

1
2
3
4
5
6

ADULT LITERACY: MONITORING AND EVALUATION FOR PRACTICE AND POLICY

DANIEL A. WAGNER*

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

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



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

42 **Zusammenfassung – ERWACHSENENALPHABETISIERUNG: DATENERHE-**
43 **BUNG UND AUSWERTUNG IN DER PRAKTISCHEN ANWENDUNG** – Seit
44 Beginn der Alphabetisierungsdekade der UN 2003 erhalten Verfahren der Datenerhebung
45 und Auswertung in der Erwachsenenalphabetisierung zunehmende Bedeutung. Der
46 vorliegende Artikel bietet einen Überblick über die meistverwendeten Messverfahren
47 der Erwachsenenalphabetisierung auf nationaler und länderübergreifender Ebene. Ein
48 besonderer Schwerpunkt des Überblicks liegt im kontrastierenden Vergleich dreier
49 Modelle zur Alphabetisierungsbestimmung: (1) das "traditionelle" Modell, das vor-
50 wiegend auf Zensus und Zählverfahren beruht und das in den letzten Jahrzehnten
51 vielfach zur Erfassung nationaler und regionaler Analphabetenraten benutzt wurde; (2)
52 flächendeckende Messtechniken, wie sie im Zuge des International Adult Literacy
53 Survey und ähnlicher Modelle benutzt werden; und (3) eine Methode, die eine Art
54 Zwischenform zwischen den beiden genannten Modellen darstellt und die sich darum
55 bemüht, Umfang und Komplexität der Untersuchungsmethoden den jeweiligen pro-
56 grammatischen Fragestellungen möglichst "maßgeschneidert" anzupassen (das sog-
57 enannte SQC-Modell). Der vorliegende Artikel vertritt die Position, dass kein
58 bestimmtes Einzelmodell oder Methodenset die universelle Eignung zur Messung und
59 Kontrolle der Erwachsenenalphabetisierung in der Welt aufweist, sondern dass es die
60 Verbindung programmatischer und komparativer Ansätze wie im SQC-Modell ist, die
61 ein stärkeres Engagement und tiefere Einblicke in die Evaluation der Erwachsenen-
62 alphabetisierung erwarten lassen.

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

82 Varying approaches to monitoring and evaluation

83 The World Conference on Education for All (WCEFA), held in Jomtien,
84 Thailand, in 1990, found that the measurement of learning achievement was
85 critical to judging the quality of education programs around the world (UN-
86 ESCO 1990). This recommendation followed on decades of evaluation stud-
87 ies that often used “participation,” “funds spent,” or “contact hours” as
88 proxy variables for determining the quality of a nation’s efforts to improve
89 education. With the arrival of WCEFA, it was agreed that what the learner
90 actually learned should be at the center of the educational enterprise, and be
91 useful for both improving practice and implementation, as well as for policy
92 development.

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

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



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

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

121 **International statistics on literacy**

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

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



148 Because of doubts about the reliability of such data collection, considera-
149 ble concern has been expressed about the credibility of literacy statistics. In
150 1986, for example, UNESCO and the UN Statistics Office (UNSO) held a
151 joint seminar in Paris to discuss the use of household surveys to improve the
152 collection adult literacy statistics; a technical report which was the basis of
153 this seminar was later published (UNSO 1989); these discussions were
154 followed by similar ones over the next decade and a half (see, for example,
155 ILI/UNESCO 1999, 2001, 2002a, b). The concern, as noted above from the
156 1990 WCEFA, was whether actual skill learning had taken place in schools
157 and/or in adult literacy programs. The traditional method of monitoring
158 international literacy levels was clearly very limited since few countries both-
159 ered to actually *measure* individual skills in a large or broad enough popula-
160 tion sample to assure that literacy levels were valid and reliable.

161 Over nearly two decades, there has been a movement to add greater
162 statistical strength to the monitoring of adult literacy, and this has meant a
163 commensurate movement toward direct measurement or assessment of liter-
164 acy skill. The two terms – monitoring and measuring – are therefore comple-
165 mentary: one can monitor with either good or poor measures, but it is better
166 to have the best measures possible within available resources. This is much
167 easier said than done; see for example, the recent efforts at the US National
168 Academy of Sciences to address these same issues in America (NRC 2005).
169 In fact, even as many specialists now agree that exclusive reliance on tradi-
170 tional indirect measures of literacy may be flawed, there is renewed discus-
171 sion of the utility of proxy measures (Desjardins and Murray 2003; Murray
172 1997), since they are may be sufficient and cost less. Indeed, cost effective-
173 ness is an issue that must be kept clearly in mind, as will be discussed below,
174 since resources are always limited, and seeking to collect too much data may
175 be as counterproductive as collecting too little.

176 Policy issues in literacy measurement

177 *Learning achievement and information stakeholders*

178 At WCEFA in 1990, a number of educational targets relating to youth and
179 adults were agreed upon, including reducing by 50% the number of adult
180 illiterates by the year 2000 and improving learning achievement to an agreed
181 percentage of an appropriate age cohort. Unfortunately, as noted in the
182 *Literacy and Adult Education* thematic paper for the 2000 Dakar EFA meet-
183 ing (Wagner 2000), the Jomtien emphasis on learning achievement had not
184 succeeded in becoming a major part of literacy work in the decade between
185 1990 and 2000. By 2000, it was agreed that the field of adult literacy requires
186 both a greater focus on the quality of literacy services and better ways
187 to measure literacy achievement. In the Dakar Framework for Action



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

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

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

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

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

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

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



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

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

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

249 Naturally, policy goals will vary across countries, as well as by gender,
250 ethnic group and region (and more) within countries. Countries concerned
251 about overcoming inequalities caused by geographical disparities may want
252 to collect more information about type of housing or community, or dis-
253 tance from the nearest school, in order to be able to identify those living in
254 remote or in inaccessible areas. Countries concerned about inequalities by
255 gender, ethnicity, or language groups will need relevant variables that cap-
256 ture group membership in this regard. Finally, countries interested in relat-
257 ing the assessment results to specific programs or educational experiences
258 will have to include relevant questions concerning program curricula.

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

263 *International comparability of data*

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



272 meaningful and potentially less applicable at the ground level. This is a natu-
273 ral and essential tension between “emic” and “etic” approaches to literacy
274 measurement (Wagner 2004b).

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

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

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

306 *Moving from monitoring to direct assessment*

307 *The end of a dichotomy*

308 There is little doubt that the use of dichotomous variables in literacy work
309 have had a deleterious effect on the field in a number of ways: from the
310 initial equating of illiteracy with uncivilized, to the continued use of “literate”
311 versus “illiterate” in census data in many countries today (Wagner
312 1990, 2001). Indeed, one can possibly argue that we are all illiterate in one
313 way or another, with the merger of such terms as reading, health literacy
314 and technological literacy; or that we are all literate, in the sense that very



315 few people in today's world have little or know knowledge of the purposes
316 and nature of literacy in at least one script. In other words, there are a great
317 deal more shades of gray than black and white in contemporary discussions
318 of literacy and its measurement. This conclusion has the positive conse-
319 quence of demonstrating that, like formal schooling, adult literacy program-
320 ming deserves a great deal more investment than it is currently getting (for a
321 discussion on sub-Saharan Africa, see Lauglo 2001). With very little doubt,
322 the next decade will see end of present use the literacy-illiteracy dichotomy
323 in discussion of literacy M&E and policy.

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

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

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

345 Since the launch of the UN Literacy Decade, the UIS has launched an
346 initiative called LAMP (Literacy Assessment and Monitoring Program),
347 where they are planning to build on some of the tools developed through the
348 IALS, but refocused on literacy assessment in developing countries (Terryn
349 2003). While the effort is still in its early stages, LAMP holds the promise of
350 being able to employ some of the best technical tools in literacy assessment,
351 and adapt them for use in poor countries. Cautionary remarks about the
352 cost-benefit trade-offs in poor countries have been made by a number of
353 literacy specialists (e.g., ILI/UNESCO 1999; Street 1996; Wagner 2003).
354 Among the issues invoked in such remarks are the low degree of transpar-
355 ency of the data when collected using the IRT methodology, the expensive
356 process of creating methods of cross-cultural comparability, the long time
357 taken to complete a given study, and the overall cost of highly sophisticated
358 methods in poor countries.



359 *Challenges to the IALS model*

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

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

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

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



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

406 *Household surveys and program evaluation models*

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

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

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

436 *Choices in survey design*

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

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



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

- 447 • *Household surveys.* These allow considerable flexibility in terms of sample
448 selection and sample reduction (one can limit the sample size through
449 appropriate stratification), and thereby provide more time for actual
450 learning assessment. One limitation is that such surveys are often “stand-
451 alone,” and require specialized training as well as relatively skilled test
452 designers. As part of this same approach, a *special literacy survey* may be
453 undertaken with a particular focus on literacy assessment, but is not a
454 part of a census sample (household or otherwise). One example of this
455 approach was that undertaken by the World Bank (Greaney et al. 1999)
456 in rural Bangladesh, or as part of research projects (Wagner 1993). A
457 recent derivative of the household survey has been termed the SQC
458 approach that will be discussed in the next subsection.
- 459 • *Post-censal sample.* The LAP India case study cited the NSSO study that
460 was of the post-censal (after a census) type. One advantage is that the
461 same personnel and data structure can be used as in the census, such as in
462 the India case study in the LAP project (ILI/UNESCO 2002b), there are
463 advantages in being able to compare the data collected from the regular
464 census with the post-censal in-depth study. One limitation is that census
465 personnel often seem to have less technical capability in designing basis
466 skills assessment instruments.
- 467 • *Program evaluation.* Evaluations of literacy programs in developing coun-
468 tries are numerous, of both formative and summative varieties (Bhola
469 1990; Carron et al. 1989; Okech et al. 2001; Nordtveit 2004). In developing
470 countries especially, such evaluations have only rarely included
471 psychometrically appropriate tests for measuring learning achievement.
472 Typically, measurement focuses on the inputs in human and infrastructural
473 resources, the pedagogical methods employed, and the outcomes in terms
474 of attendance and “successful program completion.” A lack of skill assess-
475 ment is typically a very serious gap in the current program evaluation
476 knowledge base.
- 477 • *Other issues.* Among other issues is how to meet high-stakes versus low-
478 stakes concerns on the part of participants in the assessments. In-school
479 tests are often seen as high-stakes, but this is not the purpose of the cur-
480 rent literacy assessment exercise – yet participants often view testing as a
481 simple extension of schooling. Thus, care needs to be taken, or, as in
482 some of the case study examples, one may find data falsification on the
483 part of either learners or instructors or both. Finally, some procedures for
484 post-test adjustments for sampling biases may be undertaken, though
485 these may or may not be necessary, depending on the sample and testing
486 methodology chosen.



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

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

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

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

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



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

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

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

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



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

577 *Some cost considerations in the SQC model*

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

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

599 *Some limitations of the SQC model*

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



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

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

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

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

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



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

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

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

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

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

692 Conclusions

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



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

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

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

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

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



743 tools. In a very real way, the UN Literacy Decade will be measured in prac-
744 tice only when effective tools for monitoring and evaluation can be agreed
745 upon.

746 **Notes**

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

766

767

768 **References**

- 769 Commeyras, Michelle, and Bagele Chilisa. 2001. Assessing Botswana's First National
770 Survey on Literacy with Wagner's Proposed Schema for Surveying Literacy in the
771 "Third World". *International Journal of Educational Development* 21: 433-446.
- 772 Desjardins, R., and R. S. Murray. 2003. Introduction to Special Issue on *Literacy*
773 *Proficiency: Investigations Using Direct Measures of Skills*. *International Journal of*
774 *Educational Research*.
- 775 Greaney, Vincent, S. R. Khandker, and M. Alam. 1999. *Bangladesh: Assessing Basic*
776 *Learning Skills*. Washington: DC/Dhaka: World Bank.
- 777 ILI/UNESCO. 1998. *Literacy Assessment For Out-of-School Youth and Adults*. ILI/
778 UNESCO Technical Report from Expert Seminar, Paris, June 1998. Philadelphia:
779 International Literacy Institute, University of Pennsylvania.
- 780 ——. 1999. (D. Wagner, Principal Author). *Assessing Basic Learning Competencies in*
781 *Youth and Adults in Developing Countries: Analytic Survey Framework and Implemen-*
782 *tation Guidelines*. ILI/UNESCO Technical Report. Philadelphia: International Literacy
783 Institute, University of Pennsylvania.
- 784 ——. June 2001. *LAP First Experts Meeting Report*. Philadelphia: ILI.
- 785 ——. April 2002a. *LAP Second Experts Meeting Report*. Philadelphia: ILI.



- 786 —. April 2002b. *Analytic Review of Four LAP Country Case Studies*. Philadelphia:
787 ILI.
- 788 Kirsch, Irwin, A. Jungeblut, L. Jenkins, and A. Kolstad. 1993. *Adult Literacy in*
789 *America: A First Look at the Results of the National Adult Literacy Survey*. Washington,
790 DC: US Department of Education, National Center for Education Statistics.
- 791 Lauglo, John. 2001. *Engaging with Adults: The Case for Increased Support of Adult Basic*
792 *Education in Sub-Saharan Africa*. Washington DC: World Bank.
- 793 Lavy, Victor, Jennifer Spratt, and N. Leboucher. 1995. *Changing Patterns of Illiteracy in*
794 *Morocco: Assessment Methods Compared*. LSMS Paper 115. Washington DC: World
795 Bank.
- 796 Levine, Kenneth. 1998. Definitional and Methodological Problems in the Cross-
797 National Measurement of Adult Literacy: The Case of the IALS. *Written Language and*
798 *Literacy* 1: 41–61.
- 799 Murray, T. Scott. 1997. Proxy Measurement of Adult Basic Skills: Lessons from
800 Canada. In: *Adult Basic Skills: Innovations in Measurement and Policy Analysis*, ed. by
801 Tuijnman. Cresskill, NJ: Hampton Press.
- 802 Murray, T. Scott, Irwin S. Kirsch, and L. B. Jenkins (eds.). 1998. *Adult Literacy in*
803 *OECD Countries: Technical Report on the First International Adult Literacy Survey*.
804 Report NCES 98-053. Washington DC: US Department of Education.
- 805 National Research Council. 2005. *Measuring Literacy. Performance Levels for Adults*.
806 Washington, DC: National Academies Press.
- 807 Nordtveit, Bjorn H. 2004. *Managing Public–Private Partnership: Lessons from Literacy*
808 *Education in Senegal*. Washington DC: World Bank.
- 809 OECD/Statistics Canada. 1995. *Literacy, Economy and Society*. Paris: OECD.
- 810 —. 1997. *Literacy Skills for the Knowledge Society: Further Results from the*
811 *International Adult Literacy Survey*. Paris: OECD.
- 812 —. 2000. *Literacy in the Information Age*. Paris: OECD.
- 813 Okech, Albert, R. Carr-Hill, A. R. Katahoire, T. Kakooza, A. N. Ndidde, and
814 J. Oxenham. 2001. *Adult Literacy Programs in Uganda*. Washington DC: World Bank.
- 815 Oxenham, John. 2005. The Quality of Programmes and Policies Regarding Literacy and
816 Skills Development. Background Paper. *EFA Global Monitoring Report 2006*. Paris:
817 UNESCO.
- 818 Street, Brian V