overall sample should be balanced in terms of gender and age composition. Here the standard used to divide low-income and medium/high income is decided based on the practical situation of Kunming by us.

(b) Translate in [Chinese] at least the questions to be analyzed in the in-depth interviews (provided by the World Bank), and lead-in questions as required.

(c) Choose 2-4 key, high-level members (such as field coordinator and supervisors for the full-length survey) of the firm to be trained during a one-day training session carried out by video link from World Bank HQ.

(d) Conduct up to **25** open (in-depth) interviews in [the urban areas of Kunming] within a maximum of 2 weeks.

(e) Create a full report on the qualitative pre-test, based on a template provided by the World Bank team.

(f) Present and discuss results from the qualitative interviews with the World Bank one week after the end of the interviews by the firm.

A pilot survey

Questionnaires will be adapted as a result of information gathered during the qualitative pre-test. The questionnaires will be finalized in English and will then have the changed questions translated into [Chinese] following the same translation method as mentioned above for the questionnaires.

These finalized questionnaires and sampling methodologies will be tested in a pilot test of 20-30 households. The household questionnaire will be administered in the households, choosing the main respondent and applying all parts of the questionnaire, including the literacy exercises.

(a) The pilot test will be conducted approximately half in low -income households and half in medium/high income households. The households will be selected in areas that do not form part of the sample of the main survey.

(b) The pilot test will be led by the main trainer(s) for the interviewer training, and will involve 5-6 senior persons as interviewers for the pilot. These senior individuals will then be assisting in the interviewer training and will be team supervisors in the survey. This will allow a further familiarization of the instruments for the survey and will allow for effective debriefing from senior persons.

(c) Create a brief (5-10 pages) report on the pilot with recommendations for any changes to the questionnaire and for implementation. This report will be based on a template provided by the World Bank team. There will be a debriefing by phone of the main results of the pilot survey.

We will work in close collaboration with the World Bank team, in identifying areas that may require adjustments (wording, fluidity, translation, etc.)

Associated adjustments to the implementation manual and training program for interviewers and field operators will be made upon completion to the adjustments to the instruments. We will then update training accordingly using the adjusted questionnaire and implementation materials .

3. 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.

4. Printing of the Household Questionnaire

We will print the Household Questionnaire after receiving the final verified Chinese version. As prescribed by the STEP Consortium, 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 original English version.

I agree with the above,	
XIAOYAN LIANG	Signature:
<TAOZHEN BA>	Signature:

8.0 DATA COLLECTION

STEP Standard

Each participating 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 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.*
- 5) *Progress Reporting: Each week during the survey period, each country will submit to the WB Team a data file containing all the entered survey data to date.*

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.

8.1 Survey promotion Strategy

8.1.1 Community Communicator System

A community communicator system will be established in order to reduce the difficulties of contacting the selected households and to acquire their cooperation as far as possible.

For a selected PSU, we will select one local resident who live in the same community to serve as a communicator. Their responsibilities include:

- (1) understanding the importance and significance of this survey and propagandizing them to the selected households
- (2) contacting the selected households with interviewers at the first time and helping interviewers persuade the respondents to accept the survey from the beginning to end
- (3) helping interviewers make an appointment with the respondents
- (4) helping supervisors persuade the refusals to accept the survey

Such a communicator may be a worker of local resident committee or an enthusiastic local resident. The most important requirement is that they must be familiar with the local resident and have a certain influence on them.

8.1.2 Advance Survey Information

Advance Letter and Brochure

A major factor that can influence response is the trustfulness and perception from the respondents. Therefore, a formal advance letter will act an important role in such surveys. Interviewers typically report that they are better received if respondents have read the introductory letter and that official-looking letters on Government stationery or other formal institutions have a good likelihood of being read.

So in addition to the selected communicator, the first contacts with selected households will include a carefully worded, attractive advance letter from the Research Centre. If necessary, this letter will be sealed by Statistical Bureau of Kunming.

