

Action research: measuring literacy programme participants' learning outcomes

Results of the first phase (2011–2014)

Madina Bolly, Nicolas Jonas



United Nations
Educational, Scientific and
Cultural Organization



Maama

Maama Kpanji

Maama Nba

Maama

Ma

Ma

a a a a

aa aa aa a

a aa ma ma

ma ma ma ma

ma ma

ma ma

ma ma Ama





United Nations
Educational, Scientific and
Cultural Organization



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

Bureau à Dakar
Bureau régional
pour l'éducation
en Afrique



United Nations
Educational, Scientific and
Cultural Organization

Office in Abuja
Multisectoral Regional office for West Africa,
Benin, Cote d'Ivoire, Guinea, Ghana, Liberia,
Nigeria, Sierra Leone, Togo



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

**Swiss Agency for Development
and Cooperation SDC**

**Action research:
measuring literacy programme
participants' learning outcomes**

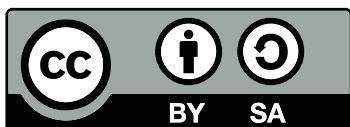
Results of the first phase (2011–2014)

Madina Bolly
Nicolas Jonas

UNESCO Institute for Lifelong Learning

Published in 2015 by
UNESCO Institute for Lifelong Learning (UIL)
Feldbrunnenstr. 58
20148 Hamburg
Germany

© UNESCO Institute for Lifelong Learning



This publication is available in Open Access under the Attribution-ShareAlike 3.0 IGO (CCBY-SA 3.0 IGO) licence (<http://creativecommons.org/licenses/by-sa/3.0/igo/>). By using the content of this publication, the users accept to be bound by the terms of use of the UNESCO Open Access Repository (<http://en.unesco.org/open-access/terms-use-ccbysa-en>).

The UNESCO Institute for Lifelong Learning (UIL) is a non-profit international institute of UNESCO. The Institute undertakes research, capacity-building, networking and publication on lifelong learning with a focus on adult and continuing education, literacy and non-formal basic education. Its publications are a valuable resource for educational researchers, planners, policymakers and practitioners.

While the programmes of the UNESCO Institute for Lifelong Learning (UIL) are established along the lines laid down by the General Conference of UNESCO, the publications of the Institute are issued under its sole responsibility. UNESCO is not responsible for their contents.

The points of view, selection of facts and opinions expressed are those of the authors and do not necessarily coincide with official positions of UNESCO or the UNESCO Institute for Lifelong Learning. The designations employed and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of UNESCO or the UNESCO Institute for Lifelong Learning concerning the legal status of any country or territory, or its authorities, or concerning the delimitations of the frontiers of any country or territory..

Authors: Madina Bolly and Nicolas Jonas

Cover design, graphic design and and typography: Sabine Siegfried

Cover photograph left: © Graham Holliday,

Cover photograph right: © compartuser

ISBN 978-92-820-1202-4

Foreword	4
Abbreviations and Acknowledgements	8
Acronyms	9
Chapter 1 Background and justification	11
Chapter 2 Action research on RAMAA	19
Chapter 3 Conceptualizing RAMAA: an innovative measurement of learning outcomes	23
Chapter 4 Data collection tools and quality assurance	33
Chapter 5 RAMAA's contribution to evidence-based advocacy and capacity development	39
Chapter 6 Results analysis	41
Chapter 7 Lessons learned	79
Chapter 8 Recommendations for RAMAA 2	83
Appendix: Member list of national teams	86
Bibliography	88

This document is the first publication resulting from the action research on measuring literacy programme participants' learning outcomes, known by the acronym RAMAA. The initial results of this action research, launched at the initiative of the UNESCO Institute for Lifelong Learning (UIL), stem from a collectively implemented programme in which members of national teams played a central role. Obtaining these results was also made possible thanks to the collaboration of the UNESCO Office in Dakar, the UNESCO Office in Abuja, the Working Group on Non Formal Education (WGNFE), the Association for the Development of Education in Africa (ADEA), external experts from the Organisation for Economic Co-operation and Development (OECD), the UNESCO Institute for Statistics (UIS), Statistics Canada, researchers from the Cadi Ayyad University in Marrakech and Hamburg University, as well as consultants. The Swiss Agency for Development and Cooperation (SDC) played a pivotal role by financially supporting this first phase of RAMAA.

Current debates on the post-2015 education agenda call for renewed efforts to improve the quality of education by taking a broader view of the challenge of continuum and life-long learning, but also by adopting an integrated and inclusive vision. In the declaration made by the 2015 World Forum for Education in Incheon, but also in that of the Sub-Saharan Africa Regional Ministerial Conference on Education Post-2015 held in Kigali, these commitments are ultimately to be reflected in the improvement of learning outcomes, whether in formal or non-formal education.

This first-generation action research on measuring literacy programme participants' learning outcomes is part of this movement and indeed at the forefront of the process, since it began in 2011. It has consequently sought to contribute to the effort to reach the 2015 goals of the Education for All initiative, but is also looking further ahead, since in addition it is working on the education objectives defined by the SDGs (sustainable development goals) that have been set for 2030.

We recognize that education is a fundamental human right and a public good. However, we are all aware that despite the efforts that have been made since 2000, the year of the Dakar commitments, massive illiteracy among young people and adults persists in sub-Saharan Africa. This population that has been left behind is a significant obstacle to implementing public policies for economic growth and social cohesion. Achieving the twin goals of education for peace and functional literacy, especially for this vulnerable population, is clearly an important strategy in Africa.

It is precisely this situation, but also the opacity and lack of visibility stemming from the absence of reliable data that prompted us to investigate the quality of the various literacy programmes and the level of instrumental and functional skills actually acquired by the participants.

The results of this action research, produced from a common methodological and contextual framework, but also from the analysis of three fundamental variables – gender, age and geographical diversity – will provide policymakers with targeted and relevant answers to better structure literacy programmes. In addition, they will act as catalysts to build fact-based advocacy cases. One further goal of RAMAA is to open up fields to applied scientific research.

We specifically opted for action research in the field of literacy measurement so as to initiate a learning dynamic and thus avoid placing the countries in situations of merely following instructions. Team members from countries participating in the first phase of RAMAA – Burkina Faso, Mali, Morocco, Niger and Senegal – were supported in building and producing their own knowledge in order to design and implement harmonized measurement, production and data-analysis tools. The objective of this participatory and co-construction-based approach is to gradually have the national teams take ownership of RAMAA and therefore develop their capacities to monitor and evaluate literacy programmes. Achieving sustainability and initiation into a culture of evaluation are the ultimate stages that will contribute to the integration of RAMAA in national education systems.

Any initiative to measure literacy accompanied by the implementation of an ‘action research project’ necessarily faces significant challenges, especially in a field as complex as that of literacy programmes. This was very much our experience, as explained in this document.

We have identified three main weak points that had an unprecedented impact on the quality of the data produced in this first phase of RAMAA: i) the difficulty in involving French-speaking or bilingual external experts to support the development of measurement tools in the field of non-formal education, ii) the contingencies of some participating countries and iii) the challenges faced by some countries to obtain funding to conduct RAMAA activities.

The results have certainly not lived up to our expectations, which is partly due to the fact that we did not succeed in developing harmonized measurement tools. However, we have learned from the first phase of RAMAA that the people who completed a literacy programme and are described as literate in the context of four participating countries, Morocco being an exception, still appear to have insufficient literacy skills. The absence of more comprehensive statistics upon which to evaluate the sustainability of outcomes is nonetheless an incentive for us to continue research into the conditions for the sustainable and transferable acquisition of instrumental skills. By focusing on the variables of gender and equality, the Morocco results show us that younger people (30 and under)

are more likely to reach the minimum level required. On the contrary, when results from all countries are compared, gender does not seem to play a leading role in sustaining outcomes. In this respect, the results from mixed-gender training centres in Niger are insightful: graduates from all-women and all-men centres scored better results than those from centres that train women and men together. Finally, with regard to the overall quality of the literacy programmes on offer, the results obtained by countries indicate that the centres offering the most hours of training are not the ones that perform the best or that the operator has no bearing on the quality of the literacy programme. These compelling conclusions all unconditionally argue in favour of further analysis of the content of literacy programmes (training of trainers, organization of meetings, number of participants, etc.).

One benefit deriving from this first phase is the positive impact the results can be seen to have in the countries. In Morocco, for instance, the available RAMAA data provided significant insights used by the country to revamp its national literacy strategy and road map for 2014–2020. In Niger, RAMAA results were used to map out a plan to speed up adult literacy, which the government recently adopted. The possibility of using RAMAA to develop a master's-level course at the University of Dakar's Ecole Supérieure d'Economie Appliquée (ESEA) is also under consideration.

All these findings point to the need to initiate the second phase of RAMAA with the ultimate goal of having a tighter focus on quality standards based on proven and established scientific principles and providing guidelines that everyone involved will understand and be able to apply consistently, at least to a significant extent. Furthermore, the learning dynamic and the cohesiveness of existing national teams must be reinforced in order to embrace a broad spectrum of contributors with profiles that suit RAMAA (national teachers' networks, existing national evaluation networks, PhD students, etc.) and thus create a link between fundamental research and practice.

In addition to the five countries involved in the first phase (Burkina Faso, Mali, Morocco, Niger and Senegal), seven other countries (Benin, Cameroon, Ivory Coast, Central African Republic (CAR), Democratic Republic of Congo (DRC), Chad and Togo) have expressed their potential interest in participating in the second phase of RAMAA, planned for 2015.



Arne Carlsen
Director
UNESCO Institute for Lifelong Learning

The results of the first phase of action research on measuring literacy programme participants' learning outcomes (RAMAA) were obtained thanks to the unconditional commitment of the policymakers of the participating countries – Burkina Faso, Mali, Morocco, Niger and Senegal – under whose supervision national coordinators and national team members played a leading role. We commend them for their active involvement and thank them also for their contribution to the elaboration of this publication. We also express our gratitude to the Swiss Agency for Development and Cooperation (SDC) without whose financial support this first phase would not have seen the light of day.

We wish also to thank the Working Group on Non-Formal Education (WGNFE) of the Association for the Development of Education in Africa (ADEA) for supporting RAMAA's national activities.

We express our profound gratitude to Professor Mohammed Bougroum, lecturer and researcher at Cadi Ayyad University, Marrakech, who has provided invaluable support in the conceptualization of RAMAA, the design of various measuring tools and the analysis of the country reports. Similarly, we wish to thank all the external experts, who, despite becoming involved in this first phase of RAMAA when it was at a fairly advanced stage, contributed extremely useful tools: Anne-Marie Charraud, education consultant, Giselle Boisvert, former educational adviser to the Montreal School Board (CSDM) – adult education section, Nicolas Jonas of the Organisation for Economic Co-operation and Development (OECD), Jean Dumais and Yves Morin of Statistics Canada.

We reiterate our thanks to Nicolas Jonas for his unconditional assistance during the drafting of the publication. Also to William Thorn of the OECD, Professor Anke Grotlüschen of Hamburg University and to our colleagues at the UNESCO Institute for Statistics (UIS) and the UNESCO National Office in Abidjan for their input and recommendations for the second phase of RAMAA.

The UNESCO Institute for Lifelong Learning (UIL) and the UNESCO Offices in Abuja and Dakar express their sincere gratitude to the former Director of UIL, Adama Ouane, who pioneered the RAMAA initiative and who made himself available to critically review this publication.

To the late Lynne Chisholm, our dear colleague, we extend our posthumous thanks for her scientific dedication that helped to advance this first phase of RAMAA.

Finally, we gratefully acknowledge the significant support provided by colleagues in the UIL's publications department, particularly Leona English.

A3F	Acquisition of basic and functional French
ADEA	Association for the Development of Education in Africa
APENF	Association for the Promotion of Non-Formal Education (Burkina Faso)
BREDA	UNESCO Regional Bureau for Education in Africa
CAENF	Centre for Literacy and Non-Formal Education in Morocco
CAP	Teaching centres
CNOAS	Senegalese National Coalition of Literacy Providers
CPS	Planning and Statistics Office Mali
CRFPE	Regional Centre for the Training of Education Personnel
CST	Training in scientific and technical culture
DALN	Senegalese Directorate for Literacy and National Languages
DLCA	Illiteracy Eradication Directorate Morocco
DNEF-LN	Malian National Directorate for Non-Formal Education and National Languages
ESEA	School of Applied Economics Dakar
EFA	Education for All
FCB	Basic complementary training
FDC-BF	Foundation for Community Development/Burkina Faso
FTS	Specific technical training
IALS	International Adult Literacy Survey
IFAENF	Institute for Literacy Training and Non-Formal Education
INS	Nigerian Institute for Statistics
INSTAT	National Institute for Statistics Mali

LAMP	Literacy Assessment and Monitoring Programme
MENA	Ministry of National Education and Literacy of Burkina Faso
MENALN	Ministry of National Education, Literacy and National Languages of Niger
NFE	Non-formal education
NFE-MIS	Non-Formal Education Management Information System
NGO	Non-governmental organization
OECD	Organisation for Economic Co-operation and Development
RAMAA	Action Research on Measuring Literacy Programme Participants' Learning Outcomes
REFLECT	Regenerated Freirean Literacy through Empowering Community Techniques
SDC	Swiss Agency for Development and Cooperation
UIL	UNESCO Institute for Lifelong Learning
UIS	UNESCO Institute for Statistics
UNESCO	United Nations Educational, Scientific and Cultural Organization
WGNFE	Working Group on Non-Formal Education



BACKGROUND AND JUSTIFICATION

Achieving a 50% improvement in levels of adult literacy by 2015, especially for women – i.e. accomplishing Education for All (EFA) Goal 4 – is a far cry from the reality in most developing countries in spite of progress so far. In sub-Saharan Africa in 2011, 182 million adults and young people aged 15 or above – two-thirds of whom were women – faced exclusion and lived in precarious social and economic conditions (EFA Global Monitoring Report, 2013–14). West Africa¹ is the most severely affected region: literacy rates there are below 50% and young people aged 15 to 24 account for 44% of its illiterate population.

Illiteracy rates are usually generated based on self-reported data that classify people into opposing categories of literate and illiterate. Such data provide vague information with limited value in terms of creating policy or addressing individual needs. (Wagner, 2005; UNESCO Institute for Statistics, 2008). In reality, we are also faced with individuals who span a broad spectrum of illiteracy categories and degrees, engage in different practices, and have various expectations and requirements vis-à-vis literacy. Taken in isolation, illiteracy rates are not enough to devise literacy programmes tailored to participants' profiles.

This lack of reliable and relevant data therefore undermines all subsequent efforts to steer literacy policy. Policymakers need data they can use to map out suitable action plans, set funding priorities, allocate additional resources, and assess the impact of their own actions as well as partners' actions.

The UNESCO Institute for Lifelong Learning (UIL) was prompted by these findings, and its aim to 'further literacy as a foundation for lifelong learning', to initiate an Action Research Project to Measure Literacy Programme Participants' Learning Outcomes (RAMAA, *Recherche-action sur la mesure des apprentissages des bénéficiaires des programmes d'alphabétisation*) in 2011. This project, which the UIL is running with the UNESCO offices in Abuja and Dakar in keeping with UNESCO's "Delivering as One" approach, stems from extensive concerted deliberations with countries and high-level experts.

RAMAA is the first generation of action research geared to measuring literacy. Its methodology centres primarily on literacy programmes. Specifically, RAMAA is aimed at furnishing policymakers and development partners with information about the quality of the various literacy programmes by measuring learning outcomes among adults and young people (age 15 or above). The tools to measure these outcomes are developed using a common methodological framework that addresses cultural, educational and linguistic specifics, as well as the country's programme design, management and implementation expertise and capacities.

In short, RAMAA intends to assist a critical mass of national experts in taking ownership of the various stages. This action-research project's goal, in other words, is not to develop

1 For the purposes of this study, Mali, Burkina Faso, Niger (29%), Guinea (38%), Sierra Leone (40%), Benin (41%), Senegal (42%) and Gambia (45%).

sophisticated measurement tools that will prove impractical in the field: it is to devise simple inexpensive tools that will further RAMAA's scientific objectives and concurrently accommodate each participating country's human, material and financial capacities.

