



United Nations
Educational, Scientific and
Cultural Organization

Organisation
des Nations Unies
pour l'éducation,
la science et la culture

Organización
de las Naciones Unidas
para la Educación,
la Ciencia y la Cultura

Организация
Объединенных Наций по
вопросам образования,
науки и культуры

منظمة الأمم المتحدة
للتربية والعلم والثقافة

联合国教育、
科学及文化组织

Internal Oversight Service Evaluation Section

IOS/EVS/PI/91
Original: English

Evaluation of the Literacy Assessment and Monitoring Programme (LAMP) / UNESCO Institute for Statistics (UIS)

Dr. Kadriye Ercikan
Rubab Arim
Maria Oliveri
Deborah Sandilands

August 2008

The views and opinions expressed in this document are those of the authors and do not necessarily represent the views of UNESCO or IOS who has contracted the study and assured adherence to evaluation standards. The designations employed and the presentation of material throughout this document do not imply the expression of any opinion whatsoever on the part of UNESCO concerning the legal status of any country, territory, city or area of its authorities, or concerning its frontiers or boundaries.

Table of Contents

<i>Executive Summary</i>	4
<i>Recommendations</i>	5
Introduction	8
I. Overview of LAMP	8
Objectives	8
Major Phases in LAMP	9
Intended Outcomes	9
Change of Strategy and Leadership in LAMP at UIS	9
II. Evaluation Methodology	11
Document review and analysis	11
Meeting with key informants.....	11
Questionnaires for Country Managers.....	11
III. Evaluation Findings	12
Technical soundness of UNESCO’s outputs	12
Target Population	12
Tests	12
Procedures	14
Content and Construct Related Validity	15
Linguistic and cultural appropriateness	16
Pre-testing.....	17
Questionnaire	18
Relevance and adequacy of organizational arrangements.....	19
Administration Procedures	21
Scoring	21
Sampling.....	22
Results	22
Risks	23

Capacity Building Efforts 24

Appendix A: List of Documents Used in the LAMP Evaluation Electronic Folders 25

Appendix B: The Questionnaires in English (also distributed in French and Spanish)
..... 26

Appendix C: Management Response..... 27

Executive Summary

The Literacy Assessment and Monitoring Programme (LAMP) has been under development by the UNESCO Institute for Statistics (UIS) since 2003, in partnership with other international agencies and technical experts. The programme was launched internationally on March 10, 2006 and its strategy adjusted in February 2007. The broad objectives of this programme are indicated as follows:¹

- 1. To develop a methodology for providing quality data on the distribution of the literacy skills of adults and young people in developing countries;*
- 2. To obtain high-quality literacy data in participating countries and to promote its effective use in formulating national policy, in monitoring, and in designing appropriate programme interventions to improve literacy levels;*
- 3. To build national capacities in the measurement of literacy, to develop and use LAMP data and methodologies.*

These objectives are intended to improve education statistics and provide methodological tools and capabilities to the participating countries.

The programme uses a sample survey methodology to select individuals 15 years of age and older. The LAMP is intended to measure five levels of literacy and numeracy skills, using a combination of household survey methods and literacy assessments. These surveys and assessments are intended to be conducted on a five- to ten-year cycle. The selected individuals are *“asked to respond to background questions as well as questions designed to assess literacy levels. The questionnaire measures a spectrum of literacy levels from very basic reading, writing, and numeracy to high level literacy skills needed to fully participate in a learning society.”*

The major phases of the LAMP can be organized as follows: (a) development, piloting and field testing phase which started in 2003 and will be completed in 2008-2009; and (b) general implementation to be started in 2009. LAMP is in the field testing phase which involved 5 participating countries (El Salvador, Mongolia, Morocco, Niger, and Palestinian Autonomous Territories). The field test has been completed in the Palestinian Autonomous Territories, Morocco, and Mongolia and is ongoing in Niger and El Salvador.

This report has been prepared based upon discussions and meetings with key UIS staff working on LAMP, a review of documents received from UIS-UNESCO, and responses to a questionnaire by the country managers of the participating countries. A full list of documents used in the preparation of this evaluation report is provided in Appendix I.

¹ Italicized text is extracted verbatim from LAMP reports, memos, and documentation.

Recommendations

LAMP is still in development and the following recommendations are intended as guidance for designing and implementing an excellent programme. Some of these recommendations are intended to address challenges UIS and the LAMP participating countries have faced. We have also made recommendations based on research and professional knowledge about conducting large-scale assessments and standards for testing and measurement for UIS to consider in the upcoming phases of the program.

1. Balancing within country validity and cross country comparability

The two goals of UIS are: (1) providing useful and high quality data at the national level; and (2) comparability of results among the participating countries. Achieving these goals simultaneously presents constraints that require careful balancing of within country validity and across country comparability. For example, within country priorities may require development of tests and instruments that are culture-specific, whereas cross-country comparability priorities require tests and instruments to be culture-free so that they may be used appropriately in all of the participating countries. We make specific recommendations regarding ascertaining and documenting this balance.

2. Test framework development needs to be based on empirically defined constructs

Assessment design principles recommend that assessment design starts with a well thought-out and well researched construct definition. This construct in turn determines the test framework. In the case of LAMP, definition of adult literacy in the participating countries and cultures should form the basis of the framework development.

Most of the test items in the current LAMP tests are originally developed in English according to research and understanding of what literacy is in English. Test framework need to take into account empirical research on reading in the languages of the participating countries.

3. Obtain and document key validity evidence

Processes and studies need to be developed to gather and document a key set of validity evidence. The following are a list of key validity evidence required for LAMP:

- Evidence for culturally appropriate definitions of literacy and numeracy as target constructs for measurement;
- Evidence that tests are designed according to these definitions of target constructs;
- Evidence that there is a good match between these target constructs and tests;
- Measurement accuracy and reliability evidence.

The documentation of validity evidence needs to be publicly available in the form of technical reports in order for users of LAMP to be informed about quality and properties of LAMP data.

4. Examine the appropriateness of the Filter Booklet as it is used in the field test

The Filter Booklet contains easy reading and numeracy items. However, these test items do require participants to be able to read and write at least at a basic level. Very low literacy participants are expected to fail this test by not being able to answer any of the questions. The order of this test administration may create some frustration and problems among the participants. The data from the field test need to be used to examine the appropriateness of the Filter Booklet in guiding respondents to Module A and B.

5. Revise current tests

Review of some of the test items by the evaluators revealed potential cultural bias.

Current tests need to be revised to minimize cultural bias.

6. Revise translation guidelines

Having instructions that are too narrow may not work for languages in which sentence or grammatical construction of the language does not allow for answers to be placed in the same location. Although the adaptation of names and places in the stimuli are not allowed, creating original items that are less culture-specific would be beneficial. We recommend for LAMP to follow an adaptation approach rather than a translation approach for minimizing cultural bias.

7. Tailor Background Questionnaires to country needs

Background questionnaires need to be developed collaboratively and tailored to country needs. This is an area where country appropriateness needs to be prioritized over cross-country comparability.

8. Follow up on specific country recommendations and needs

The following recommendations came directly from the participating countries. We note that UIS provided support, guidance and processes that are related to these issues. However, the recommendations identified by the countries indicate that these need to be reinforced and enhanced.