Moreover, a bilingual brochure will be provided to the respondents on the spot for their further understanding of this survey. This brochure will provide full information about the study, including the survey purpose, voluntary nature of participation, benefits to the respondent/population, etc. The bilingual format will be conducive to enhance the formality of the survey and then the trustfulness from the respondents.

The letter and brochure will contain the contact information of Research Centre to facilitate respondents to affirm authenticity of the survey.

What should be noted is that this brochure can be read on the spot, but for the particular consideration, it would not be left with the households. This consideration refers that this is a survey with foreign institute and now it is some sensitive in China, so we must limit the information about this survey in a limited scope.

8.2 Contact Strategy

The first contact will be by means of the community communicator and the advance materials described in the previous section.

All selected households will be firstly contacted by a personal visit from a STEP Interviewer accompanied by the communicator and the advance materials described in the previous section.

When nobody is at home, the interviewers will be instructed to try to establish a likely time when someone can be reached under the help of the communicator. At least three contact attempts will be made before coding the case as a non-contact.

Contact attempts subsequent to the first attempt will be scheduled according to the information received during the first contact attempt, or at different times of the day and different days of the week.

Trained Interviewers

Major factors in gaining respondent cooperation are the respondent's perception of the survey and his or her reaction to the interviewer. The respondent must be made to feel that he/she will be making a valuable contribution to an important research effort. The spokesperson for the study is the interviewer; therefore, each interviewer will be well-trained to discuss the merits of the STEP. To adequately prepare the interviewing team, all interviewers must attend an in-person training session, where time can be spent developing these necessary skills.

Interviewer Identification

Establishing the legitimacy and importance of the survey effort for the respondent encourages respondent cooperation. Many people are suspicious of any stranger who comes to the door, or show indifference with a survey. A number of procedures help to establish the legitimacy of an interviewer's visit and to enhance the importance of the survey.

The most effective way of overcoming suspicion and improving importance is through a good introductory statement during which the interviewer shows an ID badge and a copy of the advance letter and brochure.

8.3 Response Rate Strategy to Minimize Non-response

In an effort to obtain a response rate of 70 percent, a number of non-response strategies are being considered. They include:

- 1) **Community Communicator System:** the most important function of a communicator is to persuade the respondents to finish the whole survey. This will be a key strategy to minimize non-response.
- 2) **Interviewer Training:** The interviewers will be instructed in techniques to obtain the cooperation of individuals who initially refuse or are reluctant to participate in the

survey.

- 3) **Follow-up:** If the interviewer is unable to gain the cooperation of a selected individual, the case will be referred to the interviewer's immediate supervisor who will take steps to convert such cases to completed interviews. In addition, the communicator will help the team to persuade the respondents.
- 4) **National Pride Stimulus.** We will emphasize that this is a multi-country survey and more than ten thousand people are involved in it and the respondent will be the representative of China. We believe that national pride will be an important stimulus to make the respondents happily accept the interview.
- 5) **Token Gift and Material Incentive:** If approved and funded, the respondents will firstly be given a small gift (e.g., a bottle of shampoo) as a token of appreciation for their participation in this task. It should help obtain a higher screener response rate, and provide an upfront incentive/inducement to participate for households with eligible study participants. In addition, respondents who complete the survey will receive a material incentive for their cooperation.

A field supervisor will revisit each household in the following situations:

- 1) A household refuses or does not begin the interview because of special circumstances (result codes 1 or 2).
- 2) A household stops before finishing the Household Module, Module 1.
- 3) A household where the selected individual is not able to begin the questionnaire – for refusal, for special circumstance, absence, other reasons.
- 4) A household where the individual stops without finishing the individual modules 2-7.
- 5) A household where the individual stops without finishing the assessment exercises, Module 9.

8.4 Interviewer Hiring Plan

The Interviewer Supervisors will recruit interviewers for their regions, supervised by a Field Manager. Approximately 60 interviewers will be hired. To ensure efficient trips to the field and knowledge of the local geography and population, interviewers will be hired from the areas in which interviewing assignments are located.

