

Chapter 8

Smaller, quicker, cheaper: alternative strategies for literacy assessment in the UN Literacy Decade[☆]

Daniel A. Wagner*

*National Center on Adult Literacy/International Literacy Institute, University of Pennsylvania,
3910 Chestnut St., Philadelphia, PA 19104-3111, USA*

Abstract

In the context of the UN Literacy Decade (declared in February 2003), the present paper suggests three parameters that should be considered when new tools for assessment are considered in less developed countries (LDCs), each of which poses a special challenge to international comparative literacy assessment, such as in the International Adult Literacy Survey (IALS):

- *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. International comparative studies often run counter to this perspective.
- *Quicker*: Literacy assessments need to be completed in ‘real time’ so that results can affect policy and spending in the ‘lifetime’ of current ministerial appointments. Studies that take 3–5 years to generate results, even if robust, nonetheless fail to meet the test of timeliness.
- *Cheaper*: LDCs cannot afford either the fiscal or human resources costs of deep involvement in highly technical assessment exercises. The higher the cost, the more difficult to get to an initial ‘yes’ to participate in such an exercise, and the more difficult to gather time-series data to follow policy decisions.

In sum, this paper finds that there is a very important need for improving literacy assessment methodologies and the empirical database in developing countries, especially in light of the new UN Literacy Decade. While the IALS presents interesting and important options for

[☆]This report was supported in part by the U.S. Department of Education/Office of Vocational and Adult Education (ED-01-R-0023) to the University of Pennsylvania, under TECH21/UN Literacy Decade modification, as well as by the Spencer Foundation.

*Fax: +1-215-898-9804.

E-mail address: wagner@literacy.upenn.edu (D.A. Wagner).

methodological consideration, it also has a number of inherent limitations as discussed herein. Other options exist which should also be considered, especially for poor countries, such as the smaller/quicker/cheaper (SQC) approach.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Literacy; Measurement of learning achievement; Comparative assessment; New methods

1. Introduction

Since the World Conference on Education for All (WCEFA), held in Jomtien in 1990, the measurement of learning achievement (and more specifically, literacy and other basic skills) in developing countries has been recognized as critical to judging the quality of education programs worldwide, and in developing countries in particular. With the arrival of the UN Literacy Decade in February 2003, a principal goal is to reduce illiteracy by one-half current rates in every country. To achieve that goal will necessitate, naturally, ways to measure how far we have come in meeting rate reductions, putting a premium on empirical data collection. The International Adult Literacy Survey (IALS) is the best known of current methodologies for international assessments, and is the main subject of the papers in this Special Issue.

In the present paper, the author was asked to reflect on the IALS in the context of the UN Literacy Decade, with a special focus on developing countries. Of course, one option is to utilize the IALS or similar instruments to reach out to an extended number of countries beyond the OECD countries in which the IALS has mainly been focused upon. Yet, there seem to be some limitations to such an approach based on the current IALS methodologies, problems that go beyond some of the technical and conceptual critiques that have already been leveled at the IALS.

The present paper suggests three main parameters that should be considered when new tools for assessment are considered in less developed countries (LDCs), each of which poses a special challenge to the IALS-type international comparative study:

- *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. International comparative studies often run counter to this perspective.
- *Quicker:* Literacy assessments need to be completed in ‘real time’ so that results can affect policy and spending in the ‘lifetime’ of current ministerial appointments. Studies that take 3–5 years to generate results, even if robust, nonetheless fail to meet the test of timeliness.
- *Cheaper:* LDCs cannot afford either the fiscal or human resources costs of deep involvement in highly technical assessment exercises. The higher the cost, the more difficult to get to an initial ‘yes’ to participate in such an exercise, and the more difficult to gather time-series data to follow policy decisions.

This paper suggests that there is a very important need for improving literacy assessment methodologies and the empirical database in developing countries, especially in light of the new UN Literacy Decade. While the IALS presents interesting and important options for methodological consideration, and also has a number of inherent limitations as discussed herein. Other options exist which should also be considered, especially for poor countries. Before arriving at the above conclusions, it is first necessary to provide some background on key aspects of the assessment of adult literacy in LDCs.¹

2. Cultural dimensions to comparative literacy assessment

Cross-national comparisons of literacy acquisition and reading skills are of rather recent vintage, beginning with Downing's (1973) seminal *Comparative Reading to Scribner and Cole's* (1981) well-known research on literacy in Liberia, to the more recent International Association for the Evaluation of Educational Achievement (IEA) study of reading in 32 school systems undertaken under the auspices of the IEA (Elley, 1992), and finally to the IALS studies reported in this Special Issue (OECD & Statistics Canada, 1995, 1997, 2000). We know that literacy is not only difficult to define in individuals and delimit within societies, but it is also charged with emotional and political meaning. It was not long ago that newspapers and scholars assigned to whole societies a single referent such as "illiterate and uncivilized"; and, "illiterate" is still a term which carries a negative connotation around the world.

Defining literacy as an individual cognitive ability was once thought to be simple: it entailed the testing of reading and writing skills. As is done in some present-day societies for statistical and methodological expediency, literacy may be simply inferred from school attendance: those with 4 (or 8 or 12) years of formal public schooling are assumed to be literate. Or, in yet other societies, literacy rates are calculated from the numbers of persons who answer "yes" to the simple question "Can you read and write?" It is now known that such approaches to presumed literacy may be quite misleading, for a host of reasons. Furthermore, when considered as a cultural—as opposed to a cognitive—phenomenon, literacy is even less well defined, since its meanings, functions, and methods of transmission may vary greatly from one cultural group to the next.

As discussed elsewhere (Wagner, 1998), the distinction between "emic" and "etic" concepts is a central feature in comparative studies. *Emic* concepts are those that can be understood only within a single cultural system or society, and are measured only according to criteria relevant to and understood within that single system. *Etic* concepts are those which are deduced or derived from a position outside of any particular system, and have as a primary goal the analysis of more than a single

¹Some of the present text is derived from earlier work by the present author published by ILI & UNESCO (1998, 1999) and Wagner (1997, 2000).

social system or society. For example, an etic perspective on literacy assumes that skills such as decoding, word-picture-matching, and reading a bus schedule ought to have substantially the same meaning to different individuals and across different cultural groups. An emic perspective on literacy would encompass skills and meanings associated with literacy within cultural groups, such as “script recognition” skills in different orthographies, as well as the knowledge about the values, meanings, uses, and even attitudes of/about print in everyday social contexts. Simply put, emic literacy skills are those which can only be adequately understood within a given society; etic skills are those that have developed out of the heuristic convenience of those who desire a common frame of reference or system of measurement.

It may be fairly assumed that the IALS methodology, like most large-scale surveys, employs a primarily etic approach to the study of literacy. This etic approach utilizes external, quantifiable, comparison-oriented measures that are clearly important in understanding how people acquire literacy, how educators and policymakers view literacy, and how economic and societal systems interact in an increasingly interdependent world. Throughout the 20th century, literacy has been thought to be associated with formal schooling, student achievement and economic success. Thus, it is not surprising that quantifiable tools are needed to compare children with other children, adults with other adults, and social and cultural groups with other such groups, depending on policy and programmatic needs. Nonetheless, the psychometric measurement of literacy does not necessarily imply a particular methodology, and that is a subject to which we return further below.

3. International statistics on literacy

In order to provide worldwide statistical comparisons, UNESCO has relied to date almost entirely on data provided by its member countries (UNESCO, 1983). These countries, in turn, typically rely on national censuses, which most often determine literacy ability by self-assessment questionnaires and/or by means of a proxy variable utilizing the number of years of primary schooling. Because of doubts about the reliability of such data—especially in the light of rapid social, economic and demographic changes—considerable concern has been expressed about the credibility of literacy statistics. In 1986, for example, UNESCO and the United Nations Statistics Office (UNSO) held a joint seminar in Paris to discuss the use of household surveys to improve 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 (e.g., ILI & UNESCO, 1999).

While many specialists now agree that exclusive reliance on indirect measures of literacy may be flawed, there is renewed discussion of the utility of proxy measures (Murray, 1997; Introduction of this issue). More importantly, there have been changes in the way that some countries have begun to view literacy measurement,

especially in North American national surveys, and in the OECD countries (OECD & Statistics Canada, 1997).

Of course, the direct measurement of basic skills in youth and adults can be technically challenging and costly. For these reasons and others, most countries (especially in the developing world) have not as yet chosen to invest in local, national or regional direct measurement survey strategies. Furthermore, it must be recognized that any change in the methodology used for calculating literacy rates in a population might result in uncomfortable political consequences; this happened not long ago in the United States when the ‘official’ US literacy rate of over 95%, as a consequence of redefinition and survey assessment via the US National Adult Literacy Survey (NALS; Kirsch, Jungeblut, Jenkins, & Kolstad, 1993) was reduced to about 75% of the adult population. Similar adjustments (in a downward direction) were a consequence of the IALS study in 1995 (OECD & Statistics Canada, 1995), and eventually led to the withdrawal of France from the final report.

Overall, the gathering of international statistics on literacy poses a variety of problems which have been matters of debate among specialists. With the advent of greater need for direct measurement, and increased technical capacity for such measurement in developing countries, some clarification of issues seems to be taking place, as described below.

3.1. Language policy and multilingualism

Most countries have formulated an explicit language policy which typically states which language or languages have official status. Often the decision on national or official language(s) is based on such factors as major or dominant linguistic groups, colonial or post-colonial history, and the importance of a given language to the interests of economic development. Official languages are also those most commonly used in primary school, although there may be differences between languages used in beginning schooling and those used later on. Further, there may be important differences between language policy in primary schooling and that of non-formal education (NFE) and adult education. For example, in Senegal, French is used exclusively in primary school, but local Senegalese languages are used in adult literacy programs nationwide. The use of mother tongue instruction in primary and adult education remains a topic of continuing debate (Engle, 1975; Wagner, 1992).

Which languages ‘count’ in literacy statistics is also controversial. When UNESCO gathers literacy statistics, it typically depends on government census bureaus to provide the number of literates and illiterates, as well as age and gender differences. Apparently, the language of literacy is not always specified by the countries concerned. That is, it would be difficult to know, taking the case of Senegal again, how many ‘literate’ individuals are literate in French, Wolof, Arabic or other languages, and how this was broken down by age and gender.

While there is usually general agreement that all official language(s) ought to be assessed in a national literacy survey (e.g., English in the United States; English and French in Canada, and German, French, Italian and Romanch in Switzerland), there may be disagreement over the assessment of literacy in non-official or semi-official

languages, where these have a recognized and functional orthography (e.g., Athabaskan in Canada, or Hungarian in Romania). In many countries, there exist a multitude of local languages which have varying relationships and status with respect to the official language(s). How these languages and literacies may be included in a national literacy survey can be a matter of serious debate. For example, in certain predominantly Muslim countries in sub-Saharan Africa (e.g., Senegal or Ghana), the official language of literacy might be French or English, while Arabic—which is taught in Islamic schools and used by a sizable population for certain everyday and religious tasks—is usually excluded from official literacy censuses. Similarly, literacy in Chinese, Spanish, Cherokee and other written languages have generally been ignored in literacy assessments in the United States.

Both emic and other perspectives would suggest that all languages and scripts should be included in national literacy assessments in order to fully understand the range of human resources (and to maximize human rights). Practically speaking, this poses various problems in fiscal and technical resources. However, assuming sufficient support for undertaking a household literacy survey, effort should be made to investigate and assess literacy in all “significant” language populations, where significant might be defined by population size or percentage (see Wagner, 1990). This is essential especially because problems of illiteracy and low literacy tend to be over-represented in linguistic minority groups. Further, for comparative research, studies which include the diversity of the languages and literacies within a country provide a more accurate measure of social variation than those with a more limited focus.²

3.2. *International comparability of data*

The comparability of data—across time and context (see Wagner, 1994 for a discussion of ‘life-span’ and ‘life-space’ literacy)³—is a major concern for policy-makers 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, for example, the hallmarks of the UN data collection and the mission of the newly created UNESCO Institute for Statistics. Nonetheless, if comparability is the primary goal, while less attention is paid to the (local and cultural) validity of the definitions and classifications of literacy, then the data may become less meaningful and potentially less applicable at the ground level.

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 is measured in terms of the number of “literate” and “illiterate”. For most countries, this dichotomous type of classification presents few

²Nonetheless, it should be noted that it is not uncommon for governments to make arbitrary language and literacy “exclusions” from nationally representative samples for reasons of efficacy, politics, ethnic strife, and so forth.

³A similar concept is used by Desjardin (Chapter 3 of this issue) in terms of ‘life-wide’ literacy.

practical (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, the fact that this dichotomous literacy variable may be regarded as simplistic and flawed, places serious limits on its usefulness in comparative statistics.

For example, educational planners may want to know about the effects of the completion of primary or secondary schooling—or of a particular literacy campaign—on levels of literacy attainment. In these cases, a simple dichotomy is a blunt statistical instrument; literacy scores or levels are clearly required for learning achievement to be adequately measured. Furthermore, precise data are needed as to which languages and which literacies are used in each region and by ethnic group, in addition to age and gender variation. The collection of such data has largely been ignored by most national and international agencies to date (with some notable exceptions).

Compared to census-based literacy estimates, household level surveys offer considerable opportunity to create a detailed picture of literacy skill profiles and their demographic correlates in national populations. As detailed in the IALS survey discussed in this Special Issue, considerable cross-national comparability can be achieved, but not without various technical problems related to item comparability (Levine, 1998), and population sampling (Kalton, Lyberg, & Rempp, 1998).

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. The 1994 US NALS found, for example, that 50% of minority group adults in America with a 10th grade education could read at only a 5th grade level. Issues of equity in the US continue to dominate educational policy, as minority youth are, on average, about 4 years behind achievement levels of same-age White children, even up through college years (a Black college graduate in the US reads about the level of a White high school graduate). In a recent World Bank national household survey in Bangladesh, Greaney found that 5 years of primary schooling resulted in only a first grade equivalent of learning achievement, and that 3 years of schooling had approximately zero value in terms of learning achievement (Greaney, Khandker, & Alam, 1998). This study may have an important impact on the kinds of investments that Bangladesh makes in the area of basic education in the future.

3.3. Measuring the consequences of literacy

Will a change from a lower to a higher level of literacy skill make a concrete difference in an individual's life? Looking at the "average literacy rate" and comparing this statistic with health indicators (such as infant mortality rates or fertility rates), or estimating "employability" from such a rate, cannot adequately illuminate the diversity of individual human conditions. The author's own work in rural and low-literate African countries has demonstrated that those with higher

literacy tend to be better off economically (Wagner, 1993; the IALS found similar evidence across OECD countries OECD & Statistics Canada, 1995).

The issue, it would seem, is not so much whether, but rather *how* to promote literacy in a way that is consistent with overall policy objectives. How can one most efficiently achieve increased literacy levels (and important social and economic impacts) within the available economic and social resources? Here, the emphasis is on efficiency and appropriateness, domains in which specialists—with appropriate measures of learning/literacy achievement—can play a crucial role. What is efficient and appropriate, of course, brings forth a whole series of qualifiers, such as: for whom, in which language, for what purpose, using what methods and so on.

There is also the issue of program impact, at the local level. Various agencies have made serious investments in NFE and literacy programs in the context of their development strategies for women, youth and disadvantaged adults. The measurement of impact of these programs not only on learning achievement, but also on social and economic outcomes, has become a matter of some urgency. The development of low-cost measures of assessment in this regard might involve not only learning achievement but also other indicators (such as attendance, attitudes and so forth) that pose special challenges, from technical and statistical perspectives. Finally, there is also the important question of literacy retention—how much learning is retained by learners for use downstream, a concept that is probably over-emphasized given the paucity of hard data on the subject (Wagner, 1998).

4. International, national and program level needs

Up until the IALS studies, UNESCO, as noted above, collected data on adult literacy from ‘national level’ data sources—whether through traditional censuses and proxy measures, or through household sample designs. The general goal of national level data collection includes at least the following practical purposes: comparison of literacy rates over time, determination of statistical relationships between literacy and other variables of interest (e.g., gender, years of schooling, etc.), and cross-national comparability (the latter of which was very much in doubt due to the diversity of data collection methods on the ground). The advent of the IALS in the 1990s and up to the present (such as in the present Special Issue) shows that both international agencies and an increasing number of countries are interested in cross-national comparisons.

Nonetheless, there are other major assessment needs in literacy work. Most of these would fall under the rubric of *program* level impact, planning and development, at the local or regional level. The specific purposes of assessment and evaluation at the program level include the following: comparison of program effectiveness over time; comparison of one type of program with others; effectiveness of teacher training; and so forth. The emphasis in program assessment and evaluation is on ‘how programs work’, rather than on national policy planning at the national or global levels. At the program level, there are real and frequent needs to know what kinds of learning achievement impacts are being realized.

In the domain of non-formal basic education, adult education and literacy education, there has been a tendency to assume that the two broad measurement needs—international/national and local/program—would use quite different tools or methods for assessment. Indeed, this has most often been the case for the most part, and it is exemplified in the work up to the present in the IALS studies. Further below, it will be suggested that there is no particular reason for this situation to remain unaltered. Indeed, there are major gains to be made if the methodologies—and purposes—of international/national and local assessments were conflated, particularly in the sense of shared perspectives, tools, human and technical resources.

In sum, there is a considerable need to take a fresh look at the complementarities between international/national and local program level information needs, using surveys of direct measurement of learning.

5. Brief overview of methodologies for literacy assessment

As noted above, there have been many efforts to measure literacy achievement over past decades. Given the diverse actors in the literacy and learning achievement fields, there exist a wide variety of tools for assessment; but there is also some confusion with regard to terminology employed to describe such methods.

Assessment refers to one or more methods for judging (via some type of empirical observation or test) the actual performance of literacy or other cognitive skills. For example, the UNSO (1989) and IALS (1995) surveys included literacy assessment, as did the household survey of basic learning achievement in Bangladesh (Greaney et al., 1998). Of course, with the wider definition of WCEFA mentioned above, assessment might have to take into account broader areas of knowledge, attitudes and even values.

Household based surveys offer considerable potential in the field of literacy and basic skills learning achievement. This is the case in large part due to the many options for flexibility in design, such as: population sampling, cognitive knowledge sampling and cost management (cost can be traded off against gains in statistical reliability). For other trade-off issues, see Kolstad (1997). For example, with a focus on developing countries, Wagner and Srivastava (UNSO, 1989) produced a relatively low-cost method for the development of household surveys of literacy, based on a national household survey in Zimbabwe. While the analyses of the Zimbabwe survey were never fully completed, aspects of the methodology were adapted by others for household survey use, especially by the World Bank in Morocco (Lavy, Spratt, & Leboucher, 1995), and in Bangladesh (Greaney et al., 1998). Each of these studies utilized a national (as opposed to an international comparative) framework to produce items that could, with reasonable reliability, assign individuals to levels of literacy and numeracy achievement across multiple languages.

The IALS approach (OECD & Statistics Canada, 1995, 1997, 2000; and its predecessors, such as the 1993 US NALS) 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 especially Item Response Theory or IRT; see Chapter 1 of this issue). 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 collinearity 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](#)).

Evaluations of literacy programs in developing countries are numerous, of both formative and summative varieties ([Bhola, 1990](#); [Carron, Mwiria, & Righa, 1989](#); [Okech et al., 2001](#)). In developing countries especially, such evaluations have only rarely included psychometrically appropriate tests for measuring learning achievement ([Ziegahn, 1992](#)). Typically what is measured are the inputs in human and infrastructural resources, the pedagogical methods employed, and the outcomes in terms of attendance and ‘successful program completion’. In industrialized countries as well, there is insufficient attention to the impact of programs on learning achievement ([Wagner & Venezky, 1999](#)). This is a very serious gap in the current knowledge base.

6. Challenges to IALS and other international comparative approaches

All measures of human behavior are subject to what have been termed ‘challenges’ to their validity. This is nowhere more the case than in literacy education. In part, this is historical, in that literacy work has often been tied to politics, and thus not subject to the ordinary checks and balances of social science research and evaluation studies. Furthermore, since literacy has been seen over time as a skill that can ‘easily be taught’ (e.g., in ‘each one, teach one’ campaigns), it is only relatively recently (the last two decades) that the efficacy and efficiency of literacy campaigns and programs have come into question (and more recently, the same being true of formal schooling as well).

If literacy is so deeply tied to the integral definition of separate cultures, and if a broad definition such as that proposed by the IALS is used, then how can common understandings of literacy levels or competencies be reached across cultures? That is, literacy in poor communities may be viewed (and, in practice, used) in quite different ways than is the case in wealthy communities (see Levine’s critique of IALS, 1998). Indeed, the definition of ‘poor’ may well have a radically different operational definition in different societies, and often does. Further, when one attempts to compare individuals or groups on whether they are self-defined as “literate”, it can emerge that different definitions preclude valid comparison. This is often the case with census-derived information where large differences in schooling bias the decision matrix for responding to the question, “Can you read and write?” (cf. [Lavy et al., 1995](#)).

There are numerous technical challenges to any method of testing. These range from disputes about theoretical and operational definitions to the type of statistical tests employed to analyze data. Given space limitations, it is useful to simply note some areas that have been particularly debated in studies involving IALS assessment methodology. These include: (a) scales of literacy achievement (from dichotomous, to five 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, Kirsch, & Jenkins, 1998](#) and others); 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 selection is in fact more complex than stated above. 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 LDCs 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 shifts, the cost is high relative to local LDC budgets and/or opportunity costs. National or local household surveys can also have a similar policy impact (as in the US NALS), but this result necessitates a serious and credible 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 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 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 studies with cost figures.⁴ (This is also

⁴According to [Greaney and Kellaghan \(1996\)](#) the typical cost components of (national) assessments is roughly as follows (as percent of cost): instrument development 15; sampling and selection 10; data collection 30; data processing 10; data analysis 15; reporting and dissemination 15; and governance 5.

true with respect to costs of literacy programs themselves, as described in [Wagner, 1995](#).)

7. Moving towards smaller, quicker, cheaper approaches

In the popular media and in academic circles, it is not uncommon to hear that low levels of literacy and basic skills are among the chief problems facing contemporary societies today. And, there are numerous arguments which would support concern for such a point of view, ranging from the economic pressure on the workplace, to increases in women's health and in farmer productivity in developing countries, and the major problems endemic in urban secondary schools in Europe. Literacy and basic skills development are part of policy discussions in all of these areas, yet needed data that would help to empirically address key policy issues have often been lacking. Improved methods of literacy assessment in out-of-school youth and adults can help fill this gap.

It is also 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 ([Wagner, Venezky, & Street, 1999](#)).

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 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 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 interests for those of the 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.

The above review and analysis has led to the importance of seeking alternatives to the IALS or other international comparative approaches to literacy assessment. For simplicity, we call this the smaller/quicker/cheaper (or SQC) approach, as follows.

7.1. *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. ‘Smaller’ here 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 humans 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.

Nonetheless, many international comparative studies, such as the IALS, seem to run counter to this perspective. That is, the lowest common denominator is often to ‘go large’ so as to be sure to have enough data, especially if some data are ‘tainted’ for one reason or another. This rationale is reasonable and plausible, but would be much less likely to occur if the study was done on a smaller scale if funding were invested in more careful analysis.

Just as importantly, large international studies often cannot literally afford to focus on those ‘most in need’, such as minority ethnic groups. Here the emphasis on large and comparative may work against the stated “pro-poor” policies of international agencies such as the World Bank and UNESCO. If strong empirical science—as exemplified by numerous countries, large data sets and heavy inferential statistics—works against policy goals, then these approaches need to be given some reconsideration.

7.2. *Quicker*

Literacy assessments need to be completed in ‘real time’ so that results can affect policy and spending in the ‘lifetime’ of current ministerial appointments. Studies that take 3–5 years to generate results, even if robust, nonetheless fail to meet the test of timeliness. The NALS, IALS and IEA studies, just to name a few 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 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?

Obviously, there is middle ground that can be achieved in a number of ways. First, by focusing on national and sub-national studies, the expensive process (in time and money) of comparative ‘equity’ can be significantly reduced. Second, even at the

national or sub-national levels, greater attention can be paid to producing results that matter in the near-term, even if the larger and finalized reports take much longer. In other words, there are trade-offs that may be necessary in order for relevance and practicality to be achieved.

7.3. *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. But seen from a national or local perspective, things may be quite different. LDCs cannot afford either the fiscal or human resources costs of deep involvement in highly technical assessment exercises. The higher the cost, the more difficult to get to an initial ‘yes’ to participate in such an exercise, and the more difficult to gather time-series data to follow policy decisions.

Cheaper also means here the possibility of using local (rather than external) consultants. This can be achieved if the assessments are not constrained to use outside agencies in the US or Europe to process complex data sets. There are alternatives that are both reliable and valid (see Wagner, 2002) that are suited to such local needs and which are 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 the past. 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 contexts than on an international level.

How can both comparability and context sensitivity be appropriately balanced in literacy and basic skills assessments? Can the tools developed for household surveys be ‘shared’ with program evaluation and vice versa? What kinds of savings can be made in human and fiscal resources across these two domains? The notion of *shareability* may be one way to find a compromise between etic approaches to strict (international) comparability on the one hand, and emic approaches to strict national, subnational and local contextuality on the other. By shareability, we mean the possibility of sharing tools and methods across international, national and local assessments (see ILI & UNESCO, 1998, 2002).

8. Conclusions

There is little question that the IALS and international assessments have now put the empirical assessment of literacy on the policy map, as has been suggested for many years (see Wagner, 1990). Yet the successes of IALS have also pointed to some of its limitations both in design and in promise. With a focus on the poorest of the poor in the UN Literacy Decade, an argument has been made here for not only

pursuing the IALS approach, but for also considering approaches of a lesser scale that are more culturally and locally focused. What we are suggesting is not ethnography (nonetheless, see also Wagner, 2004, for some positive dimensions of ethnography within assessments), but rather a more ‘tailored’ approach to surveys that are the “right size” for the questions addressed.

To summarize, the SQC alternative approach would be:

(a) *Smaller and more focused*

- Focus on national and local interests (avoid international comparability).
- Collect focused/strategic samples—reach the unreached.
- Just enough psychometric items to give operational and/or policy guidance.

(b) *Quicker*

- Speed of delivery is important for policy makers.
- Shorter time horizons.
- Rapid prototyping of simpler tests.
- Frequency is important.

(c) *Cheaper*

- Needs to be affordable.
- “Shareability” can produce economies of scale.
- Transparency and simplicity.
- Capacity building is easier.
- Sustainability: lower cost means that replication is more affordable.

To move forward on the goals of the UN Literacy Decade, the SQC approach would allow us to more effectively reach the unreached and un-(-under) schooled (gender, minorities); obtain 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 larger/slower/expensive international comparative studies. But, one might say, they are not the only game in town; SQC approaches are likely to play an increasingly important role in the future.

References

- Bhola, H. S. (1990). *Evaluating ‘literacy for development’: Projects, programs and campaigns*. Hamburg: UIE/DSE.
- Carron, G., Mwiria, K., & Righa, G. (1989). *The functioning and effects of the Kenyan literacy program*. IIEP Research Report No. 76. Paris: UNESCO/IIEP.
- Downing, J. (1973). *Comparative reading*. New York, NY: Macmillan.
- Elley, W. B. (1992). *How in the world do students read?*. Hamburg: International Association for the Evaluation of Educational Achievement.

- Engle, P. L. (1975). Language medium in early school years for minority language groups. *Review of Educational Research*, 45, 283–325.
- Greaney, V., & Kellaghan, T. (1996). *Monitoring the learning outcomes of education systems*. Washington, DC: World Bank.
- Greaney, V., Khandker, S. R., & Alam, M. (1998). *Bangladesh: Assessing basic learning skills*. Washington, DC: 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.
- ILI & UNESCO (Wagner, D., Principal author). (1999). *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.
- ILI & UNESCO. (2002). *LAP analytic review of four country case studies*. Philadelphia: ILI.
- Kalton, G., Lyberg, L., & Remp, J.-M. (1998). Review of methodology. In: T. S. Murray, I. S. Kirsch, & L. B. Jenkins (Eds.), *Adult literacy in OECD countries: Technical report on the first International Adult Literacy Survey*. Report NCES 98-053. Washington, DC: US Department of Education (Appendix A).
- Kirsch, I., Jungeblut, A., Jenkins, L., & Kolstad, A. (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.
- Kolstad, A. (1997). Information trade-offs in large scale assessments. In A. Tuijnman, I. Kirsch, & D. A. Wagner (Eds.), *Adult basic skills Innovations in measurement and policy analysis*. Cresskill, NJ: Hampton Press.
- Lavy, V., Spratt, J., & Leboucher, N. (1995). *Changing patterns of illiteracy in morocco: Assessment methods compared*. LSMS Paper 115. Washington, DC: The World Bank.
- Levine, K. (1998). Definitional and methodological problems in the cross-national measurement of adult literacy: The case of the IALS. *Written Language and Literacy*, 1(1), 41–61.
- Murray, T. S. (1997). Proxy measurement of adult basic skills: Lessons from Canada. In A. Tuijnman, I. Kirsch, & D. A. Wagner (Eds.), *Adult basic skills: Innovations in measurement and policy analysis*. Cresskill, NJ: Hampton Press.
- Murray, T. S., Kirsch, I. S., & Jenkins, L. B. (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.
- OECD & Statistics Canada. (1995). *Literacy, economy and society*. Paris: OECD.
- OECD & Statistics Canada. (1997). *Literacy skills for the knowledge society: Further results from the International Adult Literacy Survey*. Paris: OECD.
- OECD & Statistics Canada. (2000). *Literacy in the information age*. Paris: OECD.
- Okech, A., Carr-Hill, R. A., Katahoire, A. R., Kakooza, T., Ndidde, A. N., & Oxenham, J. (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 M. C. Smith (Ed.), *Literacy for the 21st century: Research, policy and practice*. Boulder: Greenwood.
- Scribner, S., & Cole, M. (1981). *The psychology of literacy*. Cambridge: Harvard University Press.
- UNESCO. (1983). *Statistics on educational attainment and literacy*. Paris: UNESCO.
- United Nations Statistical Office (UNSO) (Wagner, D. A., & Srivastava, A. B. L., Principal authors). (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, D. A. (1990). Literacy assessment in the Third World: An overview and proposed schema for survey use. *Comparative Education Review*, 33(1), 112–138.
- Wagner, D. A. (1992). *Literacy: Developing the future*. UNESCO Yearbook of Education. Paris: UNESCO.

- Wagner, D. A. (1993). *Literacy, culture, and development: Becoming literate in Morocco*. New York: Cambridge University Press.
- Wagner, D. A. (1994). Life-span and life-space literacy: National and international perspectives. In D. Keller-Cohen (Ed.), *Literacy: Interdisciplinary conversations*. Cresskill, NJ: Hampton Press.
- Wagner, D. A. (1995). Literacy and development: Rationales, myths, innovations, and future directions. *International Journal of Educational Development*, 15, 341–362.
- Wagner, D. A. (1997). Adult literacy assessment in comparative contexts. In A. Tuijnman, I. Kirsch, & D. A. Wagner (Eds.), *Adult basic skills: Innovations in measurement and policy analysis*. Cresskill, NJ: Hampton Press.
- Wagner, D. A. (1998). Literacy retention: Comparisons across age, time and culture. In H. Wellman, & G. P. Scott (Eds.), *Global prospects for education: Development, culture and schooling*. Washington, DC: American Psychological Association.
- Wagner, D. A. (2000). *Literacy and adult education*. Global Thematic Review prepared for the U.N. World Education Forum, Dakar, Senegal. Paris: UNESCO.
- Wagner, D. A. (2002). *Towards guidelines for the improvement of literacy assessment in developing countries: Conceptual dimensions based on the LAP project*. Philadelphia: ILI & UNESCO.
- Wagner, D. A. (2004). Literacy(ies), culture(s) and development(s): The ethnographic challenge (a five-volume review). *Reading Research Quarterly*, 39(2), 234–241.
- Wagner, D. A., & Venezky, R. L. (1999). Adult literacy: The next generation. *Educational Researcher*, 28(1), 21–29.
- Wagner, D. A., Venezky, R. L., & Street, B. V. (Eds.), (1999). *Literacy: An international handbook*. Boulder, CO: Westview Press.
- Ziegahn, L. (1992). National literacy assessments: Present practice and alternatives. *International Journal of Educational Development*, 12, 223–232.

Daniel A. Wagner is Professor of Education and Director of the National Center on Adult Literacy at the University of Pennsylvania, which includes the federally funded US National Technology Laboratory for Literacy and Adult Education. He is also Director of the International Literacy Institute, co-founded by UNESCO and the University of Pennsylvania. He received his Ph.D. in psychology at the University of Michigan, was a 2-year post-doctoral fellow at Harvard University, a Visiting Fellow at the International Institute of Education Planning in Paris, a Visiting Professor at the University of Geneva (Switzerland), and a Fulbright Scholar at the University of Paris.