- *UIS to provide guidance, communication, and coordination*
Reinforcement of capacity of LAMP teams to lead them to become more autonomous, provide guidance and support in obtaining further funding for LAMP and help coordinate communication and collaboration among countries within regions.
- *UIS to evaluate country team capabilities*
UIS should accumulate better knowledge about each national team (their strong points, their weak points, their problems). A diagnosis of the situation at each country level is needed to have this kind of information.
- *More opportunities for country needs to be expressed and considered*
UIS should give participating countries more opportunities to express their needs and requirements and more involvement in the implementation of these needs and requirements in the action plans. The set up of a scientific committee including a representative from each participating country could address this need.
- *Effective cooperation between countries*
UIS should give more importance to capitalize upon and value successful experiences of participating countries and encourage effective cooperation between these countries (South-South cooperation) by creating synergies so that each country can benefit from the experiences of others.
- *Technical advising process*
Strengthen the technical advising from UIS process in a direct way that would enable countries in full implementation of LAMP.

9. Refine scoring procedures

In the scoring process, double scoring is conducted to check the accuracy of the scoring process. The scoring process needs to be refined so that discrepancies between scores from the two sets of scores (from two independent sets of scorers) be reconciled through an agreement or averaging process.

10. Establish an international technical advisory panel

We recommend UNESCO to establish an ongoing external review process. This can be in the form of an international technical advisory board, consisting of psychometricians,

education policy experts, and statisticians with international reputation. This board needs to have an arms-length relationship with LAMP and review activities, processes, and products and provide a report to UNESCO Headquarters on a yearly basis.

11. A realistic determination of financial and other demands for participation in LAMP

The cost, time frame, and processes involved in LAMP participation need to be identified and communicated to participating and potential participant countries. This goal has been challenging for UIS due to the several changes in leadership, processes and procedures it experienced in relation to LAMP. The goal is expected to be more achievable to refine and finalize procedures based on the field test data.

12. A consistent team at UIS

Changes in LAMP leadership lead to delays and frustration among countries. In the upcoming phases of LAMP, a system and a team need to be in place that maintains stability in LAMP. It is essential for this team to include expertise in measurement, in particular expertise in international assessments and adult literacy. We recommend UIS/UNESCO to establish processes to minimize disruptions in LAMP by maintaining a key set of team members and expand the existing membership to include the two key areas of expertise identified above.

Introduction

This evaluation report of the Literacy Assessment and Monitoring Programme (LAMP) contains an overview of LAMP, methodology used in the evaluation, and a summary of findings from our evaluation of the Technical Soundness of UNESCO's Outputs related to Monitoring and Assessing Learning Achievement. These sections are followed by evaluations of the relevance and adequacy of organizational arrangements, the effectiveness of the quality control and quality assurance procedures, and assessment of capacity development.

The evaluation is based on discussions and meetings with key UNESCO Institute for Statistics (UIS) staff working on LAMP, a review of documents received from the UIS-UNESCO, and responses to a questionnaire by the country managers of the participating countries. A full list of documents used in the preparation of this evaluation report is provided in Appendix A.

I. Overview of LAMP

LAMP has been under development by the UIS since 2003 in partnership with other international agencies and technical experts. The programme was launched internationally on March 10, 2006 and its strategy adjusted in February 2007.

Objectives

The broad objectives of this programme are indicated as follows²:

1. *To develop a methodology for providing quality data on the distribution of the literacy skills of adults and young people in developing countries;*
2. *To obtain high-quality literacy data in participating countries and to promote its effective use in formulating national policy, in monitoring, and in designing appropriate programme interventions to improve literacy levels;*
3. *To build national capacities in the measurement of literacy, to develop and use LAMP data and methodologies.*

The programme uses a sample survey methodology to select individuals 15 years of age and older. LAMP is intended to measure five levels of literacy and numeracy skills, using a combination of household survey methods and literacy assessments. These surveys and assessments are intended to be conducted on a five- to ten-year cycle. The selected individuals are "*asked to respond to background questions as well as questions designed to assess literacy levels. The questionnaire measures a spectrum of literacy levels from very basic reading, writing, and numeracy to high level literacy skills needed to fully participate in a learning society.*"

² Italicized text is extracted verbatim from LAMP reports, memos, and documentation.

Major Phases in LAMP

The major phases of the LAMP involve the developmental and pilot phase which started in 2003 and will be completed in 2008-2009 and general implementation to be started in 2009. The national implementation of the LAMP intends to involve the following major phases:

1. *The development and approval of a national implementation plan;*
2. *The development and certification of the content and design of survey documents in national languages;*
3. *The conduct of a field test to establish the psychometric and measurement properties of the survey instruments and to confirm key aspects of survey cost and quality;*
4. *Processing and analysis of the field test results and subsequent revision of the survey instruments and procedures to reflect the results of the field test;*
5. *Administration of the final instruments to a probability sample of the adult population;*
6. *Processing, analysis, and reporting of the main assessment results.*

LAMP is in the third phase of development for the five participating countries (El Salvador, Mongolia, Morocco, Niger, and Palestinian Autonomous Territories). The field test has been completed in the Palestinian Autonomous Territories, Morocco, and Mongolia. These countries have started data capturing, coding, and scoring. In Niger, the field testing is ongoing in the Zinder and Tillabéry regions. In El Salvador, the country manager reported (April, 2008) that they are presently in the planning stage (instrument adaptation, reproduction, and selection of field personnel). The field test data will be collected from approximately 500 adults per test language. The final instruments will be administered to a minimum of 2,250 individuals per test language.

Intended Outcomes

LAMP intends to provide

- (1) adult literacy and numeracy skill data at the national level (and at sub-national levels if desirable and feasible);
- (2) the ability to benchmark literacy progress cross-nationally particularly with countries at similar levels of development;
- (3) more detailed data than separating populations into literate versus illiterate by describing the percentage of the population at various levels of literacy; and
- (4) data trend analysis of literacy development. The main uses of the data are to help countries make evidence-based policy decisions based on the LAMP data and develop capabilities within countries to develop their own assessments.

Change of Strategy and Leadership in LAMP at UIS

In mid-May, 2007, a new LAMP Head was appointed by UIS. He and his staff reviewed the status of LAMP, started communicating with the participating country managers regularly including through a monthly memorandum in three languages, English, French, and Spanish. In his September 2007 communication, the LAMP Head informed the country managers that “LAMP implementation involves a huge number of complex tasks both at national and international level.” This translates into a significant workload that goes well beyond the existing and foreseeable UIS operational capabilities. This situation has led to a review of the LAMP initial strategy. This review took place in February 2007 and concluded that:

- i. *UIS should concentrate its efforts in completing the pilot phase as soon as possible. The desirable scenario requires field-test data collection to be completed in pilot countries by 2007, and main assessments to be conducted in 2008. Results from the pilot phase would be released around mid-2009. This would mean that the pilot phase would take 6 years since originally launched.*
- ii. *UIS should design a strategy to proceed with LAMP implementation on a larger scale. This strategy should take into consideration the following elements:*
 1. *UIS major mandates and essential tasks pertaining to LAMP (quality assurance and overall leadership)*
 2. *UIS operational capabilities*
 3. *UNESCO field structure and UIS distributed field staff (Regional Advisors and their units)*
 4. *Linguistic and cultural communalities*
 5. *Existing expertise including the experience of pilot countries, resulting in the development of major partners at global and regional level.”*

These considerations and the review led to a re-conceptualization of LAMP to be comprised of the following components:

LAMP-S (standard version)	Oriented towards producing evidence on literacy skills at national level including only an essential set of possible breakdowns (gender, age 15-24, 24+). The production of a limited set of breakdowns would make LAMP more affordable.
LAMP-X (expanded version)	Oriented towards producing evidence on literacy skills at a national level including all the breakdowns a country can be interested in. This option could be far more expensive.
LAMP-E (estimations)	A LAMP estimation model in order to generate estimates of the distribution of literacy skills for subgroups of population when LAMP-S is implemented.
LAMP-P (projections)	A mechanism to project data for periods in between LAMP implementations (within a five to ten-year cycle) based on a particular model that could combine historical data (including LAMP measures), current information on changes in the education system and demographic changes.
LAMP-HHS (household survey module)	A regular and systematic collection of literacy-related data in order to have evidence on literacy on a more frequent basis (LAMP measurements should be implemented only every five to ten years) and to feed into LAMP-E and LAMP-P.
LAMP-SB (school-based)	Measures of learning achievement at the end of primary education, using the same conceptual framework and scales in place for LAMP.
Linking these elements:	UIS will test as part of the pilot study if LAMP-S would yield enough information for having reasonable estimations as described above (LAMP-E paragraph). If that is the case, LAMP-S can be the logical standard, and LAMP-HHS would be the best way of promoting a regular collection that would provide the data required for the estimations (LAMP-E) and projections (LAMP-P).