Interviewer Supervisors will begin interviewer recruitment by drawing from experienced interviewing staff. They will use the interviewer database which is established from previous projects. The Interviewer Supervisors will consider experienced interviewers who had high evaluations on previous projects and live in the areas in which the interviewing assignments are located.

Other sources of candidates will include applications from individuals with and without interviewing experience. These applications will be generated by ads placed on notice board in university or local resident committee. When considering applicants for an interviewing position, supervisors will assess the applicants' basic skills and personality traits. They will look for interviewers who have basic reading and computational skills, who are able to follow instructions, and who have an aptitude for working with a computer and using a keyboard. Desirable personality traits include endurance of hardships, receptivity to others' ideas, open-mindedness, and motivation.

All candidates will be screened by face-to-face talking for availability, level of interest in the project, related job experience, and general ability to communicate personably and effectively. Those who meet the basic qualifications of availability and interest and who have generally good communication skills will be interviewed in depth by supervisors, who will then select the best candidates for interviewers for the STEP.

8.4.1 Number of Interviewers

Approximately 60 interviewers will be hired for data collection. A priority list will be established during the hiring process to identify suitable candidates in the event of interviewer turnover during the STEP.

8.4.2 Method of Payment

The interviewers will be paid based on the number of finished questionnaires rather than the number of hours worked. Because the amount of households to be interviewed is predetermined and the selected households are identified, we think this payment way is the best way to improve the work efficiency and to ensure the progress. Especially such a payment way is easy to be understood, counted and accepted by interviewers. In addition, the interviewers will be compensated for any travel costs that are necessary to carry out an assignment.

8.5 Interviewer Training Plan

Before beginning the actual household and individual survey in the field, we will respectively carry out a rigorous training of field supervisors and interviewers. Each training should last a minimum of 5 full training days, including field practice where each trainee will interview at least two households and two selected individuals.

Firstly we will carry out the training of field supervisors. This training is expected to last a minimum 5 full training days, including introduction to the project, operational progress and requirements for the field work, explanations of the questionnaire, discussions of implement strategies, skills and other related problems, and field practice. At least 15 field supervisors are expected to participate this training as well as the scoring and data entry supervisors. This training will be carried out at the centre office by Project Manager before the Spring Festival holidays. After the training, a final list of team supervisors will be submitted to the World Bank for approval.

After the Spring Festival holidays, we will carry out the training of interviewers. This training will be carried out twice for two groups of interviewers. Each training is expected to last a minimum 5 full training days, including introduction to the project, operational requirements for the field work, explanations of the questionnaire, discussions of interview skills, and field practice. 60 interviewers will be divided into two groups to participate the training at the centre office. The training staff will be consist of Project Manager, Data Collection Manager and trained supervisors. Every 4 interviewers will be directed by 1 supervisor during the training. After the training, a final list of team interviewers will be submitted to the World Bank for

approval.

All the trainings will follow the training plan and contents provided by the World Bank except for the time (because of the Chinese Spring Festival Holidays) and personnel arrangement.

There will be written exercises and test done throughout the training. These will be marked by training staff. And there will be supplemental after hour sessions available for trainees who want more practice, or who have been identified in training as needing help.

Throughout training, interviewers will receive close attention and will be given extensive hands-on experience with the STEP materials. Training scripts and exercises will be designed to address situations that the interviewers are likely to encounter and to build in complexity.

At the end of training, trainees will be paired for scripted role plays and lastly, set up with a “live/paid respondent” interview. All training will be conducted in a classroom or conference type facility.

Training staff should hold daily debriefing sessions to discuss progress, problem trainees and any technical or administrative issues that have been flagged during the day.