RAMAA results will not only spur continuous improvement in the quality of adult and youth literacy programmes outside school systems: they will also serve as catalysts to build fact-based advocacy cases. RAMAA will also enhance national capacities in terms of ownership and sustainability, and will open up new prospects for applied scientific research.

What decision-makers need to know

There are prerequisites:

- ◆ Literacy is a human right that must be available to all without discrimination.
- ◆ The right to literacy has a powerful impact on a country's economic, social and cultural development inasmuch as it contributes to substantially reducing poverty, decreasing child mortality, curbing demographic growth, achieving gender equality, and securing sustainable development, peace and democracy.
- ◆ Literacy is the stepping stone to lifelong learning and therefore the foundation to start building instrumental skills leading to employment and citizenship, and fostering sustainable development.

Therefore:

- ◆ Persistent massive illiteracy is seriously jeopardizing efforts to honour international commitments (EFA, Millennium Development Goals, Belém Framework for Action, recent sustainable development objectives, etc.).
- ◆ The fact that the number of functionally illiterate people remains high is undermining any efforts to bring demographic policy into effect.
- ◆ Low literacy rates are hampering public policies for economic growth and social cohesion.
- ◆ Current national literacy policies show that policymakers are determined to make headway on this front. But these policies lack substance: they do not bring about national medium- and long-term strategies accompanied by significant budgetary commitments.
- ◆ The improvements as regards access to literacy programmes (in terms of course enrolment) do not provide sufficient grounds to hope that the current trends will change direction.

This context explains why setting up a system to manage literacy programme quality is justified. This steering mechanism, again, will flesh out the case for the right to education and, beyond that, generate reliable data that policymakers and development partners can use for guidance. It will additionally rank basic literacy programmes for adults and young people outside the school system as an educational priority on a par with basic schooling.

Overall, the breadth of the issues in the field of basic education, and the challenges awaiting on that front, call for a comprehensive approach. This comprehensive approach should fully integrate formal and non-formal education, and thereby provide real opportunities for all to access high-quality education and lifelong learning. Transferring this comprehensive approach into operations necessarily entails optimizing resources. This in turn means pooling resources and experiences across formal and non-formal education, and coordinating sector-wide policy (for health, employment, education, etc.).

The scope of RAMAA

RAMAA targets developing countries in general and focuses on French-speaking African countries in particular. It was devised taking the following into account:

- ◆ Low adult literacy levels are prevalent in most African countries, hence the need for a statistical system to measure, monitor and assess programmes.
- ◆ The RAMAA project's priority is to cater to national policymakers striving to improve literacy process efficacy.
- ◆ The project should produce recommendations that central and decentralized policymakers are likely to follow and factor into their decision-making processes. Also, their continuous involvement in all project phases is a sine qua non for RAMAA rollout.
- ◆ RAMAA's action-research approach should bolster national capacities in the participating countries, shed light on the needs that each of those countries share and encourage those countries to pool their experiences. This approach will help set up a statistical system geared to literacy programmes that will remain relevant and viable over the long term.

The first phase of the project (RAMAA 1) involved five African countries: Burkina Faso, Mali, Morocco, Niger and Senegal.

Illiteracy is a structural reality in all RAMAA 1 countries, as is rapid demographic growth. Illiteracy rates exceed 50% in all of them except Morocco, where the figure is around 38% (*EFA Global Monitoring Report, 2013–14*).

Dysfunctions in national education systems also impact this widespread illiteracy.

In these countries, formal education is indeed fuelling basic illiteracy. Limited access to the first year of primary education, low retention rates and declining teaching standards in this sector are doing little to increase literacy rates. At the same time, efforts to provide literacy programmes for adults and young people, substantial though those efforts are, have not yet brought baseline functional education, enabling self-reliance and empowerment, to a majority of the population outside the schooling system.

These countries are therefore aware that they need to step up their efforts. At the same time, reliable statistics are becoming more crucial than ever to steer public action in the non-formal education sector. This is precisely what RAMAA aims to provide.

This project's novel approach, and the advantages that the countries saw in enhancing their capacity for assessing adult literacy programmes, prompted the Swiss Agency for Development Cooperation (SDC) to provide critical funding for RAMAA 1 (the SDC encourages efforts on this strategic front in the non-formal basic education sector, with a view to contributing effectively to sustainable development in African countries).

RAMAA 1 protocol

The various steps were approved and taken as follows:

- ◆ **Preparation work**, i.e. consulting, reaching consensus on and building a conceptual and operational framework, establishing a framework for partnerships with participating countries, and appointing national teams. The consultation workshop to agree on the concept note and research plan was held in 2008 in Marrakech, Morocco. In 2010, Burkina Faso, Mali, Morocco, Niger and Senegal signed memoranda of understanding stating their political and financial commitment to the project.
- ◆ **Tool development**, encompassing the conceptual, methodological and operational aspects of the measurement and data-collection instruments. However, before initiating this step, it was necessary to develop national communication plans and national capacity-building plans.

The national capacity building plans aim to ensure that (a) national teams acquire the technical skills they need to manage each research phase; (b) capacity-building efforts cascade into infra-national (regional and local) levels; (c) the project effectively harnesses the potential in South-South cooperation, starting with exchanges among national teams; (d) all the stakeholders who have a key role to play in the action-research project's success (subsector administrative heads, national statistics institutes, universities, local authorities, civil society and so on) are on-board; (e) every step in the project is documented.

The national communication plans in particular leverage new technologies for (a) communication within and among teams; (b) communication targeting the various outside audiences (policymakers, technical and financial partner organizations, and NGOs) through the various channels (documents, forum participation, etc.).

Once these plans were finalized, work in 2011 focused on discussing and preparing the skills descriptions and related tools, namely the reference framework, core questionnaire and background questionnaires, sample design and procedure manual.

- ◆ **The pilot research**, using the measurement tools that had been developed on a participant cohort. Burkina Faso, Morocco and Senegal completed this step in 2012. Niger and Mali did so respectively in 2013 and 2014. The countries proceeded at different paces but, in most cases, managed to make the necessary adjustments. Morocco was the only one that did not, because it had already reached a very advanced stage at that point.
- ◆ **The main research**, which Burkina Faso, Morocco and Senegal completed, paved the way for database analysis and national reports on findings. The ensuing recommendations from these countries, and the experts at the workshop in Hamburg in 2014, concurred that the measurement tools needed to be reviewed. Mali and Niger, as well as the other countries, agreed.
- ◆ **Result dissemination** among the various target audiences (international organizations, national policymakers, selected national players, and the general public). We did not reach this stage because findings during the first phase of the project suggested that measurement tool quality warranted review. The tools and results have nevertheless had an impact at national level. In Morocco, for instance, RAMAA data provided meaningful insights that the country used to revamp its national literacy strategy and road map for 2014–2020. In Niger, RAMAA results were used to map out a plan to speed up adult literacy, which the government recently adopted. There is also talk about using RAMAA in tentative plans to develop in a master's-level course at the University of Dakar's École Supérieure d'Économie Appliquée (ESEA).

The purpose of this document analysing RAMAA 1

This document's main purpose is to help policymakers, programme providers, development partners and the general public to understand the RAMAA project and the results its first phase has yielded. One advantage is that it could spur exchanges on the lessons learned and recommendations, with a view to improving measurement tool quality and possibly involving other countries in RAMAA 2. That is why this report on RAMAA 1 is written in simple, non-technical language that the various stakeholders will readily understand.



ACTION RESEARCH ON RAMAA

Action research is a participatory and collaborative form of research (Catroux, 2002; Kemmis, 1997; Liu, 1992) involving a practitioner and a researcher working in tandem. They aim to fully understand the issues under review, by intertwining the theoretical and practical knowledge they build by dealing with those issues, and to try out and test new approaches. Their goal is to achieve a deeper understanding of the programmes, and thereby improve the programmes' effectiveness and efficiency (Lefrançois, 1997).

From the start, this approach implies that the practitioner is acknowledged as a skilled contributor and that both – the practitioner and the researcher – contribute meaningful social insights (Berger, 2003, p. 14). The practitioner and the researcher, in other words, pool their complementary skills (Coenen, 2001). Their action-research project's success, intrinsic quality and soundness hinge, to a large extent, on their teamwork (Savoie-Zajc, 2001).

The practitioner's role, moreover, involves experimenting and challenging action from within. The practitioners, in other words, are the main architects of their emancipation (Coenen, 2001; Elliott, 2007).

In short, action research cannot be considered a methodology wherein one group remote-controls changes in another group from the outside. It thrives on co-construction dynamics (Guillemette et al., 2012).

These are the postulates that have shaped RAMAA. As it is ushering in a novel approach to literacy programme assessment, it aims to support a critical mass of national experts in taking ownership of all its stages and to embed RAMAA in their country's national mechanisms for the long term. RAMAA's action-research approach therefore provides a framework that helps to achieve these objectives.

The members of the national teams are experts selected based on the various profiles that the project requires. They include assessors, education economists, statisticians, sociologists and education specialists.

The learning dynamics are set in motion through exchanges between UNESCO and the national experts, fuelled by experience-sharing among countries. Action research within these interconnected exchanges involves an interactive and iterative cycle comprising three steps: (a) preparing and approving matrices (guidelines); (b) producing the countries' measurement instruments and data collection tools based on the approved matrices; (c) assessing measurement tools, adjusting them if and as appropriate, and reviewing each country's production in perspective.

Another advantage of action research is that it remains in flux and can therefore accommodate adjustments and critique, which in turn drive research forward. The adjustments are not preordained: they arise amid consultation and controversy during stakeholder meetings. These meetings are all the more important as they provide opportunities to share experiences and build a common language (Beaupère et al., 2010).

The RAMAA project has integrated this important dimension. UNESCO staff, international experts and national team members gather for international seminars during each project stage. At these seminars, they discuss, adjust and approve the previous stage's results, and agree on the terms of reference for the following stage. Participating countries also organize national meetings to prepare, discuss, approve or report on developments in the project.

Generally speaking, action research is viewed sceptically in scientific circles on account of its co-constructive dimension. We nevertheless maintain that the intercultural and multicultural professional views that national experts, external experts and UNESCO pool nurture cross-fertilization. The fact that a variety of profiles share their perspectives, as Lucile Courtois (2013) points out, makes it possible to interlink scientific and social concerns, theoretical and practical issues, and academic knowledge and field-experience feedback. This cross-fertilization provides the keys to unravelling complexity.

RAMAA action research is therefore scientifically sound, and indeed ethically sound as it likewise promotes democratic values and principles (respect for participants, inclusion in research and quality of exchanges) (Gohier, 2004; Savoie-Zajc, 2001).

- Loowdi Haatumeeere
 Janori Kunaari
 Golle Burda nafude

Golle Burda
 Golleade.

Aden

Jonnde
 Keedi

Jaagu

Awgal

Golle juude

Margol
 Daaba
 Raji

Janori	Jannde	T'ellal	Sandya	Yellal	Jannde
4	2	1	1	-	
-	-	-	-	5	
2	-	-	5	3	
5	-	-	-	1	



CONCEPTUALIZING RAMAA: AN INNOVATIVE MEASUREMENT OF LEARNING OUTCOMES

A methodological choice for literacy management

RAMAA is the first action-research project focused on measuring literacy. As such, it sets out to fill an information gap that no other project currently addresses: assessing the quality of the various literacy programmes by measuring learning outcomes. This action-research project develops standardized/harmonized measurement tools (more about the terms 'standardization' and 'harmonization' in RAMAA below) in order to generate results.

In this regard, RAMAA chose to use a different methodology for literacy management, and complements two other UNESCO projects:

1. The Literacy Assessment and Monitoring Programme (LAMP), run by the UNESCO Institute for Statistics (UIS), which measures literacy and numeracy standards among entire adult populations in given settings.
2. UNESCO's Non-Formal Education Management Information System (NFE-MIS), which produces and disseminates literacy indicators. One aspect of setting up this system is to collect data on the processes used to implement literacy programmes.

Together, these three projects produce robust data and thereby considerably underpin the fact-based advocacy case for literacy.

RAMAA and LAMP results could feed the NFE-MIS. We also need to consider synergies between LAMP – its tools draw extensively on surveys by the Organisation for Economic Co-operation and Development (OECD) – and RAMAA on the method for measuring administrative procedures². Moreover, RAMAA can leverage a number of NFE-MIS activities, in particular its work mapping non-formal education programmes with a view to facilitating fieldwork.

The teams in Morocco and Niger have set up the three initiatives mentioned above, and the team in Senegal is working with the NFE-MIS. This is opening up promising prospects for South-South cooperation with a view to consolidating national capacity.

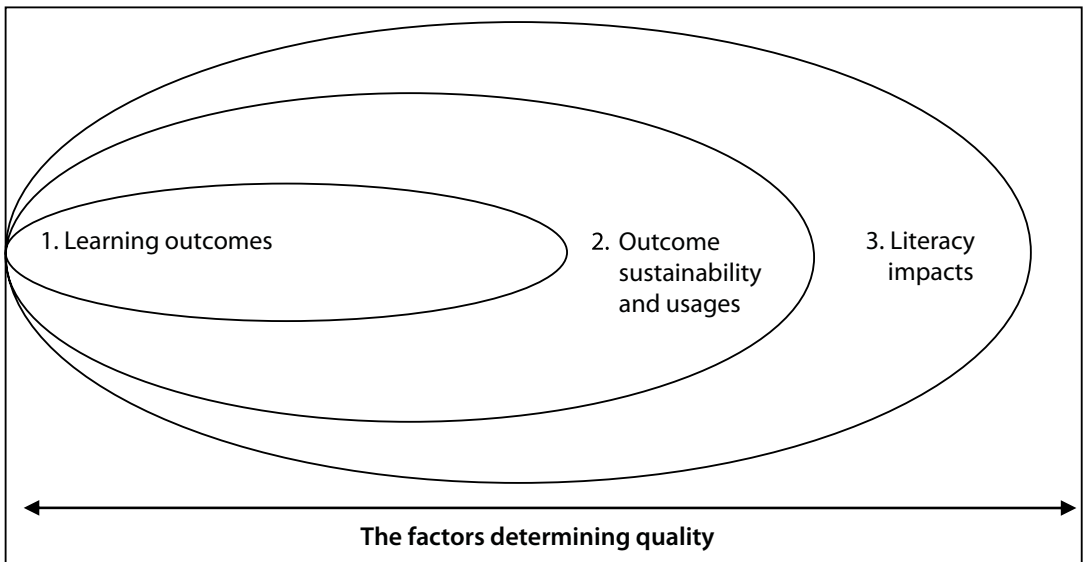
2 LAMP's design is principally based on the OECD's International Adult Literacy Survey (IALS) and Adult Literacy and Lifeskills (ALL) Survey.

Assessments based on external measurements of learning outcomes

RAMAA is an external assessment inasmuch as it appraises literacy-programme quality by measuring outcomes on successful programme completion, not by reviewing learning processes. These literacy programme assessments can span three levels: (a) outcomes on literacy programme completion; (b) outcome sustainability and usages; (c) literacy impacts on participants, their families and their communities (see Figure 1).

Another dimension – the factors determining quality in the environment – cuts across the three assessment levels mentioned above. These factors determining quality explain result dispersion within the same programme and between various literacy programmes.

Fig. 1. The sequential levels of outcome quality analysis



The three learning-outcome assessment levels, as depicted in the figure above, tally with the information needed by decision-makers tasked with steering adult literacy. However, it is difficult to provide answers on all these levels at the same time in a single action-research project such as RAMAA.

The first phase of the project (RAMAA 1), therefore, focused mostly on the first level. The central questions in the research follow:

- i) What have adults learned on completion of a literacy programme?
- ii) What factors determine the variability of those outcomes?

The second measurement level – outcome sustainability – was introduced at the participating countries' request, as an optional aspect of fieldwork. This was completed with an additional optional question aimed at providing empirical evidence to challenge the notion that literacy programmes are not the only (nor the best) channels to learn these basic skills, as conveyed in certain political messages. These two additional optional questions follow:

- i) How are literacy programme outcomes evolving over time?
- ii) To what extent does attending a literacy programme determine skill acquisition?

As well, RAMAA focuses on measuring outcomes directly, meaning that it assesses learning directly face-to-face with participants, using written documents ('paper-and-pencil' tests).

Assessments based on standardized measurement tools

RAMAA set out to develop standardized measurement tools, i.e. harmonize its development methods, documents and questions. This action-research project's standardized tools are designed to gather and measure the same datasets in different countries and in different contexts. In other words, the goal is to develop operational quality-measurement tools that strike a balance between international standardization and contextualization. This approach does not preclude comparing results across various countries but its focus, at first, is on comparing programmes within each country.

A reference framework based on a novel skills description

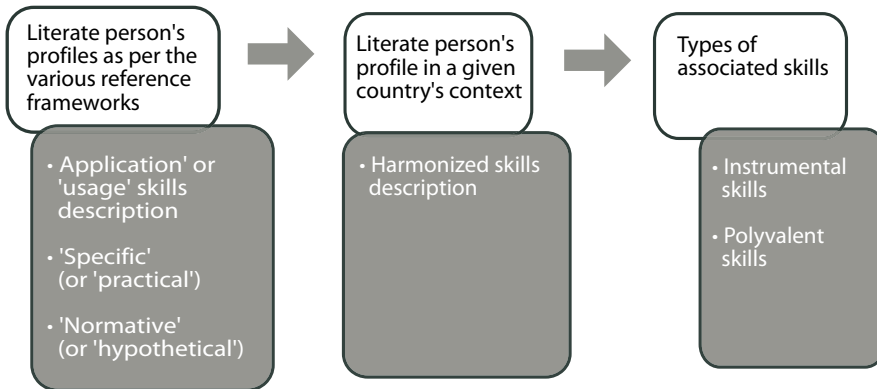
The reference framework, by definition, provides the key skills that are to be measured, the type of documents to be used and the contexts to be taken into account (RAMAA Technical Paper, 2012). Reference frameworks are usually based on skills descriptions, as Périsset Bagnoud (2007) and Nadine Postiaux et al. (2010) point out, and prescribe the skills that participants should possess on completion of their learning programmes. In other words, reference frameworks are basic qualitative tools to align skills and training requirements (UNESCO, 2013).

The RAMAA 1 reference framework – referred to as the harmonized framework – was developed in light of the skills profiles that adults require to be deemed 'literate' in their specific contexts, in each country taking part in this action-research project. This tool was developed from a common matrix and was adopted by the national teams working on RAMAA 1. It combined the three levels below (see Figure 2):

- ◆ *'Application' or 'usage' skills descriptions*, which refer to literacy skills used in the social, educational and professional activities that literacy programme graduates can access in each of their countries;

- ◆ *'Specific' skills descriptions*, which are based on the reference frameworks used by the principal literacy programme operators (governments, NGOs, the voluntary sector and the private sector);
- ◆ *'Civic' or 'normative' skills descriptions*, which describe the literate citizen's ideal profile, as portrayed in documents laying down public policy and in citizens' perceptions.

Fig. 2. RAMAA harmonized skills descriptions and the associated skills



The literacy skills and levels required in the skills descriptions

A large corpus of work consistently highlights two salient aspects of a skill in its definitions (Report on the item map review, 2013):

- A skill is the ability to efficiently and effectively complete a task. It therefore relates to a person's ability to take action with a view to achieving a purpose, i.e. a goal. To assess someone's reading skills, for example, we need to ask them to read.
- A task, however, lacks meaning and purpose if it is not embedded in a context. In other words, the various particular aspects of the context are included in the description of the skill or contribute to the level of achievement.

In short, developing skills involves deepening action and adjusting to specific contexts. Skills are therefore aimed at attaining suitable results in the given context.

RAMAA agrees: it defines a skill as an individual's ability to use literacy programme outcomes to deal with a real-life situation in a given context.

We initially considered three key skillsets in the skills description for RAMAA 1 (Framework document for the reference framework and background questionnaires, 2012):

- i) Instrumental skills, meaning adequate proficiency in basic aspects of literacy and numeracy, namely the ability to read, write, calculate and communicate in at least one language.
- ii) Functional skills, meaning an individual's ability to use written documents and literacy skills to carry out tasks related to the various roles they are required to fulfil as a member of a family and community, and/or as a producer, consumer and citizen.
- iii) 'Transversal' (or 'polyvalent') skills, meaning an individual's ability to identify and process information, apply strategies to solve problems, and use technology to consolidate and/or deepen instrumental skills and to implement functional skills.

We later reviewed these skills in order to underline the contexts where they materialize. This involved grouping functional and instrumental skills under 'instrumental skills', and reclassifying transversal skills under 'knowledge' because questionnaire-based interviews such as the ones we used in RAMAA were unable to measure them. (*RAMAA Report 2012, Report on the item map review, 2013*).

Following revision, these skills were defined thus:

- ◆ Instrumental skills refer to an individual's level of proficiency in the basic aspects, and the ability to use them productively in everyday tasks.
- ◆ Knowledge refers to general life abilities (communicating effectively, working methodically, and thinking critically) and the specifics are inventoried within each country context. Knowledge relates to culture and people may acquire new knowledge in their daily lives, i.e. outside literacy programmes (*Report on the item map review, 2013*). The principal topics in all RAMAA 1 countries fell into one of four areas of knowledge: health and well-being, citizenship, the environment and work.

Once we had defined the key skills thus, we proceeded to measure them, on a scale ranging in a **continuum** from Level 1 to Level 3, as follows:

- ◆ **Level 1** participant has very limited basic skills.
- ◆ **Level 2** participant can apply the target skill in practice.
- ◆ **Level 3** participant can transfer the skill or use it in an appropriate situation serving a purpose in their daily life, in a manner denoting self-reliance.

The levels of proficiency in instrumental skills and knowledge were built and used on this basis in RAMAA 1. The details follow:

Instrumental skills (see Table 1)

Language skills

- ◆ **Level 1** (words): participant can associate a written word with an object, concept or situation.
- ◆ **Level 2** (sentences): participant can combine words to understand or form a sentence.
- ◆ **Level 3** (texts): participants can combine a number of sentences to understand or express a topic or idea.

Mathematical skills

- ◆ **Level 1** (vocabulary): figures, numbers or measurement units.
- ◆ **Level 2** comparisons between two quantities or two measurements.
- ◆ **Level 3** operations using quantities or measurements of space or time.

Knowledge (see Table 2)

1. *Describing* a situation, e.g. the components in a set of symptoms suggesting malnutrition or types of identity documents.
2. *Explaining*, which implies knowledge of the rules (on healthy diets or the procedure to obtain identity documents) used to function in a given situation.
3. *Solving problems* in a given situation can constitute a third level. This one is the closest to a skill in that it involves applying knowledge in a specific situation. Producing an invoice in a work-related situation is one example of problem-solving. This stage almost invariably requires instrumental skills, and it is important to take this into account when preparing the items.

In short, RAMAA is not an attempt to assess literacy programme participants in yes-or-no (literate or illiterate) terms: it is rather an attempt to define participants' profiles on the three levels.

Table 1. Reference frame: instrumental skills

	LEVEL 1	LEVEL 2	LEVEL 3
READING	Identify/recognize written words	Read and understand the meaning of a sentence (instruction, rule, etc.)	Read and understand a brief text
WRITING	Spell dictated words or copy words	Write or complete a full sentence conveying meaning	Write a brief text on a given subject
ARITHMETIC	Count objects or currency; write numbers	Compare numbers of objects or amounts of currency; more than, less than, equal to, as much as, each, etc.	Perform operations with numbers and money
MEASURING	Read and write units of time (hour, day, week, etc.), distance and volume (kilometre, metre, kilogramme, etc.)	Situating oneself in time and space (before, during, after, above, inside, outside)	Measuring time and space

Source: Reference framework, 2013

Table 2. Reference framework: polyvalent skills/knowledge

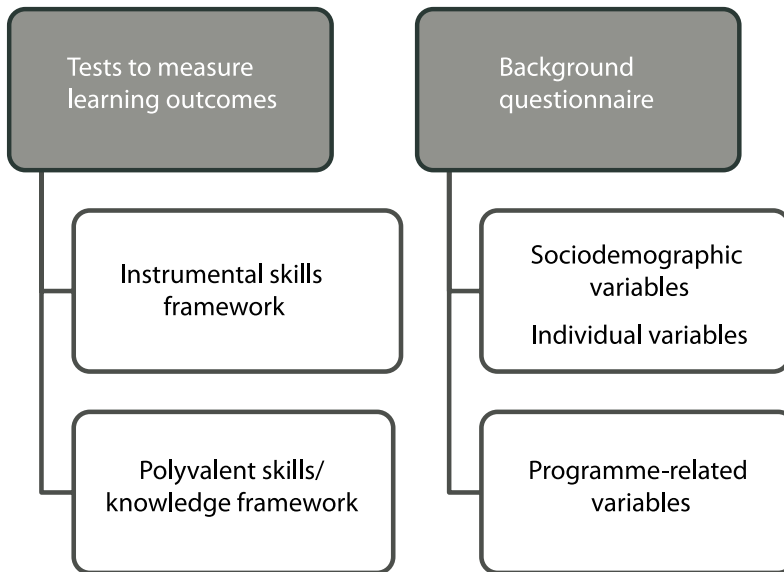
TOPIC	LEVEL 1	LEVEL 2	LEVEL 3
T1 HEALTH AND WELL-BEING	<ul style="list-style-type: none"> • Describe aspects of nutrition best practices • Name signs of malnutrition • Identify symptoms of infectious diseases 	<ul style="list-style-type: none"> • Explain the rules for remaining healthy • Explain the causes and consequences of malnutrition • Explain the causes of certain infectious diseases 	Identify problems in an illustration of a real-life situation, suggest more beneficial courses of action
T2 ENVIRONMENT	<ul style="list-style-type: none"> • Describe sanitation practices for living quarters • Name actions aimed at protecting and restoring the environment 	<ul style="list-style-type: none"> • Explain the environmental protection principles and rules 	Devise a plan to create a healthier and more productive living environment
T3 CITIZENSHIP	<ul style="list-style-type: none"> • Name official documents and the places where they can be obtained • Name the rules of citizenship • Identify a citizen's rights and responsibilities 	<ul style="list-style-type: none"> • Explain the purpose and importance of official documents • Explain the importance of voting • Explain the importance of certain cultural values in one's surroundings • Explain actions to promote citizens' rights and responsibilities 	Devise a plan to create a healthier and more productive living environment
T 4 WORK	<ul style="list-style-type: none"> • Name the tools used to plan and manage an income-generating activity (IGA) • Name the steps in the administrative procedure 	<ul style="list-style-type: none"> • Explain the steps to prepare an IGA project • Explain the IGA planning and management rules • Explain how a technique or product improves production 	

Source : Reference framework, 2013

Contextualizing information on learning outcomes using a detailed questionnaire

RAMAA provides information on the level of skills that participants have acquired on successful completion of various literacy programmes. It also provides policymakers, programme providers and development partners with insights into the factors that influence variability in these outcomes. The factors determining quality that we singled out when preparing the background questionnaire focused on three types of variables: (a) individual (the participant's motivation and outcome usage); (b) sociodemographic (the participant's household and community); (c) programme-related variables (course management and content relevance as regards participants' expectations).

Fig. 3. The principal tools to measure adult learning outcomes in RAMAA





DATA COLLECTION TOOLS AND QUALITY ASSURANCE

In the interest of harmonization, we also developed guidelines for data collection tool operationalization. The associated matrices applied to the target population, sampling frame and sample design.

The target population

RAMAA's target population consists of literacy programme participants and comprises adults and young people age 15 or above up living in the participating countries at the time the survey was conducted.

To answer the two core questions in the survey (as mentioned above), fieldwork targeted the participants who had completed adult literacy programmes most recently.

To appreciate the factors causing outcome variability more accurately (i.e. answer the second core question), we also gathered data from key contributors to the literacy process, namely trainers and people managing programmes locally.

The answers to the other two (optional) questions came from comparison groups. The first comparison group provided answers to the optional question about outcome sustainability. Measuring this would otherwise have involved conducting a longitudinal study, i.e. monitoring literacy programme graduates in various situations relating to their social, family and economic settings over several periods. For practical reasons, this first comparison group was built from the population that had completed literacy programmes 24 to 36 months prior.

The second comparison group, which we used to address the second optional question, reflected the adult population assumed to possess no literacy skills, meaning that the people in this group had never attended formal or non-formal basic education programmes.

Sample design

Sample design encompasses all the steps required to identify the units that will be included in the sample. To provide robust results, the sample must represent the entire population under consideration. For RAMAA 1, however, population representativeness was defined thus:

The principal sample is derived from the entire population of the most recent group of graduates. The definition of this sample may vary from one country

to another in light of two parameters: (a) the study's geographic scope (national or infra-national); (b) the study's linguistic configuration (cf. Sample design 2012–2013).

A three-tier sample design was then tendered based on these parameters. It took into account: (a) the literacy programmes; (b) the geographic area; (c) the graduates' profiles. All types of literacy programmes need to be represented regardless of their objectives, duration and target populations. Samples must also reflect geographic diversity and consider the structure of the participant population in light of at least two stratification variables, namely age and gender.

- ◆ Comparison group 1 is built in the same way as the principal sample.
- ◆ Comparison group 2 is selected from the same environment as the principal sample and comparison group 1.

Sampling frames

A sampling frame is tantamount to a list of individuals from which a sample is selected. This list determines the population under observation.

The principal sample in the RAMAA sampling frame is derived from a list of all the individuals who had enrolled in literacy programmes and passed the test at the end of the latest cycle (2012–2013).

Comparison group 1 comprises the graduates from the two previous cycles (2011–2012 and 2010–2011).

No sampling frames are available for comparison group 2. Field teams select this cohort non-probabilistically among the relatives of the people selected for the principal sample or comparison group 1.

Sample size

Due to time restrictions and budget constraints, the following sample sizes are recommended:

- ◆ Principal sample: 1,000, 1,500 or 2,000 people
- ◆ Comparison group 1: 250 to 500 people (who took literacy courses in previous years)
- ◆ Comparison group 2: 250 to 500 people (who have never taken literacy courses)

RAMAA quality assurance

Surveys to measure adults' skills are relatively problematic to conduct, and teams need to follow a number of procedures to ensure the data they collect are reliable and usable. RAMAA teams have therefore taken several measures to ensure action-research quality. These measures mirror RAMAA's two objectives, namely (a) develop context-sensitive measurement tools to improve literacy policy and programme quality; (b) reinforce national team capacity in this area.

The main guidelines for the research and measurement tools are therefore honed through a process encompassing consultation, implementation, assessment and building consensus.

Reinforcing capacity is one of the goals embedded in this process. Besides international meetings, 'tailored' technical support is provided for each national team with field missions by the UIL and by RAMAA-partner UNESCO offices, backed by external experts. Country output was also reviewed with countries, UNESCO and external experts.

We devised a steering system geared to optimizing RAMAA coordination efficacy and quality assurance, and subsequently adjusted it during action-research rollout.

RAMAA 1 initially included an international steering committee and a scientific committee.

The steering committee was tasked with supervising, leading and directing research, and its members were specialists from the UIL, the Regional Office for Education in Africa (UNESCO-BREDA)³ and experts representing the participating countries' ministries of education, which oversaw the selected national coordinators and national teams conducting the research.

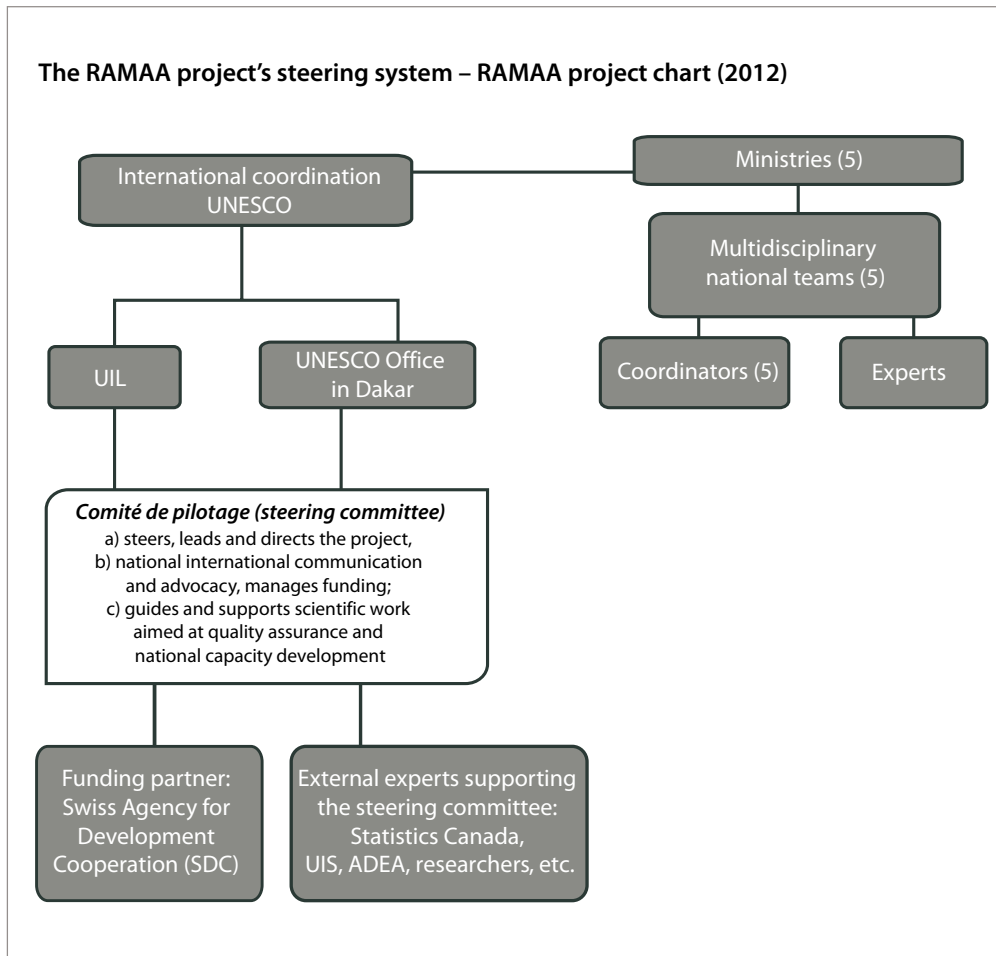
The scientific committee provided methodological, technical and scientific support. This committee comprised a scientific coordinator, UIL and BREDA experts, and external experts appointed at each research stage when and as specific requirements warranted doing so.

³ Following UNESCO reform, RAMAA is currently piloted in partnership with the UNESCO Regional Office in Dakar and the UNESCO Regional Office in Abuja.

Governance was reviewed in August 2012 to optimize efficiency. The new organization follows:

- ◆ The *international steering committee* (UIL and BREDA) serves the same purpose as before;
- ◆ The *project coordinator* (a UIL member) coordinates and monitors the project;
- ◆ The *scientific advisor* supports the steering committee;
- ◆ The *international experts/consultants* are selected at each research stage.

The RAMAA project's steering system – RAMAA project chart (2012)





RAMAA'S CONTRIBUTION TO EVIDENCE-BASED ADVOCACY AND CAPACITY DEVELOPMENT

RAMAA results and public action: overall objectives

RAMAA is the first action-research project centred on assessing literacy programme quality in participating countries.

Results in terms of skill levels acquired on literacy programme completion, measured in light of contextual factors, will inform policymakers on literacy programme quality.

These RAMAA results will streamline existing literacy programmes in countries taking part in the project, with a view to using funding more efficiently and effectively.

RAMAA results can also inform national and regional debates on accreditation frameworks, and thereby foster a comprehensive approach to basic education by reinforcing ties between formal and non-formal education.

RAMAA also provides examples of best practices in applied research: in this case for research-based training and, more broadly, university involvement in literacy drives. As student researchers will be involved, they will be able to use this study in their doctoral theses.

RAMAA results and raising professional standards in adult-literacy management: the medium- and long-term objectives

RAMAA will have a positive impact on national capacity development and outcome sustainability inasmuch as this action-research project follows a horizontal dynamic based on a participative approach (building with country teams), an integrated approach (involving national teams in all project stages) and a long-term approach (building momentum). The RAMAA project thus departs from the vertical top-down outlook that permeates many international programmes, and which relegates national teams to following instructions (*RAMAA, a Brief Presentation, 2014*). The goal here, in other words, is to support participating countries as they build and produce their own knowledge, thereby attain ownership of the project and learn how to lead it over time, and, ultimately, embed it in their national systems.



1. **Tableau 1** : Les données de base
 2. **Tableau 2** : Les données de base
 3. **Tableau 3** : Les données de base
 4. **Tableau 4** : Les données de base
 5. **Tableau 5** : Les données de base
 6. **Tableau 6** : Les données de base
 7. **Tableau 7** : Les données de base
 8. **Tableau 8** : Les données de base
 9. **Tableau 9** : Les données de base
 10. **Tableau 10** : Les données de base

Jamde Avandeece
 1. **Tableau 1** : Les données de base
 2. **Tableau 2** : Les données de base
 3. **Tableau 3** : Les données de base
 4. **Tableau 4** : Les données de base
 5. **Tableau 5** : Les données de base
 6. **Tableau 6** : Les données de base
 7. **Tableau 7** : Les données de base
 8. **Tableau 8** : Les données de base
 9. **Tableau 9** : Les données de base
 10. **Tableau 10** : Les données de base

RESULTS ANALYSIS

In this section, we analyse RAMAA 1 results in two steps. First, we provide a critical review of the development of measurement instruments (harmonized skills descriptions, frames of reference and questionnaires) and data collection tools (sample designs and sampling frames) in participating countries. Second, we provide a critical review of the data collected in each of those countries.

Overall, this analysis is based on the outputs that participating countries reported to the UIL and UNESCO offices working with RAMAA. These data arise from the main assessment in Burkina Faso, Morocco and Senegal, and from the field test in Mali and Niger. The first thing that stands out when reading country survey reports is that the objectives and research questions are worded in different ways (see Table 3).

Table 3. RAMAA objectives and research questions in each country

Country	Objectives and research questions
Burkina Faso	<ul style="list-style-type: none"> - Research aims to achieve two principal objectives: (a) develop a working tool to assess quality, balancing international standardization and contextualization; (b) ensure sustainability by fostering ownership among national (central, regional and local) stakeholders. - The research questions: (a) To what extent do the programmes enable sustainable proficiency in instrumental skills, with regard to the participant's skills profile, and with a view to enabling lifelong learning? (b) What factors determine variability in outcome quality levels?
Mali	<ul style="list-style-type: none"> - General objective (RAMAA): support five French-speaking African countries in the processes they are following to develop a system to measure, monitor and assess basic-literacy and post-literacy programmes for adults, in order to remedy the lack of reliable statistical data. This system will help establish an effective system to monitor literacy programme quality, which will in turn furnish a compelling fact-based case to guide policymakers more accurately. - Specific objectives: develop and test tools to measure outcomes, using common frameworks. The two principal research objectives: (a) develop a high-quality working assessment tool, balancing international and national standardization; (b) assess experience ownership among stakeholders in the non-formal education sector, at central and decentralized levels.
Morocco	<ul style="list-style-type: none"> - General objectives: (a) source reliable information about the characteristics of target populations and literacy programmes; (b) develop a system to measure learning levels among literacy programme participants; (c) support decisions on policies and programmes tailored to the target populations' expectations and requirements. - Specific objectives: (a) measure learning outcomes among literacy programme participants; (b) measure learning outcome sustainability and usage among programme participants; (c) highlight the factors determining literacy programme quality.

Niger	RAMAA intends to (a) assess literacy programme quality on course completion; (b) study the factors determining outcome variability. Secondly, research will examine outcome sustainability by observing former learners' current skills and the role that literacy programmes play as regards leveraging outcomes.
Senegal	<ul style="list-style-type: none"> - General objectives: develop a robust system to manage literacy programme quality and build a fact-based case for advocacy among policymakers, and technical and financial partners. - Specific objectives: (a) develop a methodological approach for each stage in this measurement process; (b) reinforce participating countries' national capacities in each stage in this process; (c) develop a context-sensitive working tool to assess quality. - The overarching question/problem: Are literacy programmes participants trained with a view to enabling lifelong learning? - Specific questions/problems: (a) What skills do participants possess when they complete their literacy programme? (b) Are those skills sufficient for participants to embark on a life-long learning process? (c) What factors determine outcome variability?

The following definitions from the various concept papers also shed light on the path that RAMAA followed and how it has evolved:

In the paper presenting RAMAA when it was launched in 2010

This research project is the first step in initiating a medium- and long-term process that will lead to national evaluation systems for non-formal education programmes in developing countries. In the light of the magnitude of this domain and the little that is known about it as well as the enormous need in terms of national capacities, the study focuses on adults' (older than 15 years) learning outcomes, focusing on quality determinants for basic competences and their use.

The RAMAA technical paper (2012) and brief presentation (2013) state that:

The overall objective of this research-demonstration is to support participating countries in the process of setting up a system to steer the quality of literacy services based on two research questions revolving around ownership and national capacities reinforcement.

The three specific objectives follow: (a) develop data collection tools; (b) set up a survey system to implement these tools; (c) analyse collected data and list recommendations for public action.

In its current phase, the project is organized around two central research questions, considered sequentially: (a) What skills to adults possess on successful completion of a literacy programme? (b) What factors determine outcome variability?

*These central questions naturally entail two **optional subsidiary** questions: (a) How do the skills that adults possess on completion of literacy programmes evolve over time? (b) To what extent does attending literacy programmes determine the acquisition of these skills? (UIL translation)*

As the research objectives and questions diverge, preparing questionnaires and using results is particularly difficult. The sections below illustrate this.

1. Analysis of national measurement instruments

The harmonized national skills descriptions

Each of the teams from the five countries that took part in RAMAA 1 produced a report on the results of their work on the harmonized skills description. They finished these documents in 2011, using a common matrix that the national team members helped to compile.

All the teams agreed to produce this skills description, but designing it proved complex. It involved building a framework aimed at a 'virtual' point of reference. This point of reference encompassed the profiles expected on programme completion, factoring in the specific contexts and synthesizing various components derived from different categories and dissimilar parameters, which were then supposed to yield measurable skill levels. Doing this involved gathering information on policies in effect, the programmes already under way and the measures taken to target participants (*Proposal for a methodological framework based on analysis of international and national RAMAA outputs until 2012, 2012*).

Unsurprisingly, the national reports on the skills descriptions attest to variations in the degree to which countries have taken ownership of the concept underlying this tool. Indeed, the skills classifications used in the harmonized national skills descriptions vary considerably, and depart somewhat from the instructions provided in the common matrix (see Table 4). This predictably foreshadows difficulties when building RAMAA's harmonized skills description .⁴

4 One of the project's original goals was to pool harmonized national skills definitions in an attempt to compile a common harmonized skills definition.

Table 4. Categories used in harmonized national skills descriptions

BURKINA FASO	MALI	MAROCCO	NIGER	SENEGAL
Skills: - instrumental - psycho-social - socio-professional	Skills: - instrumental (disciplines) - socio-economic - polyvalent ('transversal')	Skills: - instrumental - psycho-social - professional	Skills: - basic - management of economic activity - knowledge for everyday life	Skills: - instrumental - socio-economic - technical and technological - socio-emotional

The classifications in Table 4 cause problems. Psycho-social and/or professional skills, unlike instrumental skills, are specific to each country's context. Harmonizing them is therefore difficult (citizenship, for instance, may be very important in one country and much less so in another). Then, the notion of 'skill' as defined in RAMAA (see Chapter 3) implies that the questionnaire should put respondents in situations where they are required to solve problems. The only way to assess polyvalent skills is to ask respondents to deal with a situation. RAMAA's interview-based approach does not include this (*Report on the 2013 Workshop; Report on the item map review, 2013*).

This analysis shows how important it is to ask *what we want to measure* from the start. This question is crucial in action-research projects such as RAMAA because it shapes all subsequent development work on measurement tools, data collection, research questions, and result quality (in light of objectives).

In short, RAMAA 1 national teams submitted very well documented reports on their work building their harmonized skills descriptions. The methodology provided to produce these tools is nevertheless relatively complex, and this no doubt explains some of the variety of outputs.

Reference frameworks by country

The answer to the question about what we want to measure would have warranted more thorough analysis and confirmation during the skills description preparation stage, and in any case before building the reference framework and moving on to the operational stage with the measurement tests.

The 2014 technical workshop report underlines this concern, which had already been voiced during the international workshop in 2013:

Three questions about non-instrumental skills arise in the available documents and also emerged during the meetings with countries: (a) Are these skills included in the aims and syllabi of the literacy programmes we are reviewing in order to measure results? (b) Do participating countries really want to measure these skills? (c) Are these skills measurable in an assessment such as RAMAA 1?

The above, compounded with the lack of a harmonized RAMAA skills description and contingencies in countries such as Mali and Niger, led to differences in national skills descriptions. What was missing was a harmonized agenda for preparing and managing the measurement tools. That would have provided a 'cursor' for the skills description.

The reference framework, it is important to point out, is the tool used to specify the key skills that will be measured, establish the type of media that will be used to measure those skills, and factor the relevant contexts into the equation (*RAMAA Technical Paper, 2012*).

Burkina Faso, Morocco and Senegal, which had reached an advanced stage at that point in time, finally developed reference frameworks to test them during the field test. These frameworks covered (a) instrumental skills (reading, writing and calculating); (b) psycho-social skills; (c) socio-professional skills.

Table 5. Skills in three RAMAA 1 countries' reference frameworks

		Burkina Faso	Morocco	Senegal
INSTRUMENTAL		READING	READING	READING
		WRITING	WRITING	WRITING
		COUNTING	COUNTING	COUNTING
		SITUATING IN TIME WELL -BEING	SITUATING IN TIME WELL - BEING	WELL - BEING
POLYVALENT	Psycho-social	CITIZENSHIP	CITIZENSHIP	CITIZENSHIP
		SOCIAL SERVICES	SOCIAL SERVICES	
	Professional	INCOME - GENERATING ACTIVITIES	INCOME - GENERATING ACTIVITIES	INCOME - GENERATING ACTIVITIES
			TECHNICAL AND TECHNOLOGICAL	

Source: Report on the item map review, 2013

An expert's critical analysis of these outputs – at a much later stage – led to a series of pointers for progress. The instrumental skills, which are exactly the same in every country regardless of language, were rightly selected as a common baseline. The polyvalent skills, on the other hand, were reclassified as knowledge and four common topics, overarching country contexts, were identified: health and well-being, the environment, citizenship and work.

The next question is therefore *how to assess the selected key skills*. The instrumental skills can be measured by asking participants to deal with a situation, as per the RAMAA definition. Measuring reading skills, for example, involves asking a person to read in order to understand and carry out a task in a given context. As for measuring knowledge, assisted mode – i.e. asking the person to reply orally to instructions – appears to be the most suitable approach because it sidesteps a person's possible reading and writing difficulties, and focuses specifically on their ability to convey knowledge about the various topics under review.

To specify the development of each acquired skill along a continuum, the teams in three countries – Burkina Faso, Morocco and Senegal – used three interdependent skill levels to prepare the reference frameworks they used in their field tests (see Table 4).

Table 6. Proficiency levels identified by Burkina Faso and Morocco

Elementary	Learner is aware of the basics, e.g. spells correctly, deciphers sentences, calculates
Advanced	Learner applies the target skill in practice but is unable to use this skill in everyday life.
Transfer	Learner can transfer or use the skill in a relevant situation, serving a purpose in everyday life, in a manner denoting self-reliance.

Source: Report on the item map review, 2013

The team in Senegal essentially defined the proficiency levels along the same lines as the ones in Burkina Faso and Morocco, and labelled them 'Introductory,' 'Consolidation' and 'Advanced' respectively.

The descriptive proficiency levels that the teams in the three countries developed, and the continuum that they built into their scales, would have entailed issues during subsequent assessment because a heavy subjective component could have clouded the way

in which tasks were analysed. These levels were therefore revised and assigned numbers (Level 1, 2 and 3) instead of names.

It is also difficult to assess polyvalent knowledge on a scale from least to most difficult. Naming the causes of a disease is not, in itself, harder than knowing the rules for protecting the environment. It all depends on what a person has learned. Cross-referencing the levels of instrumental skill development and the contexts of polyvalent knowledge development is one way of sidestepping this hurdle. This approach (a) takes into account the contexts surrounding instrumental skill acquisition and assessment; (b) sheds light on the contexts surrounding the acquisition of the cultural knowledge that is desirable in that given setting (*Report on the item map review, 2013*).

The teams in Burkina Faso and Senegal were able to make these adjustments, and thus use a robust reference framework to prepare their questionnaires and then conduct their main assessments in a 'suitable' manner. Morocco, on the other hand, had already started its main assessment at that point in time and was therefore no longer in a position to make changes.

The teams in Mali and Niger used the experiences from the countries that had already reached more advanced stages, in their field tests.

National questionnaires

RAMAA questionnaires fall into one of two categories – measurement tests or background questionnaires – depending on the research questions they contain. When we reviewed country tools, however, we realised that they did not address literacy programme variability, meaning they did not include interviews with local managers and trainers, as prescribed in the *Technical paper, 2012*:

The discussions about research questions and sample design lead to the following options for the survey's protocol:

Version 1: only conduct the survey among a representative sample of graduates from the most recent adult literacy programmes (Test section and Background Questionnaire section);

Variation 2: Variation 1, plus a sample of local managers and/or trainers, etc. in the Background Questionnaire.

Version 1 is required. All other versions enhance Version 1. These enhancements are accretive. The enhancement in Variation 2 provides clearer insights into factors causing variability by gathering information from other key players in the literacy process (local programme managers, trainers, etc.).

As RAMAA's reference framework had not yet solidified as much as other frameworks, work focused on trying to develop 'common' questionnaires by making minor modifications in each country. Generally speaking, the key guidelines for doing so were as follows:

- ◆ The team in each country needs to choose what it wants to measure in terms of knowledge in the four identified topics (health and well-being, environment, citizenship and work) and to what level it wants to measure them.
- ◆ They need to refer to situations that can be illustrated and used effectively to assess instrumental skills and knowledge.
- ◆ They need to word questions or assign tasks that reflect as closely as possible the way in which respondents use the associated knowledge and skills in their everyday lives.
- ◆ We agree that oral evaluation is the most suitable approach to assessing knowledge of the topics identified by the team in each country.

An example of a question to measure literacy

Affiche 1



Affiche 2



Affiche 3



- Pour se prémunir du paludisme, il faut dormir sous une moustiquaire.
 A
- Avant de manger, il faut se laver les mains au savon
 B
- Allons au dispensaire pour nous Soigner
 C
- Un enfant bien nourrit est en forme
 D

Translation

Instructions

The health committee has put three posters up in the clinic.
Write the number of the poster in the box by the corresponding sentence.

- A: Sleeping under a mosquito net prevents malaria
- B: Wash your hands with soap before eating
- C: We have to go to the clinic for treatment
- D: A well-fed child is healthy

An example of questions to assess knowledge

Tell the participant: we are going to ask you questions about malnutrition, the environment, citizenship and income-generating activities. You can choose more than one answer. If you don't understand the question, tell me and I will ask you again.

NUTRITION

Instructions for the assessor: circle the letters by the participants' answer.
The boxes to the right will be used later for codification purposes.

Question 101: Which of these signs suggest malnutrition?

- | | | | |
|-------------------------------------|--------------------------|----------------------|--------------------------|
| A. Slow growth | <input type="checkbox"/> | B. Weight loss | <input type="checkbox"/> |
| C. Abdominal edema (fat belly)..... | <input type="checkbox"/> | D. Hair loss | <input type="checkbox"/> |
| E. none of the above | <input type="checkbox"/> | | |

2. Analysis of collection tools and collected data

In the countries that completed the project (Burkina Faso, Morocco and Senegal), 'data' refers to the data recorded during the main assessment. For countries that were unable to complete RAMAA (Mali and Niger), 'data' refers to the data recorded during the field test. This difference provided useful insights into the type of tools used and into the data collected during these two phases, despite the fact that the countries and contexts differed.

The target population and sampling frames

The principal objective for RAMAA 1 was to measure skill levels among adult literacy programme graduates in five African countries. The goal, ultimately, was to assess literacy programme quality and purpose, based on the levels of skills that graduates had attained and in light of graduates' contexts.

The population of interest in this study, it follows, did not encompass the entire population in any of the countries involved in RAMAA. Research was restricted to adults and, specifically, to two categories of adults:

- ◆ Adults who had taken one of the courses available from literacy programmes.
- ◆ Adults who had never taken literacy courses and who can be considered illiterate on account of their level of skills.

The people in the population of interest were eligible or not for inclusion in the target population depending on the year they took their course (see Part 2). This first selection process is relatively clear-cut. The second step, however, raises serious methodological questions because different countries – and different areas within those countries – define illiteracy in different ways. In Burkina Faso or Senegal, for instance, people are considered illiterate if they cannot read or write a simple sentence in any language. In Niger, the language is included in the definition because it only includes national and official languages, even though Arabic is also in use. In Mali, on the other hand, illiterate people are those who have never been schooled in any language. UNESCO's definition, lastly, focuses on function and sidesteps language: 'A person is literate who can, with understanding, both read and write a short statement on his or her everyday life.'⁵

The second cohort in the population of interest is therefore too variable to fit a working definition, and even more so as the participants' degree of illiteracy is only assessed based on the same participants' own views and the survey team member's observation, which heighten selection-criteria variability even further.

⁵ UNESCO, EFA Global Monitoring Report, 2006, Chapter 6: <http://unesdoc.unesco.org/images/0014/001416/141639e.pdf>.

a) Definition of the target population

A clear and accurate definition of the target population is required to ensure the target of interest is covered properly in each participating country, and to ensure countries produce consistent results. RAMAA's target population comprises all adults (age 15 or above) who are living in the country at the time of the survey, and have taken a literacy course during the year under review (2011, 2012 or 2013 depending on the country) or during the two or three previous years. This target population encompasses all graduates, regardless of their nationality. The people who will be included in the principal sample will be selected among the ones who a literacy course during the year under review. The people who will be included in comparison group 1 will be selected among ones who took a literacy course during one of the other two years. People who took literacy courses during other years are not included in the target population.

Most target populations, however, do not span every geographic area in participating countries or all adult literacy programmes available during survey years. The only exception was Morocco, where the teams conducted a representative survey covering every one of the country's 16 regions.

Lastly, literacy programme language was another criterion used in the definitions. For instance, teams selected adult programmes in four languages in Burkina Faso, in three languages in Senegal, and in one language in Mali and Morocco.

Table 7. Two characteristics of target populations – details by country

Country	Geographic areas	Languages
Senegal*	Dakar, Diourbel, Kaolack and Kaffrine	Serer, Pulaar, Wolof
Burkina Faso*	Not specified	Mossi, Dyula, Gourmanchéma, Fulfulde
Morocco*	Entire country	Not specified
Mali**	Kati circle (Koulikoro Region)	Bambara
Niger**	Dogon Dutsi and Dosso	Hausa, Zarma

**Data from the main assessment*

*** Data from the field test*

As we have seen, providing a clear and comparative definition of the functionally illiterate target population, to scope comparison sample 2 respondents, was not really possible. Moreover, national reports provide no clarification on this population.

b) Sampling frames and coverage of the population of interest

The sampling frame is the list from which the sample is selected. The quality of the sampling frame, in other words, directly affects the quality of the sample. This list must also contain the information required to adequately carry out the sampling stage, data collection stage and post-collection processing (weighting and analysing non-response) stages.

i) Sampling frames

The 2013 *Guidelines on sample design* provided participating countries with the following guidelines for preparing their sampling frames:

The sampling frame for the principal sample and comparison group is the complete list of graduates who have completed literacy programmes in each of the two years under review. Ideally, sample representativeness would be optimal if this list was complete and included all the stratification variables (each person's age, gender, type of literacy programme) and other necessary information to conduct the survey (each person's address). Likewise, random selection would be very simple if the list were available as a digital database.

In reality, databases are never both (complete and available in digital form). As soon as they adopt the sample design, national teams therefore need to start preparing the sampling frame focusing primarily on the principal sample. The samples they build will be lists of people's names along with their stratification variable and 'physical' location.

All countries indeed worked with a list of literacy programme graduates for the years under review to assemble their principal sample and comparison group 1. Compiling these lists involved substantial – and at the same time crucial – research work by national teams. To do this, they first had to list the training programmes that qualified for inclusion in the survey and then centralize the data relating to the groups of graduates.

The team in Niger, for example, visited the 94 centres running adult literacy courses in the survey area. Then it used information from decentralized government agencies to compile a complete list of all the graduates from all the programmes. The team in

Burkina Faso was also able to compile a list of all FCB, A3F and REFLECT⁶ programme graduates, adding up to more than 4,000 people. Mali followed the same approach to produce a complete list for its field test, but restricted its scope to Kati circle. Collating all the information from the region's three teaching centres (CAPs, *Centres d'Animation Pédagogique*) provided this team with a list of all the people who had attended REFLECT, *Compétence de vie courante* and *Sanmògòya* programmes. They sourced this information from the *Académies d'Enseignements* (education authorities) which provided files containing individual information about each participant's age and gender, the programme they had completed, the year they had completed the programme, and the commune where they lived.

The team in Morocco chose a different approach, assembling its sampling frame in two steps. First, it listed all the programmes available in the country to establish the stratification criteria for the first selection level. It only compiled its list of individuals during the second step, i.e. after sampling communes using participant lists drawn up by the government offices in charge of literacy and non-formal education. This team also had particularly abundant ancillary information: each person's identity, address and gender, the type of programme he or she had done, and the year he or she had done it, were included in the final list.

Table 8. Sampling frame sizes and years considered in each country

Country	Year	Number	Years	Number
Senegal*	2013	2,343	2011 and 2012	617
Burkina Faso*	2013	Not specified	2011 and 2012	Not specified
Marocco*	2012	No central base	2008 to 2011	No central base
Mali**	2013	1,444	2010 to 2012	1594
Niger**	2013	150	2010 to 2012	75

**Data from the main assessment

** Data from the field test

6 In Burkina Faso, the skills transfer system is organized into two cycles: (a) the first cycle, referred to as the literacy cycle (cycle d'alphabétisation) in turn comprises two levels: initial literacy (AI, *Alphabétisation initiale*) and basic complementary training (FCB, *Formation complémentaire de base*), which last 300 hours each (600 hours in total); (b) the second (optional) cycle comprises training in scientific and technical topics (CST, *Culture scientifique et technique*), fundamental and functional French (A3F, *Apprentissage du français fondamental et fonctionnel*), REFLECT, and specific technical training (FTS, *Formations techniques spécifiques*), which last 600 to 1,200 hours each.

As Table 8 shows, the final size of the lists varied from one country to another but, excluding Morocco – where the team did not compile a central base – sampling frames contained around 3,000 people in each of Mali and Senegal, and around 200 in Niger.

ii) Non-coverage rates

Non-coverage rates equate to the portion of the target population that is not included in the sampling frame. The main goal here, in order to generate high-quality data, is to shrink non-coverage as much as possible or, failing that, quantify it as accurately as possible while sourcing as much information as possible about the non-covered population.

In Morocco, all types of literacy programmes were taken into account. The sampling frame therefore comprised the programmes provided by ministries and other public-sector operators, and by businesses and NGOs. The team selected the provinces/prefectures (first tier), then the communes (second tier), following which it assembled a complete list of people who had graduated from the literacy programmes in those areas during the selected years.

We do not know if the teams in the other RAMAA countries compiled full lists of all literacy programmes, and therefore cannot estimate coverage rates within the populations of interest.

The information supplied by the teams in Burkina Faso, Mali and Senegal does not specify if the sampling frame comprises all the programmes available during survey years or if they used specific criteria to short-list programmes that would be simpler to analyse (e.g. permanent programmes, officially accredited programmes, programmes receiving public funding).

The team in Niger, on the other hand, pointed out that it had listed all training courses available in the survey area. So, provided the individual information is accurate and the organizations supplying that information had recorded it diligently, it is safe to assume that the coverage rate in the survey area is excellent.

Sample design and sampling

a) Requested standards and probability of selection

The *Guidelines on Sample Design* (2010, revised in 2013) recommended, but did not require, three-stage sampling. The only mandatory clause follows:

Where the survey is concerned, the reliability and relevance of the results primarily hinge on how robust the sample design is. The design must ensure the sample represents the population from which it is derived.

'Represents' is the operative word here. This means that sample design should result in selection probability and thus entails weighting each person who is finally selected. This weight – which needs to be corrected at a later phase (during non-response processing) – is absolutely essential to appraise the margin of error in all subsequent estimates provided in the survey.

b) Country sample designs

Almost all participating countries produced probabilistic multi-stratum sample designs. The details, however, vary considerably from one country to another. Overall, the teams adopted one of two strategies: assembling a list of all graduates from selected programmes in the survey area, or, due to the associated time and cost constraints, restricting their list to the geographic areas they had selected during the first stratification stages.

From a statistical standpoint, the first approach is more advisable. It leads to more robust, non-biased estimates and limits sampling errors.

The second approach, on the other hand, assumes that initial stratification captures literacy programme diversity (graduate profiles, types of programmes, number of learners, etc.) none the less. Unless additional research substantiates this assumption, there is nothing suggesting it is true per se. That said, this approach is entirely reasonable and understandable when it is a case of adjusting the theoretical requirements of sample design to fit the practical need to limit data collection costs.

i) Sampling methods

Sampling methods vary considerably from one country to another (see Table 9) and mirror different approaches. These approaches generally reflect attempts to juggle three constraints: the need to limit collection costs, the wealth of information available in the sampling frames, and the intended level of accuracy in the key variables. The wide variety of methods – in terms of the number of strata, stratification criteria, choice of statistical units, etc. – perfectly illustrates all the possibilities available to the statistics teams.

The teams in Senegal, Niger and Morocco opted for three-tier sampling. The team in Morocco, for instance, selected 29 provinces, then 111 communes in those 29 provinces, and lastly 5,369 graduates in those 111 communes. The team in Niger chose 15 of the 94 centres for its principal sample and 15 others for its comparison group 1 during its first selection step.

The team in Burkina Faso chose to use two strata (programmes then graduates) and the one in Mali settled for conducting its field test with one stratum.

Table 9. Details of a few sampling plan components in each country

	Selection stratum 1		Selection stratum 2		Selection stratum 3	
	Statistical units	Stratification criteria	Statistical units	Stratification criteria	Statistical units	Stratification criteria
Senegal*	Villages	Regions	Programmes	Urban/rural	Graduates	Programmes, gender, age
Burkina Faso*	Programmes	Literacy courses	Graduates	Literacy languages, gender, age		-
Marocco*	Provinces and prefectures	Geographic area, urban/rural	Communes	Urban/rural	-	
Mali**	Graduates	Urban area, programme, gender, age (two ranges)	-	-	-	-
Niger**	Centres	Departments (two)	Villages	Centres	Graduates	Programmes, gender

* *Data from the main assessment*

** *Data from the field test*

The stratification criteria also vary a lot from one country to another. In general, and in keeping with the survey's guidelines, they are aimed at generating data according to a number of key elements (for example, distinguishing urban and rural settings, types of programme, and male and female graduates).

The documents from Burkina Faso do not shed light on the sampling method that the team there chose, other than the fact that it used quotas for sampling. The data appear to suggest it attempted to consider age, gender, training language and perhaps the type of programme in the second stage, and the programme cycle in the first stage.

ii) Sample sizes

Sample sizes usually depend on the target level of accuracy. RAMAA 1 sample sizes on occasion appeared surprisingly small (see Table 10) but it is important to compare them with the initial sampling frame. The sampling rate in Senegal, for example, is near 30%, which is particularly high. As a point of comparison, the number of people selected in Morocco and Burkina Faso is also particularly high: combined, the principal sample and comparison group 1 comprise around 5,300 people in Morocco and almost 2,000 in Burkina Faso, which puts the overall sampling rate at 37.5%.

Table 10. Sample sizes in each country

	Principal sample		Comparison 1		Comparison 2
	Year	Size	Years	Size	Size
Senegal*	2013	750	2011 and 2012	200	50
Burkina Faso*	2013	1438	2011 and 2012	518	148
Marocco*	2012	3286	2008 to 2011	2083	1091
Mali**	2013	661	2010 to 2012	203	50
Niger**	2013	150	Not provided	75	75

** Data from the main assessment*

*** Data from the field test*

However, small samples (numerically speaking) and very high survey rates can lead to excessively complex sampling methods in certain cases – in Senegal especially. It would no doubt make sense to follow simpler approaches, which should be more attuned to the sampling frames, if the RAMAA programme moves on to its next phase.

iii) Sample quality control

Once a sample is selected, an additional process to ascertain that the sample's quality is suitable usually ensues. This process involves a number of fairly simple steps to confirm that the characteristics of the selected population match the characteristics that the statistics team had in mind when it prepared its sample design. In practice, this can mean checking the male/female ratio and age distribution by systematically comparing the raw sample, weighted sample and reference population. The teams in participating countries no doubt followed this process but very little information about quality control is available in their reports. The team in Morocco is the only one that mentions certain verifications (selected individuals' place of residence and age group, stratum representativeness, etc.) but does not provide results.

iv) Comparison group 2

Comparison group 2 warrants separate discussion at this point in the project. As it is impossible to assemble a sampling frame comprising the country's entire functionally illiterate population, the preparatory documents allow considerable latitude with regard to this comparison group. Almost all participating countries chose to apply a non-proba-

bilistic selection method. The technique involved finding a way of identifying individuals who could be considered illiterate based on their own declarations. Most often, the teams opted for a qualitative approach, looking around at some of the people who were already in one of the two other cohorts and for instance selecting their neighbours or relatives. These approaches make sense and limit collection costs. Unfortunately, however, they are biased by inevitable cluster effects and the subjective perspectives of the people tasked with selecting 'suitable' neighbours or relatives.

Table 11. Comparison group 2 selection by country

Country	Selection type	Selection method	Sample size
Senegal*	Non-probabilistic	Graduates -> household members (reasoned choice)	50
Burkina Faso*	Non-probabilistic	Not provided	148
Marocco*	Non-probabilistic	Not provided	1,091
Mali**	Probabilistic	Villages -> households -> individuals	50
Niger**	Non-probabilistic	Reasoned choice	75

* Data from the main assessment

** Data from the field test

During the field test, Mali was the only country that tendered a probabilistic sampling method: it established a three-tier, areal and stratified selection process based on the list of places of residence of the individuals in the principal sample and comparison group. There are three advantages to doing this. First, as a sampling frame is unfeasible, systematic selection sidesteps the survey team member's judgment and limits bias caused by cluster effect. Second, this approach awards respondents probability of selection (and hence weight in the survey), which is needed to estimate results representing the given population (even though it is important to avoid overestimating accuracy here). Third, it limits collection costs by selecting individuals in geographic areas that the survey will already cover with the two other samples.

c) Recommendations for future programmes

The analysis of sample designs and sampling methods leads to a number of recommendations for possible RAMAA programmes in the future:

- ◆ Before the selection process begins, it would make sense for UNESCO to liaise with participating countries in order to establish uniform standards. These stand-

ards could relate to sampling techniques (limiting selection strata, defining statistical units, producing statistics on sampling method quality and accuracy) and sampling quality (precise descriptions of sampling frames, detailed documentation on the various phases). An outside expert could also review sample designs to ensure they meet these standards before the survey proper starts.

- ◆ The UNESCO Institute for Statistics could be an invaluable resource to establish survey standards and help countries overcome the hurdles they may encounter when they attempt to meet those standards.
- ◆ It would also make sense to redefine the survey's population of interest, perhaps by restricting it to literacy programme graduates. The rationale underlying the inclusion of a comparison group 2, understandable though it is in light of RAMAA objectives, is not satisfactory from a statistical standpoint.

Data collection process and quality

a) Data collection

i) Collection time-frame

The teams in participating countries were all supposed to collect data simultaneously, but various developments unrelated to the project delayed this stage in all of them except Morocco.

Collection time-frames depend on many factors, in particular the size of the geographic area, the number of survey team members available, the workload each of these survey team members is assigned, the quality of the information in the sampling frame (people's addresses), the quality of the identification procedures, how easy or otherwise it is to access the survey areas, the management of the survey team network and, lastly, the size of the sample.

The countries that only completed the field test collected data within relatively brief time-frames: the teams in Mali took two weeks (31 July to 14 August 2014) and those in Niger one week. Data collection for the main assessment naturally took longer because the samples and geographic areas were larger. The team in Morocco, for instance – the only one that supplied precise data – took six weeks to conduct all its interviews. It is important to add that this time-frame is remarkably short and shows that the team made considerable efforts to organise its field operations to a tight deadline, which it met.

ii) Response rates

The response rate – meaning the portion of the selected sample that was effectively surveyed during the collection operation – provides the first indication of data collection quality. The framework document did not set a specific response-rate target, but it is

commonly accepted that around 70% or more will yield suitable data quality in household surveys.

Remarkably, most of the documents supplied by the teams in participating countries do not mention the response rates they aimed to achieve or, more importantly, the response rates they obtained by the end of the collection phase. This ties in with the lack of details on the exact number of people who were actually interviewed and the lack of information about the collection process in the data tables.

For example, the only information that Niger provides as an indicator to assess response rates is the number of people 'met', namely 147 (out of 150) in the principal sample, 73 (out of 75) in comparison group 1, and 75 (out of 75) in comparison group 2. The teams in Senegal and Burkina Faso report the final number of people they interviewed but do not provide the number of people they selected for their samples. In Senegal, for instance, the team interviewed 688 people in the principal sample, 184 in comparison group 1, and 37 in comparison group 2, suggesting it overshot the 50% response-rate target it set in its sample design. However, we do not know who was not interviewed or why. Elsewhere, the information from Burkina Faso on occasion lacks precision: the report states that the team interviewed 1,442 people in the principal sample and 520 in comparison group 1, whereas a separate section in that report states that those samples respectively comprised 1,438 and 518 people before collection.

Mali is the exception to this rule: its team's dataset provides the number of people selected per stratum, the number of respondents and the response rate per stratum. The problem is that the figures do not tally: on the one hand, 86% of the women and 99% of the men responded to the survey and, on the other, the report states that the overall response rate in the principal sample is 'only' 74%⁷. Similarly, in the case of comparison group 1, one column in a table states that 23 out of 203 people took the survey whereas the previous column states that the sample comprised 50 men and 154 women (meaning one is missing).

The data that the team in Morocco supplied is in a separate class altogether. Its report presents many very important quality indicators, for instance including respondent breakdowns by key variables such as gender, age and province. It also provides more precise information, which sheds useful light on the data collection process, for example including the rate of questions answered. However, the more general information about response rates is disconcerting. The Moroccan team decided not to settle for the samples it had selected beforehand. Instead, it dug into a reserve sample, sometimes to replace non-respondents in the original samples and more often to recruit new respondents. It indeed seems to have recruited new respondents almost systematically.

⁷ Apparently, data on gender are lacking for 155 selected people. This either means that this information was not included in the sampling frame or that the collection data did not match the sampling data.

The Moroccan report thus states that the response rate of the city-dwelling cohort in the province of Larache is 240% and that that of the city-dwelling cohort in the province of Zagora 185%. Such figures do not really make statistical sense because response rates, by definition, cannot exceed 100%. Tapping into a reserve sample to reach a critical number of respondents is indeed common practice in household surveys, but a number of very strict rules apply as regards measuring the biases that new respondents bring into the equation. If those rules are not followed, it is impossible to gauge distortion due to non-response or accurately calculate the survey weight of the statistical units in the database. On the other hand, adding respondents without a valid reason, and thereby departing from the original sample design, is a practice we consider questionable. In future, it would therefore be preferable to strictly supervise compliance with sample design, provide common definitions of response rates, and always bear in mind that we need to obtain survey weights once collection is complete.

iii) Survey team training

Establishing a tailored, quantified and standardized training programme for survey teams is a sine qua non to collect high-quality data. The survey team needs to fully master the survey procedures in order to conduct interviews in the best possible conditions and to reduce collection errors insofar as possible. When survey teams are comfortable with the procedures, moreover, their motivation levels are likelier to remain high and response rates are likelier to follow. Lastly, conducting surveys aimed at measuring skills requires a particular knack: survey teams need to elicit information from people who have scarcely been schooled if at all, and who may be reluctant to expose their shortcomings before an enquiring stranger.

Generally speaking, it is advised to:

- ◆ Train survey teams shortly before field operations begin.
- ◆ Train them in small groups.
- ◆ Appoint teams of trainers specializing in each of the topics required for the various stages in the survey.
- ◆ Adequately address all aspects of the survey (locating interviewees, managing rounds, background questionnaires, skills questionnaires, conducting interviews, etc.).

While Burkina Faso and Senegal did not provide details on their training strategy, the other countries supplied some very useful information showing their solid grasp of the required training standards. Niger, for example, organized a four-day training course (from 5 to 8 December 2013), led by three trainers. The team there then conducted a

pre-survey test, enabling trainees to put their instructions into practice and to make adjustments if and as required.

The team in Morocco called in experts to run the training sessions. It also provided a handbook for survey teams and specific training for supervisors, both of which bolstered the data-collection team's efficiency and effectiveness.

The team in Mali, lastly, provides an interesting angle on the strategy to optimize quality in its survey-staff network. It worked with the network of professional experts at INSTAT (the country's statistics institute), who already have extensive fieldwork experience, took tests at the end of its seven-day training courses and only kept the survey team members who displayed the sharpest skills to conduct interviews, locate individuals in samples and manage individual data (codification, filling in questionnaires, forwarding questionnaires, etc.).

Ideally, the strategies that Niger, Morocco and Mali followed should complement each other and all participating countries should apply them. They should also be systematically documented and encouraged by UNESCO. The fact that these strategies are already available and, in certain cases, have already been used, provides a very useful starting point for a possible RAMAA campaign in the future.

b) Survey network and data collection monitoring

i) The number of survey team members

The number of field workers is an important variable, which countries keen on maximizing data-collection quality will want to take into account. It is always necessary to juggle collection time-frames, sample sizes and survey team sizes. Survey teams are also the main expense item in any survey implementation budget. A sensible yet substantial effort is absolutely essential to avoid curtailing data quality and breadth before collection even starts.

In general, the participating countries made the necessary investment. Mali and Morocco, for instance, each used 24 survey team members. Niger used 16 for its field test: half from the country's Institut National de la Statistique (INS) and half from the ministry working on this project.

ii) Managing survey staff

If a survey team is to work efficiently and effectively, it needs a sufficient number of seasoned professionals to supervise it. Among other duties, supervisors organize data collection (assigning interviewees among survey team members), monitor collection progress, oversee collection work and collected data quality, and forward filled-in questionnaires to the central office.

Some participating countries did not supply any indicators shedding light on the quality of the survey teams' work, such as team-supervisor/team-member ratios or how often supervisors and team members liaised. The teams in Niger and Morocco, however, provided some information of this kind: they respectively had four and six supervisors, i.e. one supervisor for every four survey team members in both countries. Morocco added geographic divisions: the team there grouped the survey areas into six geographic zones, and appointed six groups each comprising four survey team members and one supervisor.

Non-response bias

a) Supporting the survey

To minimize the bias due to the fact that some people will be unwilling or unable to take the survey, it makes sense to map out a strategy geared to maximizing response rates before starting the survey and, more broadly, to kindling people's interest in the survey. The key, here, is to communicate actively and effectively with public services, local authorities and the target population.

Two countries appear to have built particularly successful communication strategies. The team in Mali reached out to national authorities and resource people in the local area who could play a role facilitating fieldwork (helping survey team members to locate respondents, encouraging them to respond, etc.), and told them about the survey. It reports that government services only provided limited support but that assistance from resource people was crucial to collection operation success during the field test. The team in Morocco provides another good example of an initiative to promote the survey before fieldwork began: it rolled out an extensive campaign to involve, inform and raise awareness among local and national authorities as well as the operators running the selected programmes. This awareness campaign moreover reached the micro-local level via agents d'autorité, i.e. local civil servants vested with public authority.

b) Documenting collection

The *Guidelines on sample design* provided very clear warnings on a few of the issues that teams might need to address during the collection phase:

In the case of the two samples comprising graduates (the principal sample and comparison group 1), the interval between the date graduates completed their course and the time the survey is conducted leads to the question about the selected individual's physical whereabouts. It is therefore extremely important to bear in mind the issues associated with locating individual graduates when planning the survey. This involves (a) considering the field team's work in 'locating' selected individuals and (b) arranging additional samples to replace selected individuals who, for various reasons, cannot be located at a reasonable cost.

It is essential to document the way in which collection operations unfolded in order to shed light on the difficulties hampering collection operations in the field. The team in Morocco is the only one that shared detailed information on this issue. The field workers there conducted their interviews in tandems and their interviews lasted 45 minutes each on average, which is optimal to keep respondents interested while shortening collection time-frames and limiting survey costs. A few of the selected interviewees were unavailable – which is a recurring issue that teams conducting household surveys know well – but, on the whole, the people in the samples rarely declined to take part. Lastly, from a very practical perspective, the team reports that the sampling frame from the public operators was not always up to date or complete.

It would be a good idea to complete these indicators and, especially, use them systematically in all participating countries. One would expect to see interview time-frames and methods clearly documented in each country context. One would also expect databases to include information about data collection process monitoring, i.e. which individuals replied or declined, which individuals the survey team was unable to contact or interview, and so on. Information on the circumstances surrounding the interview could also be included, such as the time, place and other people present. These recommendations tally with several concerns already mentioned in this report: information about the collection process would make it possible to calculate exactly what proportion of the original sample was interviewed, monitor and calculate survey-weight distortion, and track collection progress in order to fine-tune approaches to maximizing response rates and managing the survey team's work, and so on. In short, this information is absolutely crucial and the fact that it is missing is particularly regrettable for this RAMAA phase.

c) Documenting non-response

Documenting non-response is another important aspect of ensuring the data produced are reliable. There are several reasons why an individual originally selected for inclusion in a sample may not be interviewed. It is important to know why, in order to distinguish the people who should have responded (and therefore belong in the survey scope) and the ones who should not have been selected (because they do not belong in the survey scope). It is also important to know a number of things about non-respondents (age, gender, etc.) to make sure non-response does not alter the respondent population's final structure. Lastly, we need to quantify non-response in order to correct respondents' survey weight at the end of the process.

Niger was the only participating country that provided satisfactory – but not detailed – information (the number of non-respondents and their reasons).

Fieldwork quality

a) Data validation

The quality of the recorded data hinges to a great extent on the survey team member's diligence in noting down replies during interviews. This is why it is important that supervisors verify and validate the teams' fieldwork on a regular basis – in addition to the provision of suitable training for survey staff.

In Morocco, for instance, the validation process comprised various steps. Each day, the team leaders (or supervisors) verified the questionnaires that their survey teams had submitted. Each week, supervisors held meetings to update on progress in the collection process, with a view to helping survey team members to overcome any difficulties they had encountered in the field and to monitor their productivity. At a second stage, the head office manually checked questionnaires before keying replies into the information system, focusing especially on question completion rates.

The team in Mali produced a supervisor handbook to ensure supervisors were aware of the importance of validating data and to guide them through the verification process. Before forwarding the questionnaires to the head office, supervisors manually reviewed survey team members' work, checking the number of questionnaires they had filled in, and ensuring they were complete and, more importantly, consistent.

The teams in the other countries did not provide details on their processes.

b) Other verifications

Besides the quality of the data, teams verified the quality of the survey team's work and its compliance with survey guidelines.

Here again, the teams in Mali and Morocco are the two that explain their processes in most detail. The eight Malian supervisors were explicitly tasked with providing suitable working conditions for survey team members and with helping them to solve practical problems (reaching fieldwork areas, dealing with the logistics, filling in questionnaires, conducting interviews, etc.). In Morocco, the project manager and executives from the government agency in charge of fighting illiteracy monitored supervisors.

Quality assurance and control

a) Quality assurance

National teams were entrusted with building their own quality standards. Only one document – the *Template for the RAMAA survey procedure manual* – provided a few

guidelines on the documents that the national teams needed to produce throughout the survey process, namely:

- ◆ A set of guidelines on sample design
- ◆ A guide on item administration
- ◆ A guide on conducting the background questionnaire
- ◆ A guide on scoring
- ◆ A guide on codifying
- ◆ A guide on recording information
- ◆ A guide on training field workers
- ◆ An analysis plan

b) A guide on quality assurance

It would have made sense to send all these documents to the UIL. These documents are essential to understanding the quality of the data collected. The reports from the national teams do not really show which documents were produced in each country, with one notable exception, Morocco.

c) Quality control

Quality control, in this case meaning the analysis of the extent to which each participating country's project fulfilled the standards set down in the quality assurance plan, widely eluded the UIL. Each country developed its own quality assurance plan – or not.

Data management and processing

a) National-level and international-level responsibilities

Because the questionnaires were not harmonized across countries, and because the UIL team assigned to RAMAA was short-staffed, data management and processing were delegated entirely to country teams. Guidelines required that tables be submitted in formats that could slot easily into the software applications that statisticians typically use (SAS, SPSS or STATA), but provided no prescriptions on the variables or how to process them.

b) From data collection to import

i. Codifying data

The raw information supplied by interviewees sometimes needs to be codified (or at least harmonized) before it can be used for statistical analysis. This is especially the case for

open-ended questions and questions for respondent classification (place of residence, profession, etc.).

This issue is even more sensitive in the psychometric questionnaire because codification may involve correcting exercises and doing this requires specific training and experience. For this reason, Mali produced a guide on scoring, enabling four team members to process this information correctly. Niger tackled this issue by entrusting its experts with this task: the members of the team that prepared these items were the ones who manually scored the answers to the psychometric exercises.

ii. Extracting data

Extracting data, i.e. transferring information on an individual questionnaire into a collective digital file, is a strategic step:

- ◆ It provides an opportunity to make additional corrections when it becomes clear that the replies do not necessarily match the expected format or that there are some outliers.
- ◆ New errors often occur during this step if the extraction process is not properly supervised.
- ◆ Adding a format, heading and general structure produces a preliminary template to read the data.

The three reasons above explain why a rational extraction process and quality standards are necessary.

One of the methods that was used most often – in particular in Mali and Niger – involved programming input masks. In other words, team members entered handwritten replies via a software application that predetermined reply parameters (authorised replies, maximum and minimum admissible values), the format of variables (numeric or alphanumeric) and the order of questions. This way, the team cleaned up the data, which were then fit for export to SPSS format, enabling a statistician to review them in a suitable software application.

Morocco was the only participating country that developed a Common ISDN Application Programming Interface (CAPI). Survey team members used a computer to manage their rounds, key in replies, and forward data to the national team. This method automated data extraction and limited errors.

iii) Verifying data (duplicates, missing data, etc.)

Once the data are grouped in a digital file, it is possible to deal with errors more systematically. This clean-up step in particular involves tests to ascertain consistency. The team in Morocco

entrusted this step to an outside consultant, and the one in Niger to a statistician tandem. The teams in these two countries also checked completeness to ensure the database matched the information about the sample (the number and types of questionnaires).

Processing the data

a) Background questionnaire

Following clean-up, teams can perform several operations to make background-questionnaire data easier to read and enhance their consistency. The team in Morocco, for example, codified replies to open-ended questions during this processing phase (in particular replies to the question about the choice of literacy programmes). It also dealt with filter questions and anonymized the file.

b) Psychometric questionnaire

The methods for running psychometric data through statistical processing vary extensively. The wealth of available data, preferred psychometric theory, questionnaire characteristics and survey goals open up a huge array of techniques for national teams to calculate scores or proficiency in each skill, and to appraise the accuracy of score estimates.

Generally, the teams in all countries used Classical Test Theory, first of all to assess the questionnaire's psychometric quality. They provided statistics – varying in accuracy and complexity – that make it possible to calculate overall test homogeneity (Cronbach's alpha), question difficulty (success rates) and consistency (discrimination coefficient). Then, most country teams estimated individual scores, in most cases using a simplified two-parameter model based on Item Response Theory.

Unfortunately, the reasons why this method was chosen were never provided. This leaves several questions unanswered. (What is the point of trying to provide a continuous score for a skill? What does that score really mean? Where are the meaningful thresholds to assess an interviewee's level of proficiency?). Without the answers to those questions, scores remain abstract data. Participating countries should confer and agree on a unified method to process psychometric data and adopt a consistent grid to interpret data, explicitly based in a theory (or skills reference framework).

The team in Niger is the only one that did not become entangled in excessive complexity. It decided to calculate individuals' scores in each of the three skillsets by adding up the weighted correct answers. It would be interesting to understand why it applied certain weights to certain questions, but this method has the advantage of being clear and easy to understand for people outside the project who lack specific expertise in statistics and nevertheless wish to use the study's findings.

Finalization and publication

a) Data transmission

The goal for an international survey such as RAMAA, by the end of the project, is to feed into multilateral discussions and analyses on the basis of known, shared and comparable databases. To do this, it is essential to circulate results from all project participants to the various stakeholders.

Then, the obtained data needs to be easy to understand in order to facilitate discussion. This means that a number of documents, enabling users to understand the database properly, should be provided. These in particular include:

- ◆ A dictionary of codes providing names for the variables and describing the types of replies;
- ◆ A presentation of the questionnaire;
- ◆ A technical report explaining the way in which the database was created and built, the underlying objectives and the database's limits;
- ◆ A series of variables shedding light on the quality of the data obtained (survey weights, interview times).

It would also be useful to format databases to the same model, even if the questionnaires vary from one country to another. In other words, insofar as possible, it is preferable to give similar variables similar names, use the same formats for replies (numeric, alphanumeric, etc.), deliver databases in standard formats (SPSS, SAS, etc.) and so on. The original plan included a deadline for submitting data to UIL-UNESCO, but local difficulties considerably disrupted the tentative timeline and all participating countries but Morocco were unable to deliver on time. At this point, only field-test data is available from the five participating countries. Three countries have also completed collection on the main assessment and submitted all their data.

b) Country report publication

National teams were solely responsible for their country reports (as they were for their databases). This explains why the nature of the information in each of those reports varies so widely, as discussed in this document. Approaches to promoting results also varied considerably.

The participating countries published their reports and submitted them to the UIL at different times (see Table 12).

Table 12. Country report publication months and years

Country	Published
Senegal	September 2014
Burkina Faso	January 2015
Morocco	November 2013
Mali	February 2015
Niger	December 2013

When they published their reports, national teams also organized a variety of events to disseminate their findings. That was the case on Morocco especially, but also in Niger, where the team convened a seminar from 14 to 17 February 2014 for public officials, literacy specialists and stakeholders, and the RAMAA team, to discuss progress and map out the road forward for the fight against illiteracy in that country.

General results

Given the reservations expressed throughout this document, relating to the lack of explicit criteria informing opinions about the quality of data and, especially, the lack of survey weights, which prevents calculation of the representative data and estimation of their margins of error, this section of the report will be succinct. It will primarily seek to extract a few major trends based on the analysis of results from participating countries. These initial trends are provided as pointers for additional research and investigation rather than precise facts and figures.

a) Outcomes among adult literacy programme participants

Comparing results is difficult because the teams in some countries did not restrict themselves to assessing the three fundamental areas and, more importantly, because the techniques they used to estimate individual skills and the scales describing proficiency in those skills vary from one country to another. We will therefore restrict results to three areas: reading, writing and mathematical reasoning.

In Morocco, the psychometric questionnaire contained 12 questions, i.e. 3 in each area (and a fourth area, 'situating in time and space', which we do not analyse here). The questions in each area are arranged from the least to the most difficult: the first denotes 'elementary' proficiency (3 points), the second 'advanced' proficiency (3 points) and the third

'transfer' proficiency (4 points)⁸. Graduates are deemed proficient in the skill if they score at least 7 out of 10. The data suggests that adult literacy programme graduates obtained high scores for reading but not for writing (76%) or mathematics: around 70% displayed 'transfer' proficiency in reading, whereas around 50% did so in writing and mathematics. Adults who had taken additional courses – post-literacy and the like – achieved substantially better scores.

The team in Burkina Faso used this same three-level (elementary, advanced and transfer) scale.

Senegalese learners who graduated during the survey years are particularly proficient at reading and their reading skills remain shaky. According to the country report, the fact that this society traditionally favours the spoken word and only uses writing sporadically could explain this gap. The report also posits that the decent scores in mathematics reflect the fact that numeracy skills are more frequently reinvested in everyday duties such as managing households or small manufacturing and trading operations. These general results, however, conceal very large disparities: 30% of the principal sample does not yet qualify for level 2 ('attaining literacy') and only 50% qualify for level 3 ('literate').

For the purposes of its field test, Niger established three levels ranked according to the percentage of correct answers provided in reading, writing and mathematics (0% to 50% is unsatisfactory, 50% to 60% is satisfactory and 60% to 100% is proficient). We only have overall field-test results, not results in each of the three disciplines. It appears that, even though they have completed literacy programmes, more than 70% of graduates are unable to answer more than 50% of the questions correctly. These figures are nevertheless considerably higher than those in the functionally illiterate comparison group. The report's authors ponder these below-average results but do not really explain them. It is moreover impossible to determine whether these results reflect disparate standards in literacy-programme quality or flaws in the measurement tool.

b) Outcome sustainability

The question as to whether literacy programmes provide baseline skills that learners will be able to use enduringly in their everyday endeavours is central in the RAMAA initiative. The answers appear to vary from one country to another.

⁸ There is some inconsistency here: according to the report, a graduate is expected to score least 7 points out of 10, i.e. attain 'Advanced' level. Based on the schedule provided, however, the points in the first two levels add up to 6, not 7. The '7 points out of 10' threshold is therefore difficult to interpret. We do not provide the figures here.

The survey in Morocco – the only one that compared graduate performance over the past five years – provides interesting perspectives for analysis. The country report states that 63% to 67% of the adults who took literacy courses scored 28 points overall out of 40 (the threshold for basic skill acquisition). The scant score variability – which may be non-significant – suggests that basic training courses results are sustained over time. In Senegal, graduates in the principal sample obtained slightly higher scores than those in comparison group 1. Initial analysis may suggest that there is some loss in reading, writing and mathematical skills, but that this loss is small. This in turn suggests that, under certain conditions – which are as yet unclarified – participants retain their skills one or two years after completing their training courses. However, lacking a corpus and more specific information about developments in literacy programmes over the past two years, two other possible explanations warrant consideration:

1. This change may reflect changes in the quality of the available programmes;
2. It may reflect changes in the profiles of the people who are recruited to do these programmes.

Ultimately, only half the people in the comparison group displayed level-2 or level-3 proficiency (and therefore know how to 'read and write simple sentences and texts, and solve problems'), versus two-thirds of the graduates in the principal sample. In Burkina Faso, unlike Morocco and Senegal, graduates in the earliest cohort (2011) scored slightly better results than those in the most recent one (2013) and much better results than those in the intermediary cohort (2012).

The case in Niger is somewhat particular: 36% of graduates in comparison group 1 achieved level-2 or level-3 scores, versus only 28% of graduates in the principal sample. We do not know the statistical significance of this gap, which may be due, in part, to bias in the principal sample (certain centres had not finished their programmes when data collection started). In any case, if it is confirmed, it suggests that (a) the graduates who had completed their training course longest ago have since consolidated their skills; (b) the circumstances surrounding the most recent courses were less favourable (interruptions due to the domestic situation, the fact that certain operators have disappeared, changes in graduate profiles, etc.); or (c) the quality of the programmes does not yield comparable outcomes.

The question about outcome sustainability, in other words, is nowhere near settled yet. The lack of estimates relative to margins of error, the lack of more elaborate statistical calculations (inferential and the like) and the variability in findings incite us to delve deeper into the conditions that foster the acquisition of sustainable and transferable instrumental skills.

c) Are gender and age two key analysis variables?

Gender and age classically appear as key analysis variables in surveys aimed at measuring skills. However – and even though the lack of margins of error and of survey weights bar us from making a categorical statement to this effect – RAMAA results do not appear to confirm these ‘constants’ in international surveys.

In Senegal, age does not appear to be a significant variable in the analysis of results, and graduates’ genders were not accounted for in the various tables submitted.

The report from Morocco, conversely, shows that these two variables were used throughout results analysis. It provides three age groups (under 30, 30 to 50, over 50) and shows that a larger number of younger people reach the baseline level. However – and this is somewhat original – age does not seem to influence outcome sustainability. Gender seems to be even more dividing than age: in the groups that had graduated earlier than the others, the percentage of men who reached the baseline level is 22 points higher than the percentage of women who did. This gap is narrower among groups that graduated more recently, but due to a decline in the male population’s results rather than an improvement in the female population’s results (especially as the proportion of men in the principal sample and comparison group is very low). Ultimately, this context-related effect no doubt reduces the statistical significance of this gap. In any case, gender does not – according to these results – weigh on outcome sustainability.

The team in Niger did not use age as a variable for analysis but did use gender, and its findings show a fairly clear-cut difference between the male and female population. There are fewer women in the level-1 group and more in the level-3 group, suggesting that literacy programmes are slightly more beneficial for women. This result is particularly interesting if we compare results based on centres’ policy on mixed-gender training: graduates from all-women and all-men centres scored better results than those from centres that train women and men together.

d) Literacy-programme quality

As RAMAA’s primary focus is on assessing and improving literacy programmes for the adult population, it is important to ask if RAMAA 1 has yielded any compelling conclusions on this issue.

In Senegal, results vary widely according to the region where graduates live. This applies to the principal sample and to comparison group 1. Adults in the Dakar area achieved considerably higher scores than those in Diourbel and Kaffrine, who in turn obtained much higher scores than those in Kaolack. Score distribution according to other variables, such as literacy language (Pulaar, Wolof or Serer) or town category (commune, main town of rural community, village or hamlet), also reveals sizeable differences. The team in

Senegal concludes this first set of results with an operator ranking based on participant scores and posits that the operators that have performed the best are the ones that appear near the top of the list. This somewhat hasty conclusion should be compared with the previous results. Rather than cross-tabulating results, it should have used other techniques (regression analysis, multilevel modelling, component analysis, etc.) to correct for context-related effects. The operators towards the end of the list may, in fact, be the ones tasked with training the populations that compound the most disadvantages. If the team considers that it is important to establish a ranking, it would have been preferable to at least reason in terms of 'all other things being equal'. Lastly, the operators that have the most sophisticated organizational, technical, financial and logistical capabilities are the ones that train the people who live in the more urban areas, who play a more active role in the local economic and social fabric, and who speak Wolof or Pulaar rather than Serer.

In Niger, the sample is too small – the team only conducted a field test – to derive any direct conclusions as to the quality of available training programmes. That said, the results do reveal something that warrants closer examination: the centres that provide the most hours of training are not the ones that perform the best. This suggests the need for deeper analysis of literacy programme content (training for trainers, session arrangements, learner numbers, etc.).

The teams in Burkina Faso and Morocco compared population structures (language, age, gender, etc.). In Morocco, the various profile-related variables taken into account suggest that the operator has no bearing on literacy programme quality.

Conclusions

Ultimately, these few, rapidly examined results seem to confirm that it makes sense to tailor literacy programmes to the populations they target. It seems clear that an overly constrictive structure (same programme, same number of hours, etc.) is not the ideal approach to provide the entire population, in all its diversity, with baseline reading, writing and mathematical skills. It is also clear that completing a literacy course with a post-literacy course is one of the conditions for training programme success. Moreover, programme accreditation systems undoubtedly provide trustworthy guarantees as to the quality of the content that operators offer. These few points, combined with a new, slightly better calibrated and slightly more standardized survey method, could easily be enriched with additional insights and RAMAA could therefore become a relevant and useful tool to steer public policy for adult literacy.



LESSONS LEARNED

Over and above the ones discussed in previous sections, the lessons learned from RAMAA span three levels.

The first level touches on action research. RAMAA's co-construction approach involves creating tools. This means developing tools together with the national experts instead of using pre-existing ones. This RAMAA approach is challenging when it comes to supporting national teams because these teams in most cases need to depart from the assessment frameworks they generally use, which are typically based on programme content. National team members need to be aware that RAMAA assesses outcomes beyond literacy programmes, not literacy programme content per se.

As action research leads to change and, as such, generates scientific knowledge, it has a duty to help national teams to build and produce their own knowledge. It also has a duty to support teams through a learning process enabling them to take action over time and thereby embed the project in their national systems for the long term.

It is not easy to build collective learning momentum with participants who have such wide-ranging profiles and varied professional backgrounds. But, now that the first phase is complete, we can see that RAMAA has tackled this challenge. The cohesion, commitment and determination to take ownership of the project is palpable among the national teams. A member of the team in Burkina Faso attests to this: 'mobilizing a multi-disciplinary team on a national level is not simple but, at the same time, it's a collective, participative learning experience, and, with support from the international coordination team, we succeeded!'

RAMAA 2 should leverage this experience, which enabled more than 40 national experts to learn by following a complex path that involved formulating concepts, producing a variety of tools, scoping samples, and analysing and synthesizing findings. They were not merely welcome to have a look at what RAMAA was about from the outside: they experienced it from the inside, with all its difficulties and all its potential.

That said, during RAMAA 2 it will be important to work more closely alongside the existing national network of teachers and the national network of assessors, and to involve doctoral students, in order to include a broad spectrum of contributors with the profiles that suit RAMAA, and to create a link between fundamental research and practice. This phase will be crucial for the project's success from a qualitative perspective and needs to be planned carefully.

The second level of lessons learned is technical in nature. If we start from RAMAA's initial objective, i.e. to build tools that are standardized – standardized, that is, in the sense that they provide a shared perspective – it is important to begin by circumscribing research questions based on consensus, and by paying special attention to defining the various concepts that will be used to develop measurement instruments and drive result production. Work on the skills description, or in other words, answering the question what do we want to measure? will require in-depth discussion and, perhaps even more importantly, calls for the devising of a tool that is much less complex than the one developed during RAMAA 1. The quality of the tools, and therefore the quality of the results, will depend on RAMAA 2 fulfilling its original aim, i.e. adopting quality standards based on established scientific principles, and providing guidelines that everyone involved will understand and be able to apply consistently, at least to a significant extent.

Overall, RAMAA 2 will gain from drawing on practices in other surveys that explore the same issues, such as LAMP, PIAAC (the OECD's Programme for the International Assessment of Adult Competencies), etc. when it defines concepts, builds tools and so forth.

The third level, quality assurance – encompassing measurement tools and results – requires efficient and effective governance. This was not the case during RAMAA 1 because the governance mechanisms had not properly gelled by the time the project started. Selecting international experts at each research stage, moreover, appears to have undermined efficiency. Involving international experts was difficult, and that impacted project timelines. These experts also worked on their respective assignments in isolation because almost no channels for interaction were provided. We will therefore need to think again about this aspect of technical support.

The financial resources that were used to provide adequate technical support, and indeed to conduct activities in a few of the countries (for example in Mali and Niger, both of which experienced difficulties) had an unprecedented impact on the project. Financial resources are an important component in achieving project objectives, and the project must moreover rely on a solid case and robust communication.



RECOMMENDATIONS FOR RAMAA 2⁹

To deliver high-quality results, it is important to root RAMAA 2 in established scientific principles and to define the corresponding steering mechanism very clearly.

The conceptual aspects:

- ◆ Start with an inventory of all the literacy programmes available in the country and, ideally, feed them into a database. The advantage of doing this is that it will provide a full picture of all the options available countrywide. The databases will need to be built using a matrix, which could include the following: participant information (gathered before and on successful completion of a course or programme), a map showing operators and their programmes, types of programmes, etc.
- ◆ Extract information from the database to develop a common assessment grid geared to answering the question about what we want to measure. Apply this question to each type of knowledge, each type of skill and/or each type of impact we want to measure. The RAMAA 1 reference framework could be reviewed in this light.
- ◆ Conduct an inventory and critical review of best practices in national, regional and international surveys that could enhance RAMAA 2's conceptual framework, and bolster its 'capacity reinforcement' dimension.

The methodological aspects:

- ◆ RAMAA 2, like RAMAA 1, could be a collaborative research endeavour.
- ◆ Tool standardization/harmonization is of the essence.
- ◆ International benchmarking could follow. This opens up interesting prospects: it could supply very important information for participating countries, involve other players, and elicit emulation as regards quality and outcomes.
- ◆ Synchronization, governed by the project research plan, is also essential.
- ◆ RAMAA must not be considered a one-time initiative: it is an integrated process for continuous assessment. Assessments will therefore need to be conducted on a regular basis. And, as assessments are only useful if they are effectively used as

9 These recommendations might be revised after the 6th RAMAA Workshop, if necessary

tools to fine-tune steering, we need to think about assessments as systems to spur continuous improvement in education systems. In other words, it is absolutely essential to nurture an assessment culture.

The institutional aspects:

It is necessary to create a steering mechanism, before RAMAA 2 begins, encompassing but not necessarily restricted to:

- ◆ A steering committee to agree on RAMAA's overall guidelines and a scientific committee to ensure RAMAA delivers robust results.
- ◆ A pool of experts to conduct in-country operations.
- ◆ RAMAA national teams in participating countries, headed by a project leader and a scientific manager.

MEMBERS OF THE NATIONAL TEAMS DURING THE FIRST PHASE OF RAMAA**Burkina Faso**

Germaine OUÉDRAOGO, Economist APENF, National Coordinator of RAMAA
Catherine KABORÉ, Linguist, Technical Adviser, MENA

Paul DIABOUGA, Planning Officer, Senior Planning Officer for Evaluation, General Secretary, MENA

Claude DALBERA, Monitoring and Evaluation Officer, Consultant

Emmanuel GOABGA, Monitoring and Evaluation Officer, MENA

Edouard GUISSOU, Statistician and Demographer, MENA

Idrissa KABORÉ, Geographer, MENA

Rosine KABORÉ, Planning Officer, MENA

Maria KÉRÉ, Sociologist, FDC-BF

Mali

Moussa DIABY, Linguist, DNEF-LN, National Coordinator of RAMAA

Idrissa DIABATÉ, Statistician, INSTAT

Gouro DIALL, Linguist, DNEF-LN

Modibo DIALLO, Planning Officer, CPS

Abou DIARRA, Linguist, Technical Adviser of the Cabinet, MEN

Karamogo DIARRA, Computer Engineer, INSTAT

Diassé KONARÉ, Monitoring and Evaluation Officer, DNEF-LN

Ahmadou Mamou KONÉ, Monitoring and Evaluation Officer, DNEF-LN

Ouandé SOUMARÉ, Literacy Expert

Morocco

El Habib NADIR, former Director of DLCA, National Project Coordinator

Bensalem AIT BOULJAOU, Head of CAENF at the Regional Academy of Education and Training in Meknès Tafilalet

Taoufik BAH, Consultant, Statistician

Mohammed BOUGUIDOU, National Coordinator of the LAMP Project

Aziz CHAKER, Consultant in Sociology

Abdellatif FARIBI, Consultant and Adult Education Specialist

Fatima YACINE, Head of the Curriculum Department at DLCA

Niger

Laouali MALAM MOUSSA, Linguist, National Coordinator of RAMAA
Abdelkader GALY-ADAM, Teacher and Researcher, Director of IFAENF

Sabiou ARZIKA, Planning Officer, MENALN
Mahamadou CHAIBOU, Literacy Specialist, MENALN
Issoufou DANDI, Statistician, MENALN
Sibi IBRAHIM, Teacher and Researcher at IFAENF
Aboubacar MODOU AISSAMI, MENALN
Ali OUSMANE, Statistician, MENALN

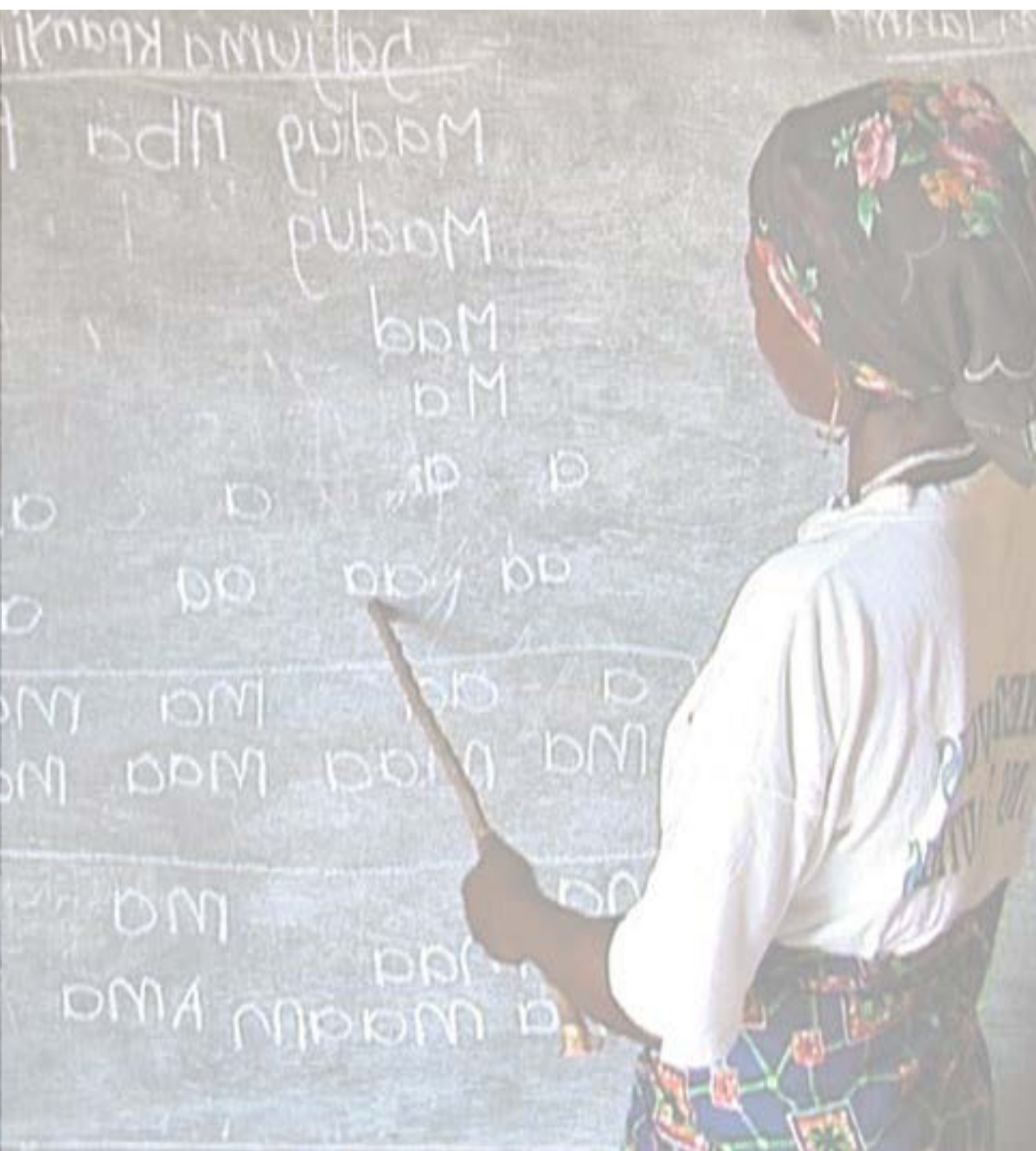
Senegal

Binta Rassouloula AW SALL, Programme Manager and Non-Formal Education Specialist, National Coordinator of RAMAA

Khalil DIARRA, Statistician, National Evaluation System
Mamadou MARA, Non-Formal Education Specialist, DALN
Boubacar NDIAYE, Statistician and Evaluation Specialist, DALN
Samba Diary NDIAYE, Evaluating Officer, CRFPE
Alioune Ngoné SECK, Linguist, University of Cheikh Anta Diop

- Beupere, N. and Podevin, G. 2010. *La "recherche-action": un processus d'émancipation et d'apprentissage pour et par l'évaluation*, JDL, 2010, Céreq, Relief No. 30.
- Berger, G. 2003. La recherche-action: épistémologie historique. In P. M. Mesnier and P. Missotte (dir.), *La recherche-action : une autre manière de chercher, se former, transformer* (pp. 13–26). Paris, L'Harmattan.
- Burkina Faso – Ministère de l'Éducation nationale et de l'alphabétisation. 2015. *RAMAA. Rapport d'analyse de l'enquête nationale* (preliminary version).
- Catroux, M. 2002. Introduction à la recherche-action: modalités d'une démarche théorique centrée sur la pratique. In *Recherche et pratiques pédagogiques en langues de spécialité*, Vol. XXI no. 3 | 2002, pp. 8–20.
- Coenen, H. 2001. Recherche-action: rapports entre chercheurs et acteurs. In G. A. Gilles and A. Lévy (dir.), *La recherche-action: perspectives internationales* (pp. 19–32). Paris, Eska.
- Courtois, L. 2013. *La recherche-action considérée comme un potentiel vecteur de changement des pratiques professionnelles du secteur associative: le cas de la recherché visant la promotion de comportements citoyens*. AREF international congress.
- Elliott, J. 2007. *Reflecting where the action is. The selected works of John Elliott*. New York, Routledge.
- Gohier, C. 2004. De la démarcation entre critères d'ordre scientifique et d'ordre éthique en recherche interprétative. In *Recherches qualitatives*, Vol. 24, pp. 3–17.
- Guillemette, S and Savoie-Zajc, L. 2012. La recherche-action et ses rapports de coconstruction de savoirs et de formation dans une perspective de professionnalisation entre acteurs praticiens et chercheurs. Formation et profession: revue scientifique internationale en éducation. In *Revue du CRIFPE* no. 20 (3), pp. 41–53.
- Kemmis, S. 1997. Action Research. In: J. P. Keeves (ed.). *Educational research, methodology and measurement: an international handbook*. Oxford, New York, Tokyo: Elsevier Science Ltd. pp. 173–179.
- Liu, M. 1992. Présentation de la recherche-action: définition, déroulement et résultats. In *Revue Internationale de Systémique* Vol. 6, no. 4, pp. 293–311.
- Mali – Centre National des Ressources de l'Éducation Non Formelle. 2015 *Enquête pilote RAMAA. Rapport final*.
- Morocco – Ministère de l'Éducation Nationale. 2013. *Enquête Nationale sur la mesure des apprentissages des bénéficiaires des programmes d'alphabétisation. Analyse et interprétation des résultats*.

- Niger. 2013. RAMAA. *Rapport d'exécution de l'enquête pilote au Niger*.
- OECD. 2013. *Technical report of the survey of adult skills (PIAAC)*.
- OECD. 2014. *PIAAC Standards and guidelines*.
- Périsset Bagnoud, D. 2007. Les référentiels de compétences, profils attendus de la professionnalité enseignante. Acteurs et points de vue en Suisse romande. In M. Bru and L. Talbot (dir.), *Des compétences pour enseigner: entre objets sociaux et objets de recherche* (pp. 87–104). Rennes, Presses universitaires de Rennes.
- Postiaux, N. et al. 2010. Référentiels de compétences à l'université, In *Recherche et formation*, 64 | 2010, pp. 15–30.
- Savoie-Zajc, L. 2001. La recherche-action en éducation : ses cadres épistémologiques, sa pertinence, ses limites. In M. Anadon and M. L'Hostie (dir.), *Nouvelles dynamiques de recherche en éducation* (pp. 15–49). Ste-Foy, QC: Presses de l'Université Laval.
- Senegal – Centre National des Ressources Éducationnelles. 2014. *RAMAA, Synthèse des résultats*.
- UIL-UNESCO. 2010. *Guidelines on sample design*. [Note sur le plan de sondage].
- UIL-UNESCO. 2010. *RAMAA presentation*. [Note de présentation de la RAMAA].
- UIL-UNESCO. 2011. *Canevas pour le référentiel des compétences* [Template for the skills description].
- UIL-UNESCO. 2012. *Framework document for the reference framework and background questionnaires*. [Quelques éléments de cadrage pour le cadre de référence et les questionnaires de contexte].
- UIL-UNESCO. 2012. *Proposal for a methodological framework based on analysis of international and national RAMAA outputs until 2012*. [Proposition d'un cadre de référence méthodologique sur la base d'une analyse des productions internationales et nationales de la RAMAA jusqu'en 2012].
- UIL-UNESCO. 2012. *RAMAA technical paper*. [Fiche technique du projet RAMAA].
- UNESCO. 2006. *Education for All Global Monitoring Report*.
- UNESCO. 2014. *Education for All Global Monitoring Report*.
- Wagner, D. A. 2015 *Des évaluations simples, rapides et abordables. Améliorer l'apprentissage dans les pays en développement*. Paris, UNESCO; IIEP and OIF.





ISBN 978-92-820-1202-4



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

Bureau à Dakar
Bureau régional
pour l'éducation
en Afrique



United Nations
Educational, Scientific and
Cultural Organization

Office in Abuja

Multisectoral Regional office for West Africa,
Benin, Cote d'Ivoire, Guinea, Ghana, Liberia,
Nigeria, Sierra Leone, Togo



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Agency for Development
and Cooperation SDC