In addition to re-conceptualizing the core components and processes of LAMP, the new strategy identified participation in LAMP as a capacity building effort as a priority. This strategy involved three key principles:

1. *Looking at country teams as knowledgeable partners, and not as “good” or “bad” students that are or are not to follow our instructions or, even worse, as “believers” in a particular methodological approach.*
2. *Relying on national and regional expertise (regarding them as knowledgeable partners) and not only on “the usual suspects” for the major elements of LAMP implementation.*

3. *Paying more attention to countries' criticisms and observations on LAMP implementation and development, as previously explained in the general section of this strategy note.*

This approach indicates sensitivity of UIS to develop respectful and collaborative relationships with the participating countries and is indeed expected to lead to self-sustainable capabilities for the countries.

II. Evaluation Methodology

Document review and analysis

The evaluation team was provided with documentation and reports on LAMP electronically. Reviews of these documents allowed us to examine the development of LAMP as a capacity building effort and as an assessment system for collecting policy relevant data; thus, assess the impact of UNESCO's work in the area of monitoring of learning achievements. The documents included large electronic folders of:

- Programme Planning
- Assessment Framework
- Assessment Instruments
- Background Questionnaires
- Translation and Adaptation
- Operational Manuals
- Workshop documentation and materials

In addition, the LAMP Head at UIS provided us with his notes on LAMP and his monthly communication with the country teams for August, 2007 through January, 2008. Appendix A provides the list of all the documents reviewed.

Meeting with key informants

The Lead Evaluator Dr. Kadriye Ercikan met with Mr. Cesar Guadalupe and Dr. Hendrik van der Pol at the UIS office in Montreal on December 4, 2007. This meeting was very helpful in obtaining most of the documents we reviewed for the evaluation. An additional meeting was held on March 7, 2008 at the UIS Montreal office with UIS LAMP staff Mr. Cesar Guadalupe, Dr. Hendrik van der Pol, Mr. Manuel Cardoso, and the IOS staff Dr. Amir Piric and Mr. Peter Maertens. This meeting was used to clarify issues and processes and obtain an update on LAMP.

Questionnaires for Country Managers

A questionnaire was developed to gather information about country managers' experiences with and perceptions of LAMP. The questionnaire addressed the following issues:

- stage of field testing the country is in
- successes of the country's participation in LAMP
- challenges in relation to country's participation in LAMP
- the extent to which LAMP is expected to meet the country's needs
- cultural appropriateness of LAMP instruments
- development process of the component skills portion of the LAMP assessments

The questionnaire was sent out to the five participating country managers. We have received responses from three of the countries. The information gained from these responses are used for different aspects of the evaluation and incorporated into the appropriate sections of the

report.

III. Evaluation Findings

Technical soundness of UNESCO's outputs

1. Have the aims of the study been stated clearly and are they relevant?

The three main aims of LAMP are listed as follows:

- 1. To develop a methodology for providing quality data on the distribution of the literacy skills of adults and young people in developing countries;*
- 2. To obtain high-quality literacy data in participating countries and to promote its effective use in formulating national policy, in monitoring and in designing appropriate programme interventions to improve literacy levels;*
- 3. To build national capacities in the measurement of literacy, to develop and use LAMP data and methodologies.*

These objectives are intended to improve education statistics and provide methodological tools and capabilities to the participating countries. They are clearly stated and relevant for the countries participating in the program. The International Planning Report provides further elaboration and guidance on how to achieve these goals. Even though the three broad objectives above do not imply cross-national comparability of literacy data from LAMP the following statement clearly indicates that it is included as part of the goals “*The methodology is designed to achieve comparability both within and among LAMP participating countries as well as IALS/ALL countries.*” The two goals: (1) providing useful and high quality data at the national level and (2) comparability of results among participating countries are important and necessary for an international survey of adult literacy. The quality of the data is critical to the extent to which data can be used meaningfully to inform policy within countries. International comparability of results, particularly within regions, allows countries within a region learning from each other's successes and challenges in their education systems.

Target Population

2. Was the defined target population appropriate? Were the target populations comparable?

The target population is 15 year-olds and older adults. This target population is consistent with the first goal of LAMP which is to provide data on the literacy skills of adults and young people. In most countries, compulsory education is up to age 15 or 16. Therefore, 15 years-old is an age at which many young people may be leaving formal education contexts and possibly entering the work force. This target population allows for the measurement of literacy skills of individuals after completion of compulsory education. The target population is fixed and therefore is comparable for all participating countries.

Tests

3.a. What was the purpose of the tests?

The measurement in LAMP consists of literacy scales on prose, document, and numeracy. In addition, it will collect information on “component skills.” The literacy scales are defined in the International Planning Report as follows:

Prose literacy – the knowledge and skills needed to understand and use information from texts that include editorials, news stories, poems, and fiction; for example, finding a piece of information in a newspaper article, interpreting instructions from a health brochure, inferring a theme from a poem, or contrasting views expressed in an editorial.

Document literacy – the knowledge and skills required to locate and use information contained in materials that include job applications, transportation schedules, maps, tables and graphs; for example, locating a particular intersection on a street map, using a schedule to choose the appropriate bus, or entering information on an application form.

Numeracy – the knowledge and skills required to apply arithmetic operations, either alone or sequentially, using numbers embedded in printed materials; for example, totaling the cost of two or more purchases, working with fractions and decimals, completing an order form, or determining the gain or loss on an investment.

LAMP will assess five components in English and the countries conducting measurements in other languages will develop their language specific components skills tests that are indicative of reading acquisition levels and integrated reading ability including:

- *alphanumeric perceptual knowledge and familiarity;*
- *word recognition;*
- *word knowledge (vocabulary);*
- *sentence processing; and*
- *passage fluency.*

Countries using LAMP to measure literacy in languages other than English will have to determine the building blocks for reading in their language and then create suitable items to measure them as part of component skills measurement.

3.b. How were the test frameworks developed?

The starting off point for LAMP is development of assessments that are comparable to International Adult Literacy Survey (IALS) so that the countries participating in LAMP could be compared to the OECD countries which have participated in IALS. Therefore, the test frameworks for the core parts of the measurements are based on those used for IALS. This approach to test framework development serves the purpose of comparability with IALS results well. However, the appropriateness of such frameworks for measuring literacy in the participating countries needs to be examined closely. Professional guidelines for testing (AERA, APA, & NCME, 1999) and assessment design principles based on more recent assessment research (Ercikan, 2005) recommend that assessment design starts with a well-thought-out and well-researched construct definition. This construct in turn determines the test framework. In the case of LAMP, definition of adult literacy in the participating countries and cultures should form the basis of the framework development. However, this did not take place for the core part of LAMP which was based on IALS items. Review of test frameworks developed for a set of other countries (OECD countries) by the participating countries has limited value in establishing cultural appropriateness and validity. This review process needs to be preceded by appropriate construct definition, framework and test development.

3.c. How were the test items written?

In August, 2004 UIS conducted a workshop where item writing according to the test frameworks was a key component. The goal was to train countries to use frameworks to write country-specific items that can be used in the pilot assessment in their countries. In addition to the IALS items, some participating countries created items according to a set of criteria outlining the type of items and the characteristics of the items. Some of these items were used as common items for all countries and others were used by specific countries only.

As a result, in addition to the IALS items, there were a set of common LAMP items and one of the countries (Morocco) developed their country specific items. Countries may choose to add nationally specific questions to the Background Questionnaire as well as to the literacy and numeracy tests to ensure that they obtain information on individual and family characteristics that may be needed to address unique national issues. All nationally specific items were submitted to the UIS for approval. These items were submitted in one of the reference languages (English, French and Spanish) and in the original language if it is not one of the reference languages. Countries were responsible for translations from their country language(s) to one of the reference languages, if necessary. The pilot testing (including pre-test and field testing) efforts are intended to provide statistical data about the properties of these items.

Components skills part of LAMP focuses on the very basic levels of literacy

One of the participating country managers reported that the participating countries were not invited to participate in the development neither of the structure nor the content of the original source version of the component skills tests. Participating countries were invited to adapt the source version to their national language(s) that would be used in the LAMP survey. Countries did have the possibility of proposing the elimination of sections that are not comparable with LAMP's national administration language(s). Similarly, the countries were authorized to introduce a slight modification in the section related to word decoding by introducing 'vowelized' words. One country reported that concerns were raised about the length of word lists and phrases and appropriateness of some phrases. These issues were not addressed to the country's satisfaction.

3.d. How valid were the tests?

The following kinds of validity evidence are needed for judging appropriateness of LAMP tests and meaningfulness of interpretations and effectiveness in guiding policy decisions:

- Evidence for culturally appropriate definitions of literacy and numeracy as target constructs for measurement;
- Evidence that tests were designed according to these definitions of target constructs;
- Evidence that there is a good match between these target constructs and tests;
- Measurement accuracy and reliability evidence.

The field test data can be used to address the above validity issues to some degree. However, the field test data is not sufficient to address these issues comprehensively and further validity studies, such as sensitivity review of test items for cultural bias, content review by an independent international group of literacy experts are needed to enhance the validity evidence for the LAMP score interpretations.

Procedures

The field testing is meant to not only gather information about the appropriateness of test items but the test procedures as well. Countries are field testing implementation of the following four instruments:

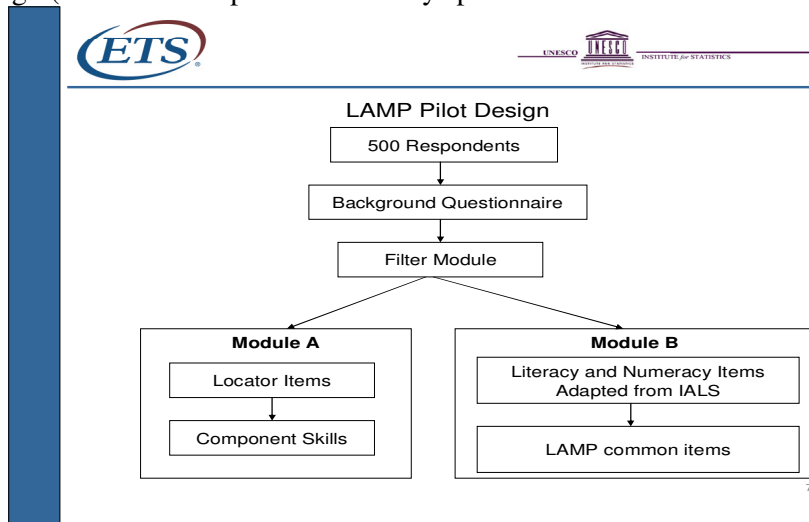
- (1) Filter test - designed to divide adults into a high and low skill levels;
- (2) Module A - designed to assess lower skills levels;
- (3) Module B - designed to assess higher skill levels; and
- (4) Background Questionnaire.

The administration procedures are described in the International Planning Report as follows:

“The interview begins by an interviewer completing a household roster designed to

identify respondents that are eligible for interview. One household member is then selected at random to complete the background questionnaire and the filter test – a test made up of relatively simple reading and numeracy tasks. Low skilled respondents who fail the filter test are asked to complete the locator test – a short test with sufficient low difficulty prose, document and numeracy tasks to place them on the LAMP proficiency scales, and the components test in the target language.”

The following diagram from the August 2004 workshop slides describes the LAMP pilot design (modified to replace the country specific item with LAMP common items).



As indicated above, the Filter Booklet contains easy reading and numeracy items. However, these test items do require participants to be able to read and write at least at a basic level. Very low literacy participants are expected to fail this test by not being able to answer any of the questions. These participants will then be guided to take Module A Locator Booklet and then the Module Components FlipBook. The Module Components Flipbook starts off with very basic literacy skills such as letter and number recognition. The order of this test administration may create some problems. Administering tests to individuals who may have minimal literacy levels may frustrate them and may provide examiners with little to no information and therefore waste time and resources. The UIS is gathering evidence about this potential problem in field testing and plans to address it accordingly.

Content and Construct Related Validity

Content and construct related validity is at the core of appropriateness and meaningfulness of measurement. For LAMP, these aspects of validity of interpretation of scores refer to the degree to which the measurement tools capture literacy skills appropriately, comprehensively, and accurately. These are determined by the degree to which the definition of literacy as measured by the tests reflects literacy in different cultures, languages, and countries appropriately, that tests capture all aspects of the literacy in these countries and that there is sufficient measurement accuracy. In the LAMP development process, the countries reviewed the test content. Even though the test content was approved by countries through this review process, some countries expressed their cultural appropriateness concerns were not addressed adequately. The field test data will provide information about the measurement accuracy provided by the tests. Country reviews combined with the field test data will form the key content and construct related

validity evidence. Further reviews by experts in literacy in each country are needed to enhance this validity evidence. One of the issues that need to be examined closely in each country is to what extent the construct definitions for literacy in English, which the core of the tests are based on, and overlaps with definition of literacy in other languages such as Arabic, French, and Spanish and many other languages and cultures the LAMP instruments are adapted to.

Linguistic and cultural appropriateness

The IALS items constitute approximately 60% - 67% of the tests. These items were created for OECD countries and the appropriateness of the contexts they are presented in needs to be examined for countries participating in LAMP. There is tremendous research in construct comparability research in measurement that indicates that familiarity with the contexts in which items are presented does affect performance levels of examinees even though previous knowledge about such contexts is not necessary for answering the item correctly (Ercikan, 2005). Currently, appropriateness of some of the contexts the test items are presented in is questionable. For example, some of these items make reference to regular and sale prices of items, gas gauge, and Italian Mozart Association. In order to increase the degree of validity, comparability, and meaningfulness of test items, there needs to be a cultural sensitivity review of all test items, including the country specific ones by content area, linguistic, and cultural experts.

UIS provides detailed guidelines to help translators deal with the challenges during the process of translating and adapting the LAMP items. Most of the examples that are provided in the document are useful in terms of providing some guidance to translators in their task. The initial versions of guidelines are being revised by the current LAMP staff at UIS to prevent problems by being too restrictive. For example, in these initial guidelines it is stated that *“It is also important that the answers in the texts appear in the same locations as the English version. For example, if the answer to a question in the English version is located in the first sentence of the last paragraph in the second column, then this should also hold true in your adapted version (p.3).”* As stated in this example, having instructions that are too narrow may not work for languages in which sentence or grammatical construction of the language does not allow for answers to be placed in the same location. LAMP staff at UIS are aware of the limitations of these guidelines may create and are in the process of revising them.

In the adaptation rules for money/currency, the rule under option 3 states that *“if numerical values must be changed to retain the realism of an item, they can only be multiplied or divided by powers of 10 (i.e., by 10, 100, 1000, etc). This restriction aims to keep the cognitive demands of the item (such as the nature of the mathematical steps and mental operations) similar in all countries (p.6).”* From a mathematical point of view, multiplying or dividing by powers may significantly change the cognitive demands required by the question. Doing calculations with four digit numbers (\$8,000) does not require the same degree of ability as using two digit (\$80) numbers in mathematical calculations. The UIS LAMP staff are aware of these problems as well and will be revising these items for the final versions of tests. This revision can be expanded to include issues like non-decimal systems used in some languages/cultures.

Although the adaptation of names and places in the stimuli are not allowed, creating original items that are less culture-specific would be beneficial. For example, Question 16 “Mozart Letter” represents a highly culture-specific item (p.18). Similarly, Question 6 “Cashmere cost” may be difficult to adapt in other cultures. Other examples include

scrambled eggs, tide tables (not all cultures may know about tides), laminated number line posters, fireworks, and skateboards. These issues have been brought to the UIS's attention by the evaluation team and the test items are going through a review to minimize such potential cultural biases.

Comments by country managers confirmed the existence of culturally inappropriate test items in the pilot versions of LAMP tests. One of the countries expressed that the cultural appropriateness was not taken into account by the UIS in their initial action plans. It took strong insistence from this particular country to come to an agreement with UIS to take cultural appropriateness into account.

Pre-testing

Pre-tests were conducted prior to the field testing in LAMP. All participating countries are expected to conduct a pre-test. Countries are instructed to pre-test assessment items and background questions on a small number of (30-40) individuals. This step is important to identify questions that may be culturally sensitive or that lack clarity in instruction and to check that the questions yield the desired information. Volunteers are asked how they reached the answer and to comment on whether the questions are clear and culturally sensitive. This is an important step in gathering validity evidence for appropriateness of test items for the countries. Most international assessments such as TIMSS and PISA do not conduct pre-tests or field tests in all of the participating countries and this is an important advancement in LAMP that needs to be acknowledged.

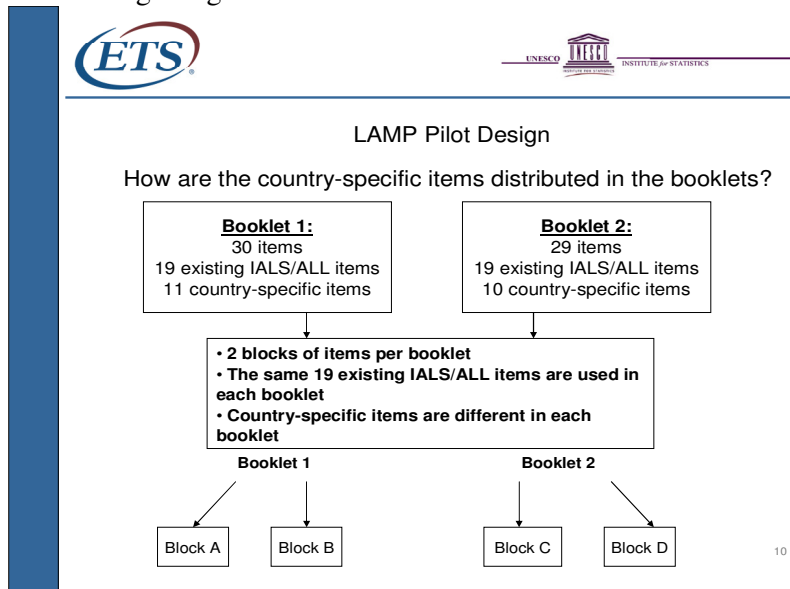
In addition to the pre-test, all participating countries are required to conduct a field test. This field test involves the administration of the entire battery of survey instruments to a carefully selected sample of roughly 500 adults in each test language. Field test objectives can be grouped as follows:

- 1) Psychometrically validate the items to be included in the LAMP assessment;
- 2) Collect sufficient data to establish the Filter 'pass' criterion for the LAMP main survey, and to examine its effect on the flow of respondents to Module A and Module B;
- 3) Ascertain the effectiveness and appropriateness of all field procedures;
- 4) Evaluate the sample design and sampling plan;
- 5) Ascertain the appropriateness and effectiveness of Interviewer training;
- 6) Assess the effectiveness of Scorer training;
- 7) Assess the scoring operation in terms of Scorer productivity and the quality of the score data;
- 8) Assess operational considerations and confirm key design and costing assumptions.

The current field test design, documentation, and implementation procedures are expected to meet these goals. The field test data will provide information about two key psychometric properties of test items: Their *difficulty* and *discrimination* levels. Both of these psychometric properties of items are key to evaluating the quality and appropriateness of the test items. The field test data can also be used to compare these properties of test items across countries and therefore help establish some level of cultural validity. The second objective of the field testing can be addressed by examining performance levels across different components of LAMP.

The samples used in the field testing are not nationally representative samples. Therefore, countries are not engaged in complex representative sample designs in this phase of LAMP and therefore, goal number 4 may be addressed in a limited way. However, the

field testing procedures can provide useful information about the countries' capabilities in sampling design and implementation. To address the rest of the field test objectives, UIS has established procedures to document the operations and procedures in field testing in the participating countries. The following slide from the UIS August 2004 workshop describes the field testing design.



Questionnaire

Do the questions provide complete coverage for the study's overall aims?

The Background Questionnaire (BQ) plays a key role in achieving the second goal of LAMP: that is providing data for making evidence-based national policy decisions. The background questions are intended to collect information in the following areas:

- demographics (e.g., gender, age, and migration pattern);
- respondent's home;
- education (e.g., educational attainment and education history);
- job-related information (e.g., employment status and history);
- literacy-related information (e.g., use of literacy skills at work and outside of work);
- language use; and
- other social information (e.g., health-related questions).

The questions in the BQ cover five major areas that are of interest to policy-makers: *population distribution of literacy skills, antecedents of literacy skills, relationship of literacy skills to social environment, relationship of literacy to other proxy variables and monitoring trends in literacy skills.* These are all relevant to key policy issues related to distribution of literacy and examining possible factors that might be associated with improving literacy. The questions provide complete coverage of issues and factors related to the study's overall aims. In addition, countries can identify and include questionnaire questions that may be relevant to their specific policy questions. The BQ is under review by the UIS LAMP staff for further improvements.

How were the background questionnaires developed?

In the August 2004 workshop, a set of background questions were developed, reviewed, and discussed by the UIS. In addition, country-specific questions were developed by

individual countries. Allowing countries to add country specific questions is very important for tailoring the BQ for country needs. Some concerns were expressed by one of the participating countries about UIS's inflexibility in tailoring the BQ to country specific needs.

We do not have data for the following terms of reference questions for LAMP due to its current development phase.

- Were the tests reliabilities at an acceptable level? (data not available yet)
- Was the item-writing process convincing? (we do not have documentation for this)
- Data (not yet completed)
- What kinds of analyses were carried out on the trial data? (analyses have not been conducted)
- How reliable were the data? (analyses have not been conducted)
- Was the data collection well conducted? (not yet completed)
- Sampling (not yet completed)

Relevance and adequacy of organizational arrangements

Between the years of 2003-2006, approximately \$2.6 million has been spent for the LAMP implementation. The information in the LAMP Finance 2003-2006 document shows that UIS has funded 94.9% of the total amount devoted to LAMP in these years. This level of financial contribution is a clear evidence of UIS's involvement and commitment to LAMP. The UIS LAMP staff developed reports, guidelines, and organized workshops that addressed key processes. These organizational support elements were detailed, clear, and well written. An International Planning report was prepared in 2004. This report covered:

- Background
- The Broad Objectives of LAMP
- How LAMP Results Can Be Used
- The LAMP Conceptual Framework
- The LAMP Design, Understandings of Country participation in LAMP
- Information About the Data Collection Tools
- Analysis, Dissemination, and Communication of Results

In addition, there were detailed manuals on translation, adaptation, and implementation procedures. Four capacity building workshops were conducted during the period of 2004 – 2006, the first workshop, Montreal 2004, focused on the following goals:

- Present a brief overview of LAMP.
- Discuss National Planning Reports as needed.
- Discuss the frameworks that underpin Prose and Document Literacy, Numeracy, and Reading Components.
- Train participants to use these frameworks to write country-specific items that can be used in the pilot assessment in their country.
- Train participants to adapt and translate the existing items for use in the pilot assessment in their country.
- Train participants to develop reading components measures that make sense for their language of assessment.

Marrakesh 2005, Nairobi 2006, Cairo 2006 workshops focused on data collection tools, items, and field testing procedures, respectively. These workshops correspond closely with the first three phases of LAMP.

Is the level adequate to produce and sustain high quality outputs?

Based on their experiences in the field testing phase the participating countries provided the following suggestions for organizational arrangements in the upcoming phases. The statements below are summaries of comments made by the three country participants:

1. Reinforcement of capacity of LAMP teams to lead them to become more autonomous, provide guidance and support in obtaining further funding for LAMP and help coordinate communication and collaboration among countries within regions.

2. UIS should accumulate better knowledge about each national team (their strong points, their weak points, their problems). A diagnosis of the situation at each country level is needed to have this kind of information. To date, UIS has treated national teams in black boxes supposing that everything works well.

3. UIS should give participating countries more room to express their needs and requirements and in the implementation of these needs and requirements in the action plans. The set up of a scientific committee including a representative from each participating country could respond to this worry and put in place a dynamic of participation that breaks away this 'top-down' approach followed thus far.

4. UIS should give more importance to capitalize upon and value successful experiences of participating countries and encourage effective cooperation between these countries (South-South cooperation). Taking into account that the countries are not in the same stage of implementation, that synergies can be put in place between these countries, so that each country can benefit from the experiences of others. This presents the advantage of leading an economy of time and means of mutual national capacity reinforcement of participating countries.

5. Strengthen the technical advising process in a direct way that would enable countries in full implementation of LAMP.

These suggestions need to be followed up on by the UIS to engage country participants in developing plans for the organizational arrangements.

To what extent do UNESCO outputs meet the technical needs and requirements of programme beneficiaries?

It is too early to evaluate the extent to which UNESCO outputs meet the technical needs and requirements of programme beneficiaries because the main instruments have not been administered and the results have not been published.

What are programme beneficiaries' needs and requirements with regard to monitoring and assessing learning achievement?

Countries reported that the LAMP will meet an important need for them that is direct measurement of literacy levels which are expected to be more accurate reflections of literacy levels than more typically used indirect measures of literacy based on surveys asking individuals about their literacy levels. The results from LAMP are expected to be used to inform important policy decisions to improve literacy levels in the participating countries. The countries also report a need for content review of some instruments and particularly the BQ and the component skills module in an effort to simplify them for the countries.

Effectiveness of the quality control and quality assurance procedures

Our evaluation team has reviewed appropriate documents to examine quality control and assurance procedures concerning the following aspects of assessments in LAMP:

- (1) test development;
- (2) appropriateness of test content in the participating countries;
- (3) comparability of versions of tests across countries;
- (4) administration procedures;
- (5) scoring procedures;
- (6) sampling design.

The LAMP has extensive procedures for documenting all aspects of the processes. This documentation can play a critical role in establishing quality control and quality assurance procedures. Since the field testing is ongoing this documentation process has not been completed and the adequacy of these procedures cannot be fully evaluated. The first three issues identified above were addressed under technical soundness of the evaluation report. These issues were also raised under validity. The final three issues (four to six) are discussed below.

Administration Procedures

LAMP Standards and Guidelines require that each country employs a sufficient number of interviewers so that the maximum assignment size for an interviewer is 40 completed assessments per month during the data collection period. The LAMP interviewers are expected to do 6 hours of Home Study training and three days of in-class training. The interviewers are expected to be supervised by a country team member who has been trained by the UIS as well as observed by the UIS staff. In Niger, a member of another country LAMP team was asked to be present at the training on behalf of the UIS staff in order to foster South-South cooperation. The interviewer supervisors' duties are described as follows:

There will be at least one interviewer supervisor for every ten interviewers. The interviewer supervisor's duties will include the following:

- a) Interviewer training*
- b) Implementation of an observation program so that each interviewer is observed at least twice during the data collection period, once as early as practically possible during the first half of the data collection period and once during the latter half of the data collection period,*
- c) On-going monitoring of each interviewer's work*
- d) At least one meeting per month with each interviewer to discuss progress and/or problems*
- e) Non-response follow-up*
- f) Visit verification and sample selection verification.*

The UIS staff were present in all of the field-testing countries (except in the Palestinian Autonomous Territories (PAT)) to provide assistance. The staff could not be present in PAT for security reasons and measures set by the United Nations' system.

Scoring

The LAMP items are scored manually. The manual scoring needs to be subjected to accuracy checks and verification procedures in order to ascertain score comparability within and across countries. The same scoring rubrics are expected to be used in all sites. The

International Planning Report states that “*the reliability of scoring will be checked by having a second scorer re-score all field test booklets and that the goal of the rescore is to improve scoring to the point where inter-scorer agreement approaches 100%.*” It is important to note, however, that in this process the guidelines indicate that it is not a requirement to reconcile individual cases. If the goal is not to reconcile individual cases, then one can ask why the re-scoring is being conducted. It is also not clear whether re-scoring will be done by an independent set of trained scorers. UIS is aware of the vagueness in the description of the scoring procedures and establishment of reliability and the staff are working on both revising the procedures and the description of procedures.

The current documentation indicates that “as an additional quality control check, a random sample of roughly 10% of test booklets will be selected and submitted to independent scoring.” This statement is in conflict with the statements about double-scoring above. If ALL the booklets are being re-scored, there is no need for a sample of 10% to be re-scored. It should also be noted that re-scoring is of no value for improving quality of scoring if the information from re-scoring is not used to verify the accuracy of scores from the initial scoring.

Sampling

The countries are expected to design the sampling framework and implement the sampling design. The LAMP has many appropriate documentation procedures in sampling design and verification and the central UIS office is planning to verify and validate the implementation of the sample design. The field testing does not require a representative sample of the country. However, the field test sample needs to cover the complete range of literacy skills in the country in order to examine the psychometric and cultural appropriateness of the measurement tools.

Is a monitoring and evaluation system in place to enable Headquarters to monitor the progress of the programme? How effectively has the monitoring and evaluation system functioned?

UNESCO Headquarters conducts regular evaluation of its programmes. This is the first evaluation of the LAMP. A key recommendation we would like to make is implementation of an ongoing external review process as a monitoring and evaluation system from UNESCO. This can be in the form of an international technical advisory board consisting of psychometricians, education policy experts, and statisticians with international reputation. This board needs to have an arms-length relationship with the project and review activities, processes, and products and provide a report to the UNESCO Headquarters on a yearly basis. In almost all international assessments we are familiar with, such committees and boards are an integral part of quality assurance and collaborative improvement of processes. UNESCO’s support of such a committee of experts will go a long way in assuring quality and revision of procedures according to up to date professional standards.

Results

What are the expected results of LAMP?

The expected results are threefold: (1) quality data on adult literacy in the participating countries; (2) use of such data in meaningful ways to inform important policy decisions; (3) capacity building in participating countries. Some steps have already been taken towards these goals by field testing procedures and tests in five countries. The outcome of the field testing and the appropriate use of data from field testing will provide some indication of how

successful this step has been and is expected to the achievement of the expected results.

What results have been achieved?

The three countries the evaluation team surveyed reported the following achievements:

- National capacity building in the area of direct measurement of literacy levels
 - item development,
 - conducting the data collection
 - the marking and data capturing;
 - effective operationalization of the partnership between several participants while data collection is taking place
- Identification of a national project leader and establishment of a LAMP country team;
- Adaptation of the LAMP guides (manuals) and items to the country languages;
- Translations were verified and corrected;
- Tests and other data collection instruments, and manuals were produced in country languages in some cases in five languages;
- Trainers, interviewers, and controllers were trained;
- Obtaining permission and help in logistics from country officials to conduct LAMP;
- Contextualizing survey instruments;
- Contribution by countries to key aspects of LAMP such as awareness and development of procedures to cultural appropriateness issues,

What factors have contributed to and/or prevented the achievement of expected results?

The expected results in this phase of LAMP are the completion of field testing. The major part of this has been completed in three countries. The following challenges were identified by the countries we surveyed:

- Unrealistic representation of LAMP to policy decision-makers as a project that could be conducted easily. This led national policy decision-makers to build short term expectations and not make plans for all stages of the project;
- The cumulative demands in putting this program in place have created a kind of frustration within the national decision-makers;
- Changes to the project leader at the UIS level has contributed to delays and to building this feeling of frustration. Further, on the scientific level, the UIS did not succeed to provide the countries complete picture of all phases of the project;
- The lack of communication on the project objectives has made room for the development of misunderstandings and misinterpretations among some participants which contributed to reinforcing this feeling of frustration;
- Lack of planning to cover financial aspects of the project. In one country, this led to trying to continue to work on a project without a budget.
- The realization of the LAMP project requires the effective and simultaneous involvement of several participants. Some countries faced the difficulty of operationalizing the partnership between the different participants involved.

These challenges contributed to delays in the field testing and frustration in some of the participating countries.

Risks

The following are some of the key potential risks that might jeopardize the outcomes of the initiatives and achievement of the goals of LAMP:

- (a) Appropriateness of tests and measurement of literacy and numeracy skills in each of the participating countries;

- (b) Accurate measurement and representation of skills and competencies in some countries;
- (c) Comparability of data and scores across countries;
- (d) Differences in gender roles in some countries, such as more men working outside of the home, may influence availability and participation of gender groups equally;

These potential risks need to be taken into account in the next steps of LAMP in order to minimize the effects of these potential risks as well as design validation studies to gather empirical evidence that address these risks.

Capacity Building Efforts

The LAMP has the potential of having a great impact on literacy policy and outcomes in the participating countries. The first is through creation of meaningful and useful data about literacy distributions. Second is through professional development and other enabling capabilities to transfer tools, skills, and knowledge to other activities in the country. The LAMP is in its initial stages and the data have not been created. Therefore, an evaluation of the data is not possible at this point. The following capacity building activities have been conducted for LAMP which are expected to not only enable participation in LAMP, but also develop capabilities for the country to conduct similar other activities:

- Provide participating countries with the LAMP framework and underlying assessment methodology;
- Help develop documentation of processes and outcomes of LAMP. Documentation includes the following: National Planning Report Template; International Planning Report; Sampling Manual; Survey Procedures Manual; Guidelines to Achieve Maximum Impact;
- Provide instruments, training programmes, and associated training materials
 - Instruments include the common items/questions on a Background Questionnaire and on the Assessment Instruments complete with instructions on their administration.
 - Training programmes include: Item Creation for Nationally Specific Items; Adaptation and Translation Guidelines for Common Items; Sample Design and Selection; Interviewer and Scorer Training; Data Capture and Analysis;
- Guidance on dissemination and communication strategies
- In addition, the UIS will provide participating countries with training, if required, in quality control procedures to be used in survey development and implementation that maximizes the quality of the results

Appendix A: List of Documents Used in the LAMP Evaluation Electronic Folders

1. Program Planning
 - Literacy Assessment and Monitoring Programme (LAMP) International Planning Report (Montreal, 2004)
 - Memorandum of Understanding between UIS and countries
 - National Planning Report Literacy Assessment and Monitoring Programme (LAMP)
2. Costing Template
 - Schedule
3. Assessment Framework
 - Framework for the Assessment of Reading Component Skills* (John P. Sabatini, Educational Testing Service)
 - The International Adult Literacy Survey (IALS): Understanding What Was Measured (Irwin Kirsch)
 - Adult numeracy and its assessment in the ALL survey: A conceptual framework and pilot results
4. Assessment Instruments (7 files)
5. Background Questionnaire (4 files)
6. Translation Adaptation
 - Literacy Assessment and Monitoring Programme (LAMP) Translation and adaptation procedures
 - Literacy Assessment and Monitoring Programme LAMP Adaptation and Translation Guidelines for the Background Questionnaire Field Test Version (March, 2006)
 - Formatting, Translation, and Adaptation Guidelines for the Reading Components
 - Literacy Assessment and Monitoring Programme
 - LAMP Translation and Adaptation Guidelines
 - General Rules for Translation and Adaptation of LAMP Materials
7. Operational Manuals
 - Standards and Guidelines for the Design and Implementation of the Literacy Assessment and Monitoring Programme (LAMP) and 10 other documents this manual referred to
8. PowerPoint slides and documents for Montreal 2004, Marakkesh 2005, Nairobi 2006 and Cairo 2006 workshops.
9. Cesar Guadalupe memo on LAMP Financial Situation 2003-2006 (11th September 2007)
10. Cesar Guadalupe notes on LAMP 310707.doc
11. Cesar Guadalupe Letters to NPL for August 2007 through January 2008

Other References

- AERA, APA, NCME, (1999). *The Standards for Educational and Psychological Testing*. American Educational Research Association, Washington DC.
- Ercikan, K. (2005). Developments in assessment of student learning. In P. Winne & P. Alexander (Eds.), *Handbook of educational psychology* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Appendix B: The Questionnaires in English (also distributed in French and Spanish)

Questionnaire on UNESCO's support to LAMP

Date:

Country:

The purpose of this questionnaire is to gather data that will be used as part of the evaluation of the UIS-UNESCO's Technical assistance and capacity building efforts in relation to LAMP. Your responses to these questions will help the evaluation team in making informed recommendations about how to improve LAMP. We have identified 6 broad questions in relation to LAMP. You can insert your comments after the questions.

Please feel free to contact the Lead Evaluator Dr. Kadriye Ercikan via e-mail or telephone if you have any questions. If you prefer to speak on the telephone instead of responding to the survey questions in writing the evaluation team can arrange a time to call you on the telephone and arrange a telephone conversation with you.

Your responses are very important for identifying strategies for improvement in these initial stages of LAMP. Thank you very much for your participation.