Data entry training: Data Processing Manager and supervisors will participate in the supervisor training, but the data entry training will be carried out separately. At least 10 data keyboarders will be trained by Data Processing Manager. The contents include requirements for data entry, practice with data entry program and related issues. This training will be hold before or at the beginning of field work and expected to last 2 days.

Scorer training: Main translators will sever as Chief Scorers and will be responsible for scorer training. This training will also be held separately during the field work but before the scoring work. In addition to Chief Scorer, 3 extra scorers will participate in this training. The contents include requirements and rules for the scoring, practice of scoring, studying and practice of data entry program for literacy assessment module and task allocation. This training will last 2 days and will be under the directions of Project Manager.

8.6 Interviewer Supervision Procedures

We will ensure there is a rigorous supervision process and mechanisms (including spot-checks) in place to ascertain an appropriate implementation of the survey (verifying adherence to the sample selected), correct implementation of tests and adherence to established interview protocols.

If any interviewer’s work is found to be suspect, the interviewer will be dismissed and all of the interviews done by that interviewer will be redone in their entirety.

8.6.1 Number of Supervisors

The STEP main study will have <15> Interviewer Supervisors who will be supervised by a Field Manager.

These supervisors will be from the graduate students in the college of Yunnan University of Finance and Economics, and the supervisors from the research companies cooperated with our Research Centre. They will be trained thoroughly and then arranged to each team.

8.6.2 Supervisor Responsibilities

Each of the Interviewer Supervisors will supervise about 4 interviewers. The supervisors' responsibilities will include:

- Attend and participate in interviewer training, and direct interviewers in their respective team during the interviewers training;
- Check the household listing and carry out the household sampling of their respective PSUs;
- Assign cases to their interviewers and establish clear production and cost goals with interviewers;
- direct the interviewers work during the field work, particularly accompanying each interviewer in their respective team to finish at least one visit;
- Hold a weekly meeting with their interviewers to review status of each of their cases, find out how much they have worked, review any problem situations, and motivate them to finish on time; they will also need to be available to receive calls from interviewers who have problems throughout the week;
- Monitor progress of data collection, review non-response reported by interviewers, and implement reassignment and conversion procedures;
- check the finished questionnaires by interviewers and fill the necessary codes, and then deliver the qualified questionnaires to data processing group;
- Review interviewers reporting of time and expenses;
- Perform validation of a designated fraction of each interviewer's work by telephoning the respondent and asking a brief set of questions;
- Edit the data collected from each interviewer's 3rd and 10th completed case; and
- Report to the Field Manager on a weekly basis (or more frequently if a problem arises) on the progress of the survey in their region.

In particular, the supervisors will carry out a verification of each interviewer's visits by a revisit to 15% of the households in each interviewer assignment and a follow-up of a further 15% of households by telephone in each interviewer assignment. The households involved in the verification process will be randomly selected within each PSU.

If any interviewer's work is found to be suspect, the interviewer will be dismissed and all of the interviews done by that interviewer will be redone in their entirety.

The supervisors will have to carry out the selection of the households (second stage of the sampling design) in their respective PSUs, as recommended in the Technical Standards for the Design and Implementation of the STEP survey. They will submit the listings of dwellings and the corresponding sample selection to the Project Manager.

The supervisor will also be in charge to follow-up with households refusing the interview, in order to try and convert these households to taking the interview. If successful, an interviewer

will be dispatched to interview that household.

8.7 Progress Reporting

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

I agree with the above,	
XIAOYAN LIANG	Signature:
<TAOZHEN BAI>	Signature:

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

<Yunnan (China)> will ensure the identification of each sampled case by including a unique sample ID on all documents pertaining to the sampled case.

1. Assessment Booklet Identification

<Yunnan (China)> will print a sequential booklet ID on each printed assessment booklet.

This ID will be used to keep track of the assessment booklet. Each booklet will be assigned to an initial selected household. If the household is replaced, then the booklets will also be replaced in order to ensure the uniqueness.

2. 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.

I agree with the above,	
<XIAOYAN LIANG>	Signature:
<TAOZHEN BAI>	Signature:

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 following codebooks will be used to code education, occupation, and industry information from the Household Questionnaire.*
 - 01 *‘1997 International Standard Classification of Education (ISCED)’ will be used to code the education variable (i.e. all questions related to the level of educational attainment).*
 - 02 *‘International Standard Classification of Occupations (ISCO 08)’ will be used to code the occupation variable. The level of disaggregation will be 3-digit.*
 - 03 *‘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.

9.2.1 Data Capture

Each item in the respondent assessment booklets will be scored and the score will be transcribed on a scoring sheet. The responses from the Household Questionnaire and the Assessment Scoring Sheets will be manually keyed from the completed questionnaire.

The 'write-in' entries from the response category "Other. Please specify" for all questions where this category has been selected will also be key-entered. Furthermore, we will be responsible for coding these 'write-in' responses and providing the code set to the STEP Consortium. Also, we will ensure that these coded responses are included in the editing of the data.

Data capture software will use data entry programs provided by the World Bank: one (written with Excel macros) for the household survey questionnaire, and a second one (provided by ETS) for the literacy module.

The data entry will be done concurrently with field work, so that information from the data entry can inform the process of the survey. For each selected PSU, the data entry will be carried out no later than 5 days after the finalization of the PSU cases.

Two teams will be responsible for the data entry respectively: data entry team for the household questionnaires entry and scoring team for the assessment scoring sheets. Each team will be trained for their work before data entry starts.

Data entry will adopt the way of double blind entry, that is, the data capture will be done twice, by two different operators, both for the household questionnaires and the assessment scoring sheets.

All the data gathered from the field will be entered but the load on the data entry operators will be different depending on the outcome of the interview. The possible outcomes are (after the required number of re-visits):

1. household refuses the interview - interview ends, but the cover page of the household questionnaire (including location, dates of contact and refusal information) will be filled.
2. household accepts interview, but there are no qualifying adults in the households - cover page, short household roster and the dwelling modules will be filled.
3. household accepts interview, there are qualifying adults, selected adult completes the interview (both household questionnaire and full literacy module).
4. household accepts interview, there are qualifying adults, selected adult completes the interview (both household questionnaire and general booklet, fails the Core)
5. household accepts interview, there are qualifying adults, selected adult starts the interview: completes the full household questionnaire but stops in the middle or refuses to take the literacy module
6. household accepts interview, there are qualifying adults, selected adult starts the interview: does not complete the full household questionnaire but completes the literacy module (either full, or only general booklet if he fails the Core)
7. household accepts interview, there are qualifying adults, selected adult starts the interview: starts the household questionnaire / literacy module, but finishes neither (not

expected to be very large percentage of interviews)

Of the situations described above, situation 3 is the most complete and situation 1 requires entering just the basic cover page. Situations 4-7 are in-betweens.

The collected data will be transmitted weekly to the World Bank, starting with the first week of data collection.

Each household and each individual who completes the individual questionnaire has a unique code for data entry. Each interviewer, supervisor and data entry person also have a unique code which will be recorded for all questionnaires.

9.2.2 Data Capture System Test

The data capture specifications and system will be tested before implementation. The testing of the data capture system will involve a thorough review of the programming specifications prior to the development of the computer programming code, and the subsequent testing of the programs prior to the start of the data capture operation. Testing will be carried out by preparing mock survey instruments (Background Questionnaires and Scoring Sheets), passing them through the data capture system, and then reviewing the resultant data file outputs. Only when satisfactory data capture results are obtained will we commence the capture of the live STEP survey data.

9.2.3 Data Capture Verification

Data capture of the Background Questionnaires (paper and pencil) and Assessment scores will be 100% verified. The data capture will be done twice, by two different operators. All differences in the captured data will be resolved.

9.2.4 Coding

Uniform coding of the questionnaire and assessment data is essential.

Pre-coded response categories on the Background Questionnaire will correspond to the International Record Layout codes.

The '*1997 International Standard Classification of Education (ISCED)*' will be followed in coding the education variable collected during the administration of the Background Questionnaire.

Each respondent's occupation will be coded using the '*ISCO Job Titles*'.

The '*International Standard Industrial Classification Of All Economic Activities, Third Revision*' will be used to code the industry variable.

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

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%.

9.2.5 Scoring Task Booklets

The assessment instrument will be scored using the Scoring Guides provided by the STEP Consortium.

9.2.5.1 Recruiting and Training Scorers

Recruiting qualified scorers to evaluate responses is crucial to the success of the assessment. The two translators who translate the questionnaires and scoring guide will serve as the Chief Scorers and 3 extra scorers will be hired and trained to score the STEP pilot instruments. The scorers selected will be required to hold at least some post-secondary education.

Inter-rater agreement refers to the consistency with which individual scorers assign a score to a question. This consistency is critical to the success of the STEP and a number of methods will be used for monitoring this level of agreement.

First, scoring supervisors will review each scorer's work to confirm that the scorer applies the scoring criteria consistently across a large number of responses and that the individual does so consistently across time. Scoring supervisors will evaluate approximately 10 percent of each scorer's work in this process.

Next, all assessment items will be subject to an agreement check involving a second rating by a second scorer across the scoring process. The procedure outlined by ETS will be adhered to so that one-third of the assessment documents is re-scored. The results of the agreement check will be monitored on a weekly basis and will provide the scoring supervisor with inter-rater agreement percentages and the percent agreement for individual exercises. Individual scorers will receive feedback on their level of performance and, if necessary, receive feedback on particular items that they may have problems scoring. If particular items seem to be giving problems to a majority of scorers, retraining will be held for those items.

Consistent performance among scorers is paramount for the assessment to produce meaningful results. Therefore, we will carefully monitor the scoring process, which will result in early identification of problems, and flexibility in training and retraining scorers. According to information from previous literacy assessments, the average percentage of agreement among the scorers within a country for all open-ended items was 97 percent. The project team feels confident that this same high standard will be maintained for the STEP.

9.2.5.2 Documenting the Scoring Process

All aspects of scoring constructed responses will be fully documented. In addition to warehousing the actual student booklets, we will keep files of all training materials and inter-scorer agreement reports. All the procedures used to assemble training packets, train scorers, and conduct scoring will be documented scoring reports. These scoring reports will also include all methods used to ensure scorer consistency, all reliability data, and all quality control measures. We will also summarize the basic scoring procedures and outcomes in the

final survey report.

9.2.6 Creation of International Data File

STEP data will be delivered to the STEP Consortium in a clean data file according to the International Record Layout (IRL) specifications. To facilitate the creation of this file, data collection has been designed with this final format in mind. Most data fields in the data collection process have been named with the same names as those fields in the IRL file. The response category codes in the Background Questionnaire were pre-coded to correspond to the codes required in the record layout.

9.2.7 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.

We will perform the edit of our STEP data file to identify and resolve errors in the data based on the computerized editing system provided by the STEP Consortium. We will be responsible for ensuring that its final data file summated to the STEP Consortium is error-free.

9.2.7.1 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.

9.2.7.2 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.

I agree with the above,	
XIAOYAN LIANG	Signature:
<TAOZHEN BAI>	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 China 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

China will provide the most recent counts of the benchmark variables to the STEP Consortium. The benchmark weights will be created using data from the <[2010 Census data of Kunming](#)>.

I agree with the above,	
XIAOYAN LIANG	Signature:
<TAOZHEN BAI>	Signature:

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.

1) Outline the China's data confidentiality requirements

The STEP Survey will be conducted on the authority of the < Statistics Law of the People's Republic of China >. All persons on the STEP project team will have professional secrecy clearance. The interviewers and interviewer supervisors must take an oath of confidentiality as a condition of employment.

The STEP data collects personal information from respondents. It will be managed according to the rules set out in the < Statistics Law of the People's Republic of China >, which requires the protection of the identities of individual respondents. This protection is assured by removing or collapsing selected variables on the data set. Once screened, the data set is considered to be in the public domain and available to all users for the cost of reproduction.

In addition, the STEP data collection is subject to the < Statistics Law of the People's Republic of China > and the guidelines from the <Data Inspectorate>, which requires very strict procedures for handling and securing personal data. Data will not be published or delivered in such a way that a respondent's identity can be revealed.

All prospective respondents in the survey will receive an introductory letter that will include information about the use of the data and any linkage to other administrative files. The letter will inform the respondents about their rights under the < Statistics Law of the People's Republic of China >, such as the right to revise or delete data and the right to withdraw from the survey at any time. Respondents are informed in the letter that participation in the survey is voluntary. If a respondent is under the age of 18, both the respondents and their

parents/guardians will receive introduction letters. Parents/guardians have the right to decline from participation in the STEP on behalf of a respondent under age 18.

2) Outline the steps to ensure data confidentiality. Ownership and sharing of STEP data.

The survey instruments, the sampling, and the information gathered by the field workers cannot be used for personal or professional goals by the local consultant firm, field workers or the coordinator and advisor without the prior request and an approval by the World Bank. The data collected is completely confidential and shall not be revealed to any source by the firm. The ownership of any information and data belongs to the World Bank.

We will submit the international data file to the World Bank STEP team (regional team and core team). Data release outside these World Bank teams may not be done until 6 months after the submission of the final data files and under the permission from World Bank.

All interviewers, data entry, coders and supervisory staff must sign affidavits of confidentiality and non-disclosure for the survey activities. A separate form of non-disclosure for the literacy booklets and training and scoring material from ETS is also required.

I agree with the above,

XIAOYAN LIANG

Signature:

<TAOZHEN BAI>

Signature:

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.

1. Quality Assurance

Quality assurance will be addressed at all stages of the STEP. Many of the steps to ensure quality are outlined in earlier sections of this document. The following sections summarize the quality assurance measures to be implemented during the STEP.

1) Team Composition

The STEP project team is comprised of experienced, knowledgeable personnel with expertise in 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.

The following key persons on the STEP project team are committed to the project for the duration of the pilot and main surveys:

1. <Mr. Bai Taozhen>, responsible for the overall management of the STEP,
2. <Mr. Li Xingxu>, responsible for survey design and implementation,
3. <Ms. Xia Weiping>, responsible for the field operations,
4. <Mr. Chen Fei>, responsible for the processing of the STEP data and the creation of the STEP International Data File.
5. <Ms. He Yu>, responsible for the translation/adaptation of all STEP materials and scoring works.

2) Expert Meetings

The National Project Manager is committed to attending the international meetings deemed

necessary by the STEP Consortium. Similarly, other team members will participate in the international meetings when requested by the STEP Consortium.

3) **Survey Instruments**

Background Questionnaire

The Background Questionnaire will include the international 'core' questions, and corresponding response categories and coding schemes developed by the STEP Consortium. The questionnaire design and layout will be consistent with the STEP Consortium requirements.

A pretest of the Background Questionnaire will be conducted with a non-probability sample of 50 members of the target population.

A copy of the Background Questionnaire, in Chinese, will be provided to the STEP Consortium for review and approval.

Assessment Instrument

The Assessment Instrument will be modeled after the master Assessment Instrument provided by the STEP Consortium. The instrument will be organized in the same way as the master instrument - 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 Instrument provided by the STEP Consortium.

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

A copy of the Assessment Instrument, in Chinese, will be provided to the STEP Consortium for review and approval.

4) **Sample Design**

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 will be used. The minimum STEP sample size goal will be satisfied – 2000 completed interviews is planned.

The sample selection of one target person within a selected household will be carried out by the interviewer using a Kish-style selection grid to ensure uniformity in selection procedures. This sub-selection procedure will be verified for a sample of five percent of the interviewer's assignment by an interviewer supervisor.

5) **Data Collection**

The Data Collection Manager goals for quality assurance are as follows:

Interviewing Staff

- a) Hiring of qualified, experienced interviewers and interviewer supervisors,
- b) Classroom training of 5 days and a home study program for all interviewers,
- c) Regular meetings between interviewers and interviewer supervisors,
- d) Interviewer Observation Program,
- e) Sample Verification of Each Interviewer's Cases.

Response Rate

- a) A response rate goal of 70%,
- b) Survey responses will be monitored throughout the collection activity,
- c) Development and implementation of a contact strategy and a strategy to minimize non-response (described earlier in this report).

6) Data Processing

The following quality assurance procedures are planned:

- a) Test of the data capture system
- b) A 100% verification of the captured data.
- c) A 100% verification of the coders' work.
- d) Development and implementation of scoring quality control procedures to ensure inter-scorer agreement.
- e) Creation of the STEP International Data File according to the record layout specifications provided by the STEP Consortium.
- f) Data Editing

I agree with the above,	
XIAOYAN LIANG	Signature:
<TAOZHEN BAI>	Signature:

13.0 SCHEDULE

STEP Standard

Each country will provide a schedule of activities for STEP.

Rationale

A schedule of activities is a key planning tool for the implementation of STEP. Since the schedule may vary from country to country it is important that each country provide a project schedule that reflects its expected activities and time period for their completion.

The development of a schedule of activities is also an important quality control task since it outlines the major activities that are required to implement the STEP Survey. A country's project team and the STEP Consortium will review these activities to ensure that all important activities have been included in the schedule of activities and to ensure that the expected timeline is realistic for the completion of the activities.

Report Requirement

1. Provide the schedule of activities for the STEP implementation.

The STEP survey is planned for November 2011 to June 2012. The following table provides the planned schedule of tasks.

Activity	Time Period
A. Survey Preparation	November, 2011 – February, 2012
1. Preparation of Sample Design and Weighting Specifications	November-December, 2011
2. Prepare National Design and Implementation Report	November-December, 2011
3. National Background Questionnaire revision	December, 2011
4. Translation and adaptation of revised survey instruments	January, 2012
5. Assessment Booklet composition	December, 2011
6. Hiring of Field Staff	January-February, 2012
7. Preparation of interviewer materials and training package	January, 2012
8. Printing of survey materials	February, 2012
9. Sample Selection	February, 2012
10. Preparation of interviewer assignments	January-February, 2012
B. Data Collection	February – April, 2012
11. Interviewer Supervisor Training	February, 2012
12. Interviewer Training	February, 2012
13. Main Survey field collection	February – April, 2012

C. Data Processing	February - May, 2012
14. Scoring	February - May, 2012
15. Data Capture	February - May, 2012
16. Coding	February - May, 2012
17. Weighting	February - May, 2012
18. Data Editing	February - May, 2012
19. File construction and delivery	February - May, 2012
D. Survey Evaluation	May- June 2012
20. Obtain feedback re survey procedures from staff	May, 2012
21. Review survey procedures	May 2012
22. Preparation of Main Survey Evaluation Report	May-June 2012

I agree with the above,	
XIAOYAN LIANG	Signature:
<TAOZHEN BAI>	Signature:

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 China is summarized in the following table:

Budget Estimate – STEP - EDITED -	
1) Project Team Salaries	
2) Translation Fees	
3) Sampling Fees	
4) Printing Costs	
5) Pilots Costs	
6) Interviewees Rewards	
7) Interviewers Payments	
8) Questionnaires Rechecking	
9) Data Entry Fees	
10) Data Processing Costs	
11) Conference Fees	
12) Office Expenses	
13) Unpredictable Fees	
TOTAL Estimated Expenditures	

I agree with the above,

XIAOYAN LIANG

<TAOZHEN BAI>

Signature:

Signature: