

## ADULT LITERACY: MONITORING AND EVALUATION FOR PRACTICE AND POLICY

DANIEL A. WAGNER\*

**Abstract** – With the advent of the UN Literacy Decade launched in 2003, there is increased emphasis on monitoring and evaluation of adult literacy around the globe. The present paper presents an overview of the main approaches that have been taken to measure adult literacy within and across countries. A particular focus of the present review is to compare and contrast three models of literacy assessment: (1) the “traditional” model based largely on census enumeration, which has been used over recent decades to collect national and regional rates of illiteracy; (2) the large-scale survey techniques employed with the International Adult Literacy Survey and similar models; and (3) an intermediate type of assessment that borrows from the previous two models, and attempts to “tailor” the size and complexity of the survey methodology to the policy questions needing answers (called the SQC model). The present paper suggests that there is no single model or set of methods that are universally appropriate to monitoring and measuring in adult literacy around the world, but that blending both programmatic and comparative models through the SQC approach may bring greater involvement in, and insight into, adult literacy evaluations.

**Résumé** – ALPHABÉTISATION DES ADULTES: CONTRÔLE ET ÉVALUATION POUR LA PRATIQUE ET LA POLITIQUE – avec l'avènement de la Décennie des Nations Unies de l'Alphabétisation lancée en 2003, il y a une insistance accrue quant au contrôle et à l'évaluation de l'alphabétisation des adultes dans le monde entier. Le présent article présente une vue d'ensemble des approches principales qui ont été faites pour mesurer l'alphabétisation des adultes dans et à travers les pays. Un accent particulier du présent passage en revue est mis sur la comparaison et la mise en contraste de trois modèles d'évaluation de l'alphabétisation: (1) le modèle «traditionnel» basé en grande partie sur le dénombrement par recensement, qui a été employé au cours des dernières décennies pour recueillir des taux nationaux et régionaux d'analphabétisme; (2) les techniques d'enquête à grande échelle mises en œuvre avec l'Enquête internationale sur l'alphabétisation des adultes et avec des modèles similaires; et (3) un type intermédiaire d'évaluation qui emprunte aux deux modèles précédents, et essaye «de façonner» la taille et la complexité de la méthodologie d'enquête en fonction des questions politiques ayant besoin de réponses (appelé le modèle SQC: Smaller/Quicker/Cheaper: plus petit/plus rapide/moins cher). Le présent article suggère qu'il n'y a pas un seul modèle ou un ensemble de méthodes

---

\*The ideas contained in this article necessarily build on a number of projects in which the author has been involved over the past decade, some of which have been supported by UNESCO and other agencies through the International Literacy Institute (see ILI/UNESCO 1998, 1999, 2002a, b; Wagner/ILI 1998), and Wagner (2005). Other work in this area was supported by Unicef, World Bank and the UN Statistical Office; and more recently by the U.S. Department of Education and the Spencer Foundation. A number of the ideas, concepts and examples presented here have appeared in various ways in previous reports.

universellement appropriés au contrôle et à la mesure dans l'alphabétisation des adultes dans le monde entier, mais que mêler les modèles programmatiques à ceux comparatifs par l'approche SQC peut apporter une participation plus grande aux évaluations de l'alphabétisation des adultes et donner un aperçu de celles-ci.

**Zusammenfassung – ERWACHSENENALPHABETISIERUNG: DATENERHEBUNG UND AUSWERTUNG IN DER PRAKTISCHEN ANWENDUNG** – Seit Beginn der Alphabetisierungsdekade der UN 2003 erhalten Verfahren der Datenerhebung und Auswertung in der Erwachsenenalphabétisierung zunehmende Bedeutung. Der vorliegende Artikel bietet einen Überblick über die meistverwendeten Messverfahren der Erwachsenenalphabétisierung auf nationaler und länderübergreifender Ebene. Ein besonderer Schwerpunkt des Überblicks liegt im kontrastierenden Vergleich dreier Modelle zur Alphabetisierungsbestimmung: (1) das "traditionelle" Modell, das vorwiegend auf Zensus und Zählverfahren beruht und das in den letzten Jahrzehnten vielfach zur Erfassung nationaler und regionaler Analphabetenraten benutzt wurde; (2) flächendeckende Messtechniken, wie sie im Zuge des International Adult Literacy Survey und ähnlicher Modelle benutzt werden; und (3) eine Methode, die eine Art Zwischenform zwischen den beiden genannten Modellen darstellt und die sich darum bemüht, Umfang und Komplexität der Untersuchungsmethoden den jeweiligen programmatischen Fragestellungen möglichst "maßgeschneidert" anzupassen (das sogenannte SQC-Modell). Der vorliegende Artikel vertritt die Position, dass kein bestimmtes Einzelmodell oder Methodenset die universelle Eignung zur Messung und Kontrolle der Erwachsenenalphabétisierung in der Welt aufweist, sondern dass es die Verbindung programmatischer und komparativer Ansätze wie im SQC-Modell ist, die ein stärkeres Engagement und tiefere Einblicke in die Evaluation der Erwachsenenalphabétisierung erwarten lassen.

**Resumen – LA ALFABETIZACIÓN DE PERSONAS ADULTAS: MONITOREO Y EVALUACIÓN PARA LAS POLÍTICAS Y LA PRÁCTICA** – En los inicios de la Década de las Naciones Unidas para la Alfabetización, lanzada en 2003, se registra un creciente énfasis en el seguimiento y la evaluación de la alfabetización de personas adultas en todo el mundo. El presente trabajo pasa revista a los principales métodos adoptados para medir la alfabetización de las personas adultas en los diferentes países y en todo el planeta. En particular, esta reseña compara y contrasta tres modelos de evaluación de la alfabetización: 1. El modelo "tradicional," basado en su mayor parte en datos obtenidos en censos; es el que se ha usado a lo largo de las últimas décadas para comprobar tasas nacionales y regionales de analfabetismo; 2. Las técnicas de los estudios en gran escala, empleadas para la Encuesta Internacional de Alfabetización de las Personas Adultas, y modelos similares; y 3. Un tipo intermedio de evaluación que toma elementos de los dos modelos anteriores y trata de "ajustar" la envergadura y complejidad de la metodología del estudio a los planteamientos relacionados con políticas (llamado modelo SQC). En este trabajo, el autor opina que no existe un modelo único ni un conjunto de métodos universalmente apropiados para observar y medir la alfabetización de personas adultas en todo el mundo, pero que la fusión de ambos modelos programáticos y comparativos mediante el enfoque SQC puede generar una mayor compromiso con la evaluación de la alfabetización de las personas adultas y un conocimiento más profundo del tema.

**Резюме – ГРАМОТНОСТЬ ВЗРОСЛЫХ: МОНИТОРИНГ И ОЦЕНКА ДЛЯ ОСУЩЕСТВЛЕНИЯ ПРАКТИКИ И ПОЛИТИКИ** – С наступлением декады грамотности, провозглашенной ООН в 2003 году, все больший акцент ставится на мониторинге и оценке уровня грамотности взрослых во всем мире. В данной статье предлагается обзор основных подходов, которые используются

для определения уровня грамотности взрослых внутри стран и между странами. Особое внимание в данной статье уделяется сравнению и противопоставлению трех моделей оценки уровня грамотности: (1) "традиционная" модель, основанная в основном на переписи населения, которая использовалась на протяжении последних десятилетий для сбора данных относительно уровня грамотности на национальном и региональном уровнях; (2) методы широкомасштабного обзора, использованные в Международном обзоре грамотности взрослых и подобных моделях; и (3) промежуточный вид оценки, который вытекает из двух предшествующих моделей и пытается «приспособить» масштаб и сложность методики проведения такого обзора к вопросам политики, которые необходимо разрешить (так называемая модель SQC). В данной статье автор предполагает, что нет единой модели или набора методов, которые были бы универсально приемлемыми для мониторинга и измерения уровня грамотности взрослых во всем мире, но, тем не менее, автор считает, что объединение прагматических и компаративных моделей посредством подхода SQC может способствовать лучшему пониманию и большему участию в оценке уровня грамотности взрослых.

### **Varying approaches to monitoring and evaluation**

The World Conference on Education for All (WCEFA), held in Jomtien, Thailand, in 1990, found that the measurement of learning achievement was critical to judging the quality of education programs around the world (UNESCO 1990). This recommendation followed on decades of evaluation studies that often used "participation," "funds spent," or "contact hours" as proxy variables for determining the quality of a nation's efforts to improve education. With the arrival of WCEFA, it was agreed that what the learner actually learned should be at the center of the educational enterprise, and be useful for both improving practice and implementation, as well as for policy development.

In the nearly two decades that have followed WCEFA, a number of major initiatives began to reinforce not only capacity building in learning assessment, but also the reconceptualization of what assessment ought to be in various education domains. Most of this effort, including fiscal support, went into the formal schooling sector. However, soon after Jomtien, efforts began to address assessment issues in non-formal education (NFE) and adult literacy in developing countries (Wagner 1990; Ziegahn 1992), and major surveys were undertaken in the United States (Kirsch et al. 1993) and internationally (OECD/Statistics Canada 1995). Some of these formed the basis for a renewed call for increased work on adult literacy, which was part of the World Education Forum, held in Dakar in 2000.

Thus, at the time of the declaration of the UN Literacy Decade in 2003, a variety of approaches or models of monitoring and evaluation of rates

(or levels) of adult literacy have been put forward, and some have been implemented in various contexts. The present paper contains an overview of the main approaches that have been taking to monitor and evaluate adult literacy within and across countries. Since other reviews exist, the primary focus of the present paper is to review the differing purposes of different models of monitoring and evaluation (M&E) in light of their intended or presumed policy goals, and implications for practice. In short, as with all measurement tools in the social sciences, there can be quite different approaches to data collection and analysis of data, each with costs and benefits – where costs are not only fiscal, but also include human resources, time, political capital, and where benefits, in a parallel fashion, may be in seen in terms of national rankings or in improving instruction or in teacher training.

In sum, there is no magic bullet in monitoring and evaluating in adult literacy work, but, as will be argued here, there are serious choices to be made depending on ones goals and means for reaching them in practice.

### **International statistics on literacy**

In the social sciences, the gathering of statistics on anything – literacy or otherwise – can imply a rather varied set of approaches to data collection. Historically speaking, and in order to provide worldwide comparisons, the international development community has relied over decades almost entirely on data provided to UNESCO by its member countries (UNESCO 1983). These countries, in turn, typically rely on a *national population census model*, which most often determines literacy ability by self-assessment questionnaires and/or by means of a proxy variable utilizing the number of years of primary schooling (i.e., 5 or 6 or 8 years of primary schooling equals a “literate” person). Such data are then collected and collated by UNESCO to create adult literacy rates, typically the number of persons over 15 years of age in the population who are designated as “literate” divided by the total number of persons in that same age category (and then often broken down by age range, gender, urban-rural residency, and so forth). Considering these same statistics over decades has provided a statistically robust way *monitoring* literacy levels on a national, regional and international basis.

Note here the use of *monitoring*, which from the Latin word *monere*, means “to warn” or to observe on a situation. This implies gathering sufficient, but minimal levels of information needed in order to judge if there is a problem to be warned about. Monitoring, thus, implies a “lite” version of data collection, sometimes using not only proxy variables (such as school statistics which may be only related to, but not directly indicative, of individual literacy levels), but also, especially in the case of self- or other-assessments (where a village leader may “say” that so-and-so person “is literate”) leave much in doubt as to the actual status of literacy within individuals or groups of individuals.

Because of doubts about the reliability of such data collection, considerable concern has been expressed about the credibility of literacy statistics. In 1986, for example, UNESCO and the UN Statistics Office (UNSO) held a joint seminar in Paris to discuss the use of household surveys to improve the collection adult literacy statistics; a technical report which was the basis of this seminar was later published (UNSO 1989); these discussions were followed by similar ones over the next decade and a half (see, for example, ILI/UNESCO 1999, 2001, 2002a, b). The concern, as noted above from the 1990 WCEFA, was whether actual skill learning had taken place in schools and/or in adult literacy programs. The traditional method of monitoring international literacy levels was clearly very limited since few countries bothered to actually *measure* individual skills in a large or broad enough population sample to assure that literacy levels were valid and reliable.

Over nearly two decades, there has been a movement to add greater statistical strength to the monitoring of adult literacy, and this has meant a commensurate movement toward direct measurement or assessment of literacy skill. The two terms – monitoring and measuring – are therefore complementary: one can monitor with either good or poor measures, but it is better to have the best measures possible within available resources. This is much easier said than done; see for example, the recent efforts at the US National Academy of Sciences to address these same issues in America (NRC 2005). In fact, even as many specialists now agree that exclusive reliance on traditional indirect measures of literacy may be flawed, there is renewed discussion of the utility of proxy measures (Desjardins and Murray 2003; Murray 1997), since they are may be sufficient and cost less. Indeed, cost effectiveness is an issue that must be kept clearly in mind, as will be discussed below, since resources are always limited, and seeking to collect too much data may be as counterproductive as collecting too little.

## **Policy issues in literacy measurement**

### *Learning achievement and information stakeholders*

At WCEFA in 1990, a number of educational targets relating to youth and adults were agreed upon, including reducing by 50% the number of adult illiterates by the year 2000 and improving learning achievement to an agreed percentage of an appropriate age cohort. Unfortunately, as noted in the *Literacy and Adult Education* thematic paper for the 2000 Dakar EFA meeting (Wagner 2000), the Jomtien emphasis on learning achievement had not succeeded in becoming a major part of literacy work in the decade between 1990 and 2000. By 2000, it was agreed that the field of adult literacy requires both a greater focus on the quality of literacy services and better ways to measure literacy achievement. In the Dakar Framework for Action

(UNESCO 2000), these approaches were reinforced by three of the six main stated Dakar EFA goals, namely:

- (iii) ensuring that the learning needs of all young people and adults are met through equitable access to appropriate learning and life skills programs;
- (iv) achieving a 50% improvement in levels of adult literacy by 2015, especially for women, and equitable access to basic and continuing education for all adults;
- (vi) improving all aspects of the quality of education and ensuring excellence of all so that recognized and measurable learning outcomes are achieved by all, especially in literacy, numeracy, and essential life skills.

Moreover, in the declaration of the UN Literacy Decade (April 2003), and reinforced in the 2006 EFA Global Monitoring Report on literacy (UNESCO 2005), evaluation processes are expected to include at least three indicators of literacy progress, namely:

- The change in the absolute numbers and in the percentages of the literate population;
- The relative contribution of formal and non-formal education to attaining literacy;
- The impact of literacy on the quality of people's lives.

Effective utilization of such indicators requires, at a minimum, a way of improved measurement tools over the traditional methods described earlier. With improved assessment methodologies and greater availability of new technologies for collection and analysis of data, it is possible, more than ever before, to increase access to credible data on learning achievement and the status of basic skills.

As in all areas of education, and perhaps more salient than in some, there are a number of diverse and even contentious stakeholders in the literacy field. There are, as noted, international and national agencies that seek to better monitor trends in literacy, so as to make improved decisions about future investments. In addition, there are an even wider variety of local (or multinational) non-governmental organizations (NGOs), program directors, and practitioners who have a need not only for government political support, but also to improve their ability to function competently and to demonstrate to themselves and others their capabilities. And, last but not least, there are learners, and their organizations, that increasingly call for improvements in the quality of literacy programs, and who wish to know what the likely impacts of participation in such adult literacy efforts.

It seems reasonably clear that neither NGOs nor adult learners living in Botswana will care very much about whether Botswana's rate of literacy is superior in a given year to that of Peru. Similarly, it is not clear, for example, that UNESCO will have the resources, under traditional literacy data

gathering, to assist Botswana in reaching marginalized ethnic groups living on its borders. The point here is simply that not all monitoring methods will likely please all stakeholders equally or at all. Conversely, there are stakeholder needs that can and should be taken into account that may go well beyond the typical national monitoring schemes currently in place.

Therefore, the problem is not – as some have occasionally said – to be only one of sufficient technical expertise or human resources in the field. Rather, there are important choices to be made about which information stakeholders will be attended to in new M&E efforts in literacy. To be concrete, let us consider the following types of stakeholder questions:

- At the national level. How can we better judge the current status of literacy levels among out-of-school children, youth and adults, irrespective of former school attendance?
- At the program level. How can we identify skill gaps and needs that may serve as better targets for interventions across diverse ethnic and linguistic groups;
- At the learner level. What am I going to get out of participation in XYZ program, especially in light of the work needs that I have in my own household or community?

Naturally, policy goals will vary across countries, as well as by gender, ethnic group and region (and more) within countries. Countries concerned about overcoming inequalities caused by geographical disparities may want to collect more information about type of housing or community, or distance from the nearest school, in order to be able to identify those living in remote or in inaccessible areas. Countries concerned about inequalities by gender, ethnicity, or language groups will need relevant variables that capture group membership in this regard. Finally, countries interested in relating the assessment results to specific programs or educational experiences will have to include relevant questions concerning program curricula.

Improved literacy measurement and data collection can provide better answers to a variety of stakeholders, but not every method or model can all needs of all literacy stakeholders. Priorities will need to be discussed and set, decisions taken, and some options foregone in order to achieve the possible.

### *International comparability of data*

The comparability of data is a major concern for policymakers and planning agencies. If definitions and classifications vary, then it can be difficult if not impossible to compare data collected through different surveys. Comparability and stability are necessarily the hallmarks of the UN data collection, including EFA Monitoring Reports and the work of the UNESCO Institute on Statistics (UIS). Nonetheless, if comparability becomes the primary goal, while less attention is paid to the (local and cultural) validity of the definitions and classifications of literacy, then the data collected may become less



meaningful and potentially less applicable at the ground level. This is a natural and essential tension between “emic” and “etic” approaches to literacy measurement (Wagner 2004).

International and national needs, definitions, and research strategies may or may not come into conflict over the issue of comparability, depending on the particular problem addressed. For example, as mentioned above, UNESCO solicits literacy data worldwide, where literacy has been measured in terms of the number of “literate” and “illiterate.” For most countries, this dichotomous type of classification presents few practical or technical problems, and is relatively inexpensive to gather (as part of nationwide censuses), while providing international agencies with a cross-national and time-series framework for analyzing literacy by geographic or economic world regions.

However, educational planners may want to know about the effects of the completion of primary or secondary schooling – such as, how much was learned in a particular area of study, or in a particular literacy campaign the levels of literacy attained. In these cases, a simple dichotomy is too blunt as a statistical instrument; skill scores or levels are clearly required for learning achievement to be adequately measured. Furthermore, precise data are needed as to which languages and scripts are used in each region and by ethnic group, in addition variation by age and gender. The collection of such data has largely been ignored by most national and international agencies to date (with some notable exceptions). See a comparable discussion of this issue in the area of M&E in educational technology (Wagner 2006, Chap. 7).

The impact on educational policy of such comparative studies is due at least in part to the realization among national policy makers that their country (or population segments therein) may be far lower in certain areas of learning achievement than would have been predicted, say, by participation in school. In a World Bank national household survey in Bangladesh, it was found that five years of primary schooling resulted in only a first grade equivalent of learning achievement, and that three years of schooling had approximately zero value in terms of learning achievement (Greaney et al. 1998). This study may have an important impact on the kinds of investments that Bangladesh makes in the area of basic and non-formal education in the future.