Kadriye Ercikan
Associate Professor
University of British Columbia, Canada

Questionnaire

1. At what stage of field testing is your country in?
2. What are the successes of your country's participation in LAMP participation so far?
3. What are your country's challenges in relation to your country's participation in LAMP?
4. To what extent is LAMP expected to meet your country's needs in relation to measuring literacy in your country? Are there modifications that may be made to better suit your country's needs?
5. To what extent do you think the LAMP instruments were culturally appropriate for your country?
6. Please briefly describe the development process of the component skills portion of the LAMP assessments.
7. Additional Comments: Please add any comments that could help UNESCO provide LAMP with better support in the future.

Appendix C: Management Response

Recommendations	If not accepted, give reasons for rejection
<p>1. Balancing within country validity and cross country comparability</p> <p>The two goals of UIS are: (1) providing useful and high quality data at the national level; and (2) comparability of results among the participating countries. Achieving these goals simultaneously presents constraints that require careful balancing of within country validity and across country comparability. For example, within country priorities may require development of tests and instruments that are culture-specific, whereas cross-country comparability priorities require tests and instruments to be culture-free so that they may be used appropriately in all of the participating countries. We make specific recommendations regarding ascertaining and documenting this balance.</p>	<p>LAMP will ensure that all new items are included in the databank and assessed by expert judges. LAMP has already ensured that regional advisory groups include item development experts in their current work plan. From 2009, LAMP will budget for all new items to be submitted to experts for technical evaluation. LAMP will provide clear instructions for item development stated in guidelines. Analysis of field test results is under way and the revision of item development guidelines is to be complete by March 2009.</p> <p>LAMP will ensure that analytical procedures include methods for identifying potential cultural bias. Analytical procedures developed for the field test include observation of potential cultural effects- Already included in current work plan</p>
<p>2. Test framework development needs to be based on empirically defined constructs</p> <p>Assessment design principles recommend that assessment design starts with a well thought-out and well researched construct definition. This construct in turn determines the test framework. In the case of LAMP, definition of adult literacy in the participating countries and cultures should form the basis of the framework development.</p> <p>Most of the test items in the current LAMP tests are originally developed in English according to research and understanding of what literacy is in English. Test framework need to take into account empirical research on reading in the languages of the participating countries.</p>	<p>LAMP will revise the National Planning Report Template including a specific section and will identify of national experts for this task, to be completed by December 2008.</p> <p>LAMP will prepare guidelines for countries to produce a technical note addressing this issue as part of their design. Guidelines on discussing construct validation at country level will be completed by March 2009 and LAMP will continue to provide technical support to countries while designing the study.</p>

<p>3. Obtain and document key validity evidence</p> <p>Processes and studies need to be developed to gather and document a key set of validity evidence. The following are a list of key validity evidence required for LAMP:</p> <ul style="list-style-type: none"> • Evidence for culturally appropriate definitions of literacy and numeracy as target constructs for measurement; • Evidence that tests are designed according to these definitions of target constructs; • Evidence that there is a good match between these target constructs and tests; • Measurement accuracy and reliability evidence. <p>The documentation of validity evidence needs to be publicly available in the form of technical reports in order for users of LAMP to be informed about quality and properties of LAMP data.</p>	<p>LAMP will prepare a technical report, including an analysis of field test results which as been included in the currently workplan. An analysis of workshop documentation has been addressed in two international workshops and will be completed by March 2009.</p>
<p>4. Examine the appropriateness of the Filter Booklet as it is used in the field test</p> <p>The Filter Booklet contains easy reading and numeracy items. However, these test items do require participants to be able to read and write at least at a basic level. Very low literacy participants are expected to fail this test by not being able to answer any of the questions. The order of this test administration may create some frustration and problems among the participants. The data from the field test need to be used to examine the appropriateness of the Filter Booklet in guiding respondents to Module A and B.</p>	<p>LAMP will review and revise the overall flow of instrument administration as needed. An analysis of field test results and a proposal of a revised flow of instrument administration will be completed by September 2008. The validation with countries is planned for November 2008, and the final flow will be included in final documentation in March 2009.</p>
<p>5. Revise current tests</p> <p>Review of some of the test items by the evaluators revealed potential cultural bias. Current tests need to be revised to minimize cultural bias.</p>	<p>LAMP will have a final version of tests available by March 2009. An analysis of field test results and draft revisions will be conducted in September 2008 and the validation with countries is planned for November 2008.</p>
<p>6. Revise translation guidelines</p> <p>Having instructions that are too narrow may not work for languages in which sentence or grammatical construction of the language does not allow for answers to be placed in the same location. Although the adaptation of names and places in the stimuli are not allowed, creating original items that are less culture-specific would be beneficial. We recommend for LAMP to follow an adaptation approach rather than a translation approach for minimizing cultural bias.</p>	<p>LAMP will have a final version of translation and adaptation guidelines by November 2008.</p>
<p>7. Tailor Background Questionnaires to country needs</p> <p>Background questionnaires need to be developed collaboratively and tailored to country needs. This is an area where country appropriateness needs to be prioritized over cross- country comparability.</p>	<p>LAMP will have a final version of Background questionnaire and guidelines for adaptation by March 2009.</p>

<p>8. Follow up on specific country recommendations and needs The following recommendations came directly from the participating countries.</p>	<p>LAMP will identify experts including member of current national LAMP teams by Sep 08. Two advisory teams will be established in Bangkok and Santiago by December 2008. As part of its ongoing work, LAMP will maintain the regular e-mail communication with national teams stimulating a more active role by them. In addition, LAMP will identify potential gaps in national teams and make national authorities aware of them.</p>
<p>9. Refine scoring procedures In the scoring process, double scoring is conducted to check the accuracy of the scoring process. The scoring process needs to be refined so that discrepancies between scores from the two sets of scores (from two independent sets of scorers) be reconciled through an agreement or averaging process.</p>	<p>LAMP will ensure that scoring procedures are properly detailed in corresponding documentation, by conducting an analysis of field test results and revision the guidelines. A final version of guidelines will be made available by November 2008.</p>
<p>10. Establish an international technical advisory panel We recommend UNESCO to establish an ongoing external review process. This can be in the form of an international technical advisory board, consisting of psychometricians, education policy experts, and statisticians with international reputation. This board needs to have an arms-length relationship with LAMP and review activities, processes, and products and provide a report to UNESCO Headquarters on a yearly basis.</p>	<p>LAMP will establish an advisory committee, identifying potential members and defining the Terms of Reference by Sept 2008. The first meeting of the advisory committee will be held in December 2008.</p>
<p>11. A realistic determination of financial and other demands for participation in LAMP The cost, time frame, and processes involved in LAMP participation need to be identified and communicated to participating and potential participant countries. This goal has been challenging for UIS due to the several changes in leadership, processes and procedures it experienced in relation to LAMP. The goal is expected to be more achievable to refine and finalize procedures based on the field test data.</p>	<p>LAMP will revise the costing template. A new version of the costing template will be completed and documented by December 2008.</p>
<p>12. A consistent team at UIS Changes in LAMP leadership lead to delays and frustration among countries. In the upcoming phases of LAMP, a system and a team need to be in place that maintains stability in LAMP. It is essential for this team to include expertise in measurement, in particular expertise in international assessments and adult literacy. We recommend UIS/UNESCO to establish processes to minimize disruptions in LAMP by maintaining a key set of team members and expand the existing membership to include the two key areas of expertise identified above.</p>	<p>LAMP will strengthen coordination with adult literacy experts. A cooperation agreement will be established with UIL and Regional Technical advisory teams established in Bangkok and Santiago by December 2008. A New Statistical Analyst will be recruited by October 2008.</p>