### *Moving from monitoring to direct assessment*

#### *The end of a dichotomy*

There is little doubt that the use of dichotomous variables in literacy work have had a deleterious effect on the field in a number of ways: from the initial equating of illiteracy with uncivilized, to the continued use of “literate” versus “illiterate” in census data in many countries today (Wagner 1990, 2001). Indeed, one can possibly argue that we are all illiterate in one way or another, with the merger of such terms as reading, health literacy and technological literacy; or that we are all literate, in the sense that very



few people in today's world have little or no knowledge of the purposes and nature of literacy in at least one script. In other words, there are a great deal more shades of gray than black and white in contemporary discussions of literacy and its measurement. This conclusion has the positive consequence of demonstrating that, like formal schooling, adult literacy programming deserves a great deal more investment than it is currently getting (for a discussion on sub-Saharan Africa, see Lauglo 2001). With very little doubt, the next decade will see end of present use the literacy-illiteracy dichotomy in discussion of literacy M&E and policy.

Similarly, there is little doubt now that there must be greater resources invested in the direct measurement of literacy. Yet, what type of investments should be made? We begin with a brief review of one of the best-known attempts to engage in direct measurement.

#### *Large-scale literacy assessments: IALS and successor methods*

The International Adult Literacy Survey (IALS) has become well-known over the past decade for its emphasis on direct measurement, and its approach to providing comparative data across countries (mainly in industrialized countries; see OECD/Statistics Canada, 1995, 1997, 2000). The IALS methodology is based on a number of national predecessors, such as the 1993 U.S. National Adult Literacy Survey, or NALS; Kirsch et al. 1993), which invested significant resources in improving the technical and psychometric properties of literacy assessment instruments, using a variety of techniques, including methods for expanding the range of items used in a survey format, including Item Response Theory (IRT). The IALS, and its predecessors, utilized a five-level categorization method for literacy, along three different scales (prose literacy, document literacy, and quantitative literacy (or numeracy). These survey scales are not without critics (e.g., Reder 2000, on the NALS, on the colinearity of the three scales); and there are criticisms, on, for example, the degree of international comparability (e.g., Kalton et al. 1998, on population sampling difference across IALS countries) or on item comparability (Levine 1998).

Since the launch of the UN Literacy Decade, the UIS has launched an initiative called LAMP (Literacy Assessment and Monitoring Program), where they are planning to build on some of the tools developed through the IALS, but refocused on literacy assessment in developing countries (Terryn 2003). While the effort is still in its early stages, LAMP holds the promise of being able to employ some of the best technical tools in literacy assessment, and adapt them for use in poor countries. Cautionary remarks about the cost-benefit trade-offs in poor countries have been made by a number of literacy specialists (e.g., ILI/UNESCO 1999; Street 1996; Wagner 2003). Among the issues invoked in such remarks are the low degree of transparency of the data when collected using the IRT methodology, the expensive process of creating methods of cross-cultural comparability, the long time taken to complete a given study, and the overall cost of highly sophisticated methods in poor countries.

*Challenges to the IALS model*

There are numerous challenges to any method of testing. These range from disputes about theoretical and operational definitions to the type of statistical tests employed to how to analyze datasets. Given space limitations, it is useful to simply note some areas that have been particularly debated in studies involving the IALS assessment methodology. These include: (a) scales of literacy achievement (from dichotomous, to 5 levels, to many levels); (b) determination of when a “level” is achieved (e.g., in IALS, is it adequate to say that a level is achieved if and only if 80% of the items in a level are completed successfully; see Levine 1998); (c) what is included in the operational definition of literacy; (d) effectiveness of the use of proxy measures (Lavy et al. 1995; Murray 1998); and (e) determination of which populations are or are not included in the population survey (e.g., are ethnic minorities, and their languages/literacies excluded?).

The issue of population sampling also poses a set of changes in the IALS model. For example, resources might be invested in a more selective fashion (directing more funds to preschool and primary schools, or to specific groups of adults), so that some individuals – those with a greater chance of success – would have access to the possibility to become literate or more literate. Indeed, recent evidence on volunteer literacy efforts in the United States suggests that the least literate portion of the population is remarkably resistant to literacy training, often exhibiting much higher rates of program attrition and lower learning achievement levels (Venezky 1992). Research in developing countries in this domain would be very illuminating, and might result in some new policy decision-making.

International surveys like the IALS have also been criticized for being too expensive and too complicated to be “owned” (that is accepted for endogenous and locally sustainable use) by national and local agencies. While comparative surveys have often received considerable media and policy attention, and have led at times to significant national educational policy impacts, the cost is high relative to local budgets and/or opportunity costs in developing countries. National or local household surveys can also have a similar policy impact, but this result necessitates a serious and expert study, followed by concrete efforts to publicize results, something often difficult to achieve without the “credibility” of external agencies and high-priced consultants.

The costs of large-scale assessment studies are quite variable. Estimates of the total cost of the IALS survey(s) run as high as tens of millions of US dollars, while the costs of the UNSO (1989) national literacy survey in Zimbabwe may be estimated at about US\$100 thousand in current dollars. Costs clearly depend on the number of countries included, the degree of external expertise required, the complexity of study design, and degree of collection and analyses undertaken. Clear trade-offs are available in costing processes, from limiting sample sizes to the length of tests created to the degree of trained personnel required. Nonetheless, there is little exactitude in

current cost estimation due to the paucity of available M&E literacy studies with cost figures.

### *Household surveys and program evaluation models*

Household surveys have been used for decades, often employing randomized samples to gather specific types of information on “target groups” within countries or regions in countries, and stratified along certain desired demographic parameters. In the literacy field, one of the first household surveys was undertaken in Zimbabwe, and referenced earlier (UNSO 1989), with numerous others to follow (e.g., in Morocco, Lavy et al. 1996; in Bangladesh, Greaney et al. 1999; in Botswana, Commeyras and Chilisa 2001). Further, in a multiyear effort just preceding the 2000 Dakar EFA Forum, the International Literacy Institute (ILI) and UNESCO undertook a series of workshops and field studies in the Literacy Assessment Project (LAP). LAP took a *pro-local* approach to surveys, trying to situate data collection more toward meeting local and national needs, and to developing practical tools that could be understood by laypersons and mid-level adult education specialists. This position was seen to contrast with the IALS model (and later the LAMP model) that was designed primarily to deliver *comparative* data to high-level national and international agencies, and well-trained statisticians. The essential tension between comparability and statistical reliability on the one hand, and local relevance, transparency and face validity on the other, is an endemic problem in international social science. For a discussion in the adult literacy field, see a set of LAP documents on this topic (ILI/UNESCO 1999, 2001, 2002a, b; ILI 2000).

A more common approach at the ground level is that of program evaluation methods, which have a long history in adult literacy and educational development work. Unfortunately, with a few exceptions (such as the interesting Okech et al. work in Uganda, 2001), most program evaluation work does not use sufficiently robust methods that would allow for serious skill assessment. This is not surprising in that evaluations are largely sponsored within NGOs that have little in the way of internal technical capacity.

## **Choice of design for M&E data collection**

### *Choices in survey design*

There are, of course, many ways to collect data on individuals and groups of individuals. Choosing among them is a major challenge that is (or should be) directly related to the policy or program questions that one seeks to answer. A brief summary of these models is provided as follows.

- *National and international sample surveys.* The IALS model, as described in some detail above, is most often undertaken as part of the work of a

national census bureau, with a focus on proper sampling across demographic parameters, and the inclusion of psychometric tests and analytic techniques. Efforts to make such surveys comparable at an international level are complex and relatively expensive.

- *Household surveys.* These allow considerable flexibility in terms of sample selection and sample reduction (one can limit the sample size through appropriate stratification), and thereby provide more time for actual learning assessment. One limitation is that such surveys are often “stand-alone,” and require specialized training as well as relatively skilled test designers. As part of this same approach, a *special literacy survey* may be undertaken with a particular focus on literacy assessment, but is not a part of a census sample (household or otherwise). One example of this approach was that undertaken by the World Bank (Greaney et al. 1999) in rural Bangladesh, or as part of research projects (Wagner 1993). A recent derivative of the household survey has been termed the SQC approach that will be discussed in the next subsection.
- *Post-censal sample.* The LAP India case study cited the NSSO study that was of the post-censal (after a census) type. One advantage is that the same personnel and data structure can be used as in the census, such as in the India case study in the LAP project (ILI/UNESCO 2002b), there are advantages in being able to compare the data collected from the regular census with the post-censal in-depth study. One limitation is that census personnel often seem to have less technical capability in designing basic skills assessment instruments.
- *Program evaluation.* Evaluations of literacy programs in developing countries are numerous, of both formative and summative varieties (Bhola 1999; Carron et al. 1989; Okech et al. 2001; Nordtveit 2004). In developing countries especially, such evaluations have only rarely included psychometrically appropriate tests for measuring learning achievement. Typically, measurement focuses on the inputs in human and infrastructural resources, the pedagogical methods employed, and the outcomes in terms of attendance and “successful program completion.” A lack of skill assessment is typically a very serious gap in the current program evaluation knowledge base.
- *Other issues.* Among other issues is how to meet high-stakes versus low-stakes concerns on the part of participants in the assessments. In-school tests are often seen as high-stakes, but this is not the purpose of the current literacy assessment exercise – yet participants often view testing as a simple extension of schooling. Thus, care needs to be taken, or, as in some of the case study examples, one may find data falsification on the part of either learners or instructors or both. Finally, some procedures for post-test adjustments for sampling biases may be undertaken, though these may or may not be necessary, depending on the sample and testing methodology chosen.

*A tailored approach to survey designs: The SQC model*

It is clear that international and local needs may not be one and the same, with “national” needs falling somewhere in-between. Countries and cultures are diverse, each with a multiplicity of groups that vary along ethnic, linguistic, social class, economic and other dimensions. Each country has its own special history of sociopolitical development, and its own experiences with formal schooling and broader educational development. The international policy community has its interests as well, mostly in trying to guide national decision-making from indices of where nations “fall” on some scale of economic productivity or worker efficiency – hence the “horse-race” concept in international comparisons.

The improvement of literacy assessment in comparative context may affect local, national and international interests in contrasting ways. National interests and “internal” considerations (involving, for example population/ethnic diversity) may be seen as nettlesome problems, or simply constraints by planners concerned with international comparison. On the other hand, national considerations about population diversity, linguistic variations, and even orthographic diversity (such as unusual features of a script) may be seen as having to be sacrificed on the methodological altar in order to achieve a larger basis for international comparison. For these and other reasons, there is ample logic for local programs and national level policy makers to hesitate in sacrificing local, cultural and/or national interests for those with an interest in regional or international comparisons.

More specifically, the level of resource investment in empirical data gathering in IEA-like studies is, for many developing countries, far greater than that made heretofore. Thus, there may be opportunities to create a research infrastructure through the carrying out of international comparisons, not so different in kind from that achieved by anthropologists working with diverse groups in developing countries. Perhaps most importantly, comparative studies can, if properly designed, help to achieve a greater understanding of cultural values and attitudes towards learning and literacy. Such an approach would enable the problem of diversity to be turned into opportunity of studying varieties of learning in context (Wagner 2004).

The above analysis has led to the importance of seeking alternatives to the technically complex and expensive IALS model on the one hand, and simple non-psychometric program evaluation on the other. What might be called a compromise model is, in reality, an approach that is tailored to the nature and size of the policy and program questions at hand – what has been termed the Smaller/Quicker/Cheaper (or SQC) model (Wagner 2003). We argue here that the SQC model has much to offer in present-day work in M&E in adult literacy.

(a) *Smaller*. Assessment methods do not need to be major entrepreneurial enterprises, but rather just robust enough to answer key policy questions at the national and local levels. Indeed the focus on “size” needs to be tailored,

as with all social science methodology, to the precise set of questions to be answered. The term “smaller” generally has two main meanings: first, the number of countries included in such studies may be only one, and in the case of large countries (e.g. India), may be at the state or sub-state level. Second, whatever the population group studied, the population of human assessed, as well as the number of items utilized in assessment instruments need only be “just large enough” to answer the relevant questions. Of course, in some cases, this may be easier said than done – it is not always possible to gauge such matters with great precision.

(b) *Quicker*. Literacy assessments need to be completed in “real time,” or thereabouts, so that results can affect policy and spending in the “lifetime” of current ministerial appointments and/or programmatic needs. Studies that take 3–5 years to generate results, even if analytically robust, nonetheless fail to meet the test of timeliness. The IALS and International Educational Achievement (IEA) studies, just to name two examples, usually take years to carry out, analyze and publish. The time-span needed to make local decisions that affect budgetary policy and programmatic change is much shorter, usually at most 18 months. Reconciling these conflicting temporal needs would not be so difficult if the studies were smaller in the first instance. Large-scale studies necessitate more time; but are large-scale studies necessitated themselves?

(c) *Cheaper*. Funding is a driving force in all development work, and is one of the reasons why large-scale assessments have received the large majority of funding for literacy assessment. It is much easier to be on the “radar screen” of the World Bank if one has a large-scale study, with the potential for a large policy report or reports. It is also an easier way to handle development agency budgets. But seen from a national or local perspective, things may be quite different. Developing countries may not be able to afford either the fiscal or human resources costs of deep involvement in highly technical assessment exercises over multiple years. Further, the higher the cost, the more difficult it is to get to an initial “yes” to participate in such a national exercise, and the more difficult to gather time series data to follow policy decisions.

“Cheaper” also means the possibility of deploying local (rather than external) experts and consultants. This can be achieved if the assessments are not constrained to use outside agencies in the industrialized countries to process complex data sets. By choosing simpler psychometric methods, one can make data and their analyses more “transparent” to local human resources. Assessments necessarily comprise a variety of expertise. These include the policy makers, psychometricians (test makers), survey and interview specialists (enumerators), and data analysts, as well as learners and instructors for preparatory pilot testing. Capacity building – the training of sufficient expert staff – has been a major impediment in past large-scale surveys. How such resources can be put into place in the future will be a major

question to address, but it will be much easier in local and national contexts than on an international level.

### *Some cost considerations in the SQC model*

In addition to the general cost-reduction features of SQC, it must be said that each design consideration in a literacy assessment has its specific associated costs. Clearly, for use in developing countries, the intent here is that such literacy assessment can be low-cost relative to other, “higher cost” approaches – hence the term “cost-effective” that has been used throughout this document. However, it needs to be said at the outset that there are no absolute certainties with respect to cost, though there are some trade-offs that can be made in terms of cost and quality. The following are specific cost considerations with respect to the SQC model:

- Limited sample household surveys can save money because they simply reduce the number of individuals that need to be interviewed in order to answer a set of particular policy questions. As noted earlier, larger scale studies – and especially those that include international comparisons – can drive costs upward.
- Lower levels of statistical reliability than sometimes used in censuses may be acceptable in literacy assessment in order to speed up data collection and reporting, hence reducing overall cost.
- Transparency and simplicity of the proposed survey design means that fewer experts are required to “interpret” (or reanalyze) the data for a non-expert audience. This translates into fewer relatively expensive external (international) experts, and more reliance on local and national staff.

### *Some limitations of the SQC model*

How can both comparability and context sensitivity be appropriately balanced in literacy M&E and basic skills assessments? Indeed, how can EFA monitoring, or UN statistics collection, be maintained as stable and reliable when localized approaches are chosen over international comparability? The answer would seem to lay, as it should, in compromise. At present, much if not most, “comparative” data from developing countries is inherently flawed by failure to use direct assessments, as we reviewed earlier. The SQC model would rectify this situation by encouraging all countries to choose some version of an SQC survey to enhance the credibility (validity and reliability) of data collected at the national level. Further, these data could be compared to similar data from other countries. However, what is *not* being required in the SQC model (but which is required in the IALS and LAMP models) is the strict (or nearly strict) item-by-item comparability. Under SQC, the advantages of size, speed and cost – and what follows from such agility – are may or may not outweigh the importance of exactitude in cross-national



comparison. Determining priorities is of course at the center of policy making in M&E assessments.

Nevertheless, we should not confuse the SQC acronym with the word “simple.” In fact, there is no doubt that SQC methodology requires some of the same skilled capabilities as does larger and more complex studies. These are more matters of degree of intensity and time of inputs. As with the IALS, SQC implementation of a household survey of out-of-school requires careful planning and human capacity building. Any survey must be planned and implemented in the field in such a way that the information collected is credible and free of error, while costs, logistical efforts, and time needed are kept at reasonable levels. At the same time, the design of data collection and reports from the survey must be planned ahead of time to focus on variables and questions of maximum interest to the stakeholders and users of the results. Planning processes need to be thought of as integral to the entire literacy survey. This is particularly the case if the data to be collected will serve both the information needs of national policy makers as well as decision-makers as well as directors of regional or local programs whose projects may serve out-of-school populations.

*A recent example of the SQC model: monitoring the literacy campaign in Morocco*

In Morocco there have been a few attempts to use “local language” or mother-tongue literacy, namely the use of *Amazigh* (Berber, in three dialects) and Moroccan dialectal Arabic (*Darija*), in non-formal and adult literacy education programs. However, none of these have been developed with scientifically designed methods that take advantage of the semantic and syntactic interrelationships between these local languages and *Fusha* (standard Arabic). Further, due to local and regional political sensitivities, government support for such local language approaches has been minimal until a new program (ALEF) began as a joint project of the Moroccan Government and a bilateral donor agency.<sup>1</sup> For a more substantive historical review of the Moroccan literacy situation, see Wagner (1993); for the current ALEF work, see Wagner (2008).

Monitoring and evaluation (M&E) are a key part of the effort to understand the effectiveness of innovations in literacy programming and impact. The M&E work was based on the SQC concept that involved scientifically valid instrumentation, as well as near-term and transparent assessment processes. In the evaluation study, a sample of female adult learners was selected from across ALEF participating literacy programs in Morocco. Each learner was interviewed and tested by a trained staff person fluent in the local language; testing took about 30 min per person. The implemented plan was designed specifically to measure the impact of the ALEF method on *Fusha* (standard) Arabic literacy learning, with part of the sample studying in *Darija*, and the other part in *Amazigh*. In addition, a control group, drawn

from the same regions and NGOs, was tested as a comparison; these learners used only the regular government curricula (in *Fusha* only).

The testing instrument – the Literacy Learning Assessment Test (LLAT) – followed the SQC design in that they would be smaller in number of items, quicker to administer and analyze, and cheaper overall to deliver (requiring few highly trained experts). The LLAT was designed to measure the learner’s acquisition of Arabic from initiation through elementary functional literacy. It was specifically developed to assess the ALEF program (of 60 h total), and the transition to and through the Moroccan “official” program of study.<sup>2</sup>

There were three main sampling parameters of the M&E study: region (16 provinces), language (Amazigh-speakers or Darija-speakers) and program of study (ALEF or official). Additional within-group variables that were measured include: age, prior schooling or literacy program experience, and other demographic characteristics. Learners were women selected randomly from each of 16 Moroccan provinces from within ALEF-sponsored participating programs in 2006.

Preliminary results of the ALEF intervention program indicated a rapid growth in skills during and following the ALEF curriculum program. Further, the ALEF program was as effective over time for the Amazigh-speaking groups when compared with the Darija-speaking groups. In addition, the control group learners (those who did not take the ALEF program) performed considerably less well than the ALEF sample. Thus, it seems that the ALEF approach was a considerable improvement over the current ‘official’ literacy program in terms of learning gains. Final results will have to await a complete rendering of the all data, as well as other predictors of skill outcomes (such as prior skill levels, years of schooling, etc.).

In sum, the ALEF literacy assessment study was created with a small team of (mainly) local experts, with instruments pilot-tested in a matter of weeks and months rather than years, and on a budget of only a fraction of what large comparative studies spend. Further, policy analyses, such as those that derive from the above results, became available nearly immediately after data collection, leading to the possibility of very timely decision-making by relevant authorities in order to improve *both* the ALEF and official government programs.

## Conclusions

Policy makers typically seek to undertake literacy M&E by utilizing the best data resources possible in order to make decisions. The problem in the adult literacy arena, as in others, is how to define “best.” Breadth (e.g., broadest demographic sampling) and depth (e.g. most reading items tested) of data collection are two ways of thinking about what is best, but so are transparency, “shareability” of data across programs and contexts, and human resources capabilities.

The movement toward improved monitoring and evaluation of literacy may be summarized through a small set of models described above:

- (a) Traditional model. Traditional census or census-like data collection as over the past several decades, usually taking the form of self-assessment or proxy variables (years of schooling).
- (b) IALS/LAMP model. This model, developed originally for use in industrialized countries utilizes advanced technical and statistical tools in order to achieve statistical comparability at both international and intra-national levels.
- (c) SQC model. The advantages of SQC are inherently the emphasis smaller, quicker and cheaper ways of gathering and analyzing data, with a focus on local and national level instruments.

If the goal is to improve literacy in the poorest countries – a major goal of the UN Literacy decade – the SQC approach may allow greater effectiveness in reaching the unreached and un- (-under) schooled (gender, minorities). Further, the SQC model can foster greater impact of data on policy (robust) than is currently available; have a more timely impact of data on policy (practical); and provide for a more sustainable future of the knowledge base. Much needs to be done to obtain a full vetting of alternative approaches that do not seek to dismantle what has been learned already (and will continue to be learned) from international comparative studies. Yet, a unique emphasis on either the traditional or the IALS approach will miss some important opportunities. The above description of the ALEF project in Morocco provides a useful example of what can be done with both methodological rigor, and empirically sound data collection and assessment.

Policy makers will inevitably undertake further surveys in order to answer the following kinds of questions: (a) what does a given population “know” in terms of literacy skills; (b) how these literacy skills are used and (c) what policy changes can be made that might influence future levels of basic learning competencies or ways to improve the acquisition of literacy skills. Such efforts will need not only the collection of skills-related (psychometric) data through the testing instruments, but also careful collection of background information that is essential in order to evaluate the determinants (independent variables) associated with observed skills. Further, if policy makers wish to understand how different levels of prior schooling impact actual (measured) knowledge or performance levels in certain skill domains, they will need to make sure they have reliable data about relevant background variables, such as language spoken at home, social, and prior educational experiences. Impact studies will also require baseline and follow-up assessment procedures.

In sum, there are real choices to be made about literacy monitoring and evaluation through improved assessment models. To achieve improved literacy rates will depend in large part on how effective and timely are our M&E

tools. In a very real way, the UN Literacy Decade will be measured in practice only when effective tools for monitoring and evaluation can be agreed upon.

## Notes

1. ALEF is a program supported by the U.S. government through USAID and MEPI, and implemented by the Academy of Educational Development, a non-profit organization. The author served as a technical advisor on this project. Comments contained in this paper are in reference to the technical aspects of the work undertaken on literacy, and are do not represent the final conclusions of the project itself, which will be described in other publications downstream.
2. Further, the LLAT was created in two complete and parallel versions, such that each group of learners was randomly assigned Version 1 or Version 2; each province, therefore, had about half the learners take Version 1, and the other half take version 2 of the test. At the time of taking of the second testing of learners, the opposite version was to be utilized, so that learners would not be able to memorize (or communicate to others) the correct answers. In this design, any individual who took Version 1 at the Pretest (Time 1) was given Version 2 at the next testing time. Learners were tested longitudinally over five different testing times. Statistically, it was found that not only was internal reliability very high (the alpha score), but also the two Versions were found to be of equivalent difficulty. This latter result was useful in that a longitudinal study is improved in terms of validity when learners do not take identical two times in a row.

## References

- Bhola, H. S. 1999. Literacy Campaigns: A Policy Perspective. In: *Literacy. An International Handbook*, ed. by Wagner, D. A., Venezky, R. L., and Street, B. V. Boulder, CO: Westview Press.
- Carron, G., K. Mwiria, and G. Righa. 1989. *The Functioning and Effects of the Kenyan Literacy Programme*. IIEP Research Report No. 76. Paris: UNESCO/IIEP.
- Commeyras, Michelle, and Bagele Chilisa. 2001. Assessing Botswana's First National Survey on Literacy with Wagner's Proposed Schema for Surveying Literacy in the "Third World". *International Journal of Educational Development* 21: 433-446.
- Desjardins, R., and R. S. Murray. 2003. Introduction. *International Journal of Educational Research* 39(3): 173-180.
- Greaney, Vincent, S. R. Khandker, and M. Alam. 1998. *Bangladesh: Assessing Basic Learning Skills*. Washington, DC: World Bank.
- . 1999. *Bangladesh: Assessing Basic Learning Skills*. Washington: DC/Dhaka: World Bank.
- ILI/UNESCO. 1998. *Literacy Assessment For Out-of-School Youth and Adults*. ILI/UNESCO Technical Report from Expert Seminar, Paris, June 1998. Philadelphia: International Literacy Institute, University of Pennsylvania.

- . 1999. (D. Wagner, Principal Author). *Assessing Basic Learning Competencies in Youth and Adults in Developing Countries: Analytic Survey Framework and Implementation Guidelines*. ILI/UNESCO Technical Report. Philadelphia: International Literacy Institute, University of Pennsylvania.
- . June 2001. *LAP First Experts Meeting Report*. Philadelphia: ILI.
- . April 2002a. *LAP Second Experts Meeting Report*. Philadelphia: ILI.
- . April 2002b. *Analytic Review of Four LAP Country Case Studies*. Philadelphia: ILI.
- Kirsch, Irwin, A. Jungeblut, L. Jenkins, and A. Kolstad. 1993. *Adult Literacy in America: A First Look at the Results of the National Adult Literacy Survey*. Washington, DC: US Department of Education, National Center for Education Statistics.
- Lauglo, John. 2001. *Engaging with Adults: The Case for Increased Support of Adult Basic Education in Sub-Saharan Africa*. Washington DC: World Bank.
- Lavy, Victor, Jennifer Spratt, and N. Leboucher. 1995. *Changing Patterns of Illiteracy in Morocco: Assessment Methods Compared*. LSMS Paper 115. Washington DC: World Bank.
- Levine, Kenneth. 1998. Definitional and Methodological Problems in the Cross-National Measurement of Adult Literacy: The Case of the IALS. *Written Language and Literacy* 1: 41–61.
- Murray, T. Scott. 1997. Proxy Measurement of Adult Basic Skills: Lessons from Canada. In: *Adult Basic Skills: Innovations in Measurement and Policy Analysis*, ed. by Tuijnman. Cresskill, NJ: Hampton Press.
- Murray, T. Scott, Irwin S. Kirsch, and L. B. Jenkins (eds.). 1998. *Adult Literacy in OECD Countries: Technical Report on the First International Adult Literacy Survey*. Report NCES 98-053. Washington DC: US Department of Education.
- National Research Council. 2005. *Measuring Literacy. Performance Levels for Adults*. Washington, DC: National Academies Press.
- Nordtveit, Bjorn H. 2004. *Managing Public–Private Partnership: Lessons from Literacy Education in Senegal*. Washington DC: World Bank.
- OECD/Statistics Canada. 1995. *Literacy, Economy and Society*. Paris: OECD.
- . 1997. *Literacy Skills for the Knowledge Society: Further Results from the International Adult Literacy Survey*. Paris: OECD.
- . 2000. *Literacy in the Information Age*. Paris: OECD.
- Okech, Albert, R. Carr-Hill, A. R. Katahoire, T. Kakooza, A. N. Ndidde, and J. Oxenham. 2001. *Adult Literacy Programs in Uganda*. Washington DC: World Bank.
- Reder, S. 2000. What Does the NALS Measure? Issues in Dimensionality and Construct Validity. In *Literacy for the 21st Century: Research, Policy and Practice*, ed. by M. C. Smith. Boulder: Greenwood.
- Street, Brian V. 1996. *Literacy Economy and Society: A critique*. Working Papers on Literacy, No. 1. Montreal: Center for Literacy.
- Terryn, B. 2003. *The Literacy Assessment and Monitoring Programme (LAMP): Background and Approach*. Paper presented at ANLCI Conference in Lyon, France.
- UNESCO. 1983. *Statistics on Educational Attainment and Literacy*. Paris: UNESCO.

- . 1990. *Final Report on the World Conference on Education For All: Meeting Basic Learning Needs, Jomtien, Thailand*. Paris: UNESCO.
- . 2000. *The Dakar Framework for Action: Education for All: Meeting our Collective Commitments*. Adopted by the World Education Forum, Dakar, Senegal, 26–28 April 2000. Paris: UNESCO.
- . 2005. *EFA Global Monitoring Report 2006, Literacy for Life*. Paris: UNESCO.
- United Nations Statistical Office (UNSO) (Principal Authors Wagner, D. A., and A. B. L. Srivastava). 1989. *Measuring Literacy through Household Surveys* (Doc. No. DP/UN/INT-88-X01/10E). New York: United Nations Statistical Office.
- Venezky, R. L. 1992. *Matching Literacy and Social Policy*. Technical Report. Philadelphia, PA: National Center on Adult Literacy.
- Wagner, Daniel A. 1990. Literacy Assessment in the Third World: An Overview and Proposed Schema for Survey Use. *Comparative Education Review* 33(1): 112–138.
- . 1993. *Literacy, Culture and Development: Becoming Literate in Morocco*. New York: Cambridge University Press.
- . 1998. Literacy Retention: Comparisons Across Age, Time and Culture. In: *Global Prospects for Education: Development, Culture and Schooling*, ed. by H. Wellman and Scott G. Paris, 229–251. Washington, D.C: American Psychological Association.
- . 2000. *Literacy and Adult Education*. Global Thematic Review prepared for the U.N. World Education Forum, Dakar, Senegal. Paris: UNESCO.
- . 2001. Conceptual Dichotomies and the Future of Literacy Work Across Cultures. In: *Literacy and Motivation: Reading Engagement in Individuals and Groups*, ed. by C. Snow and L. Verhoeven. NJ: L. Erlbaum.
- . 2003. Smaller, Quicker, Cheaper: Alternative Strategies for Literacy Assessment in the UN Literacy Decade. *International Journal of Educational Research* 39(3): 293–309.
- . 2004. Literacy(ies), Culture(s) and Development(s): The Ethnographic Challenge. (A Five-Volume Review). *Reading Research Quarterly* 39(2): 234–241.
- . (ed.). 2006. *Monitoring and Evaluation of ICT in Education Projects: A Handbook for Developing Countries*. Washington, DC: World Bank/InfoDev.
- . 2008. *Mission Report in Morocco: ALEF Literacy Campaign, Monitoring and Evaluation*. Washington, DC: AED/USAID.
- Wagner, Daniel A., and Robert Kozma. 2005. *New Technologies for Literacy and Adult Education: A Global Perspective*. Paris: UNESCO.
- Wagner, Daniel A., R. L. Venezky, and Brian V. Street (eds.). 1999. *Literacy: An International Handbook*. Boulder, CO: Westview Press.
- Ziegahn, Linda. 1992. National Literacy Assessments: Present Practice and Alternatives. *International Journal of Educational Development* 12: 223–232.

### The author

**Daniel A. Wagner** is Professor of Education and Director of the International Literacy Institute, co-founded by UNESCO and the University of Pennsylvania. He is also

Director of the National Center on Adult Literacy at the University of Pennsylvania. He has been a visiting fellow at the International Institute of Education Planning in Paris, a visiting professor at the University of Geneva, and a Fulbright Scholar at the University of Paris. His most recent multi-year projects have been in India, South Africa, and Morocco. Dr. Wagner has more than 150 professional publications, and currently specializes in literacy, basic education, and new information technologies.

*Contact address:* International Literacy Institute, University of Pennsylvania, 3910 Chestnut Street, Philadelphia, PA 19104-3111, USA. E-mail: [wagner@literacy.upenn.edu](mailto:wagner@literacy.upenn.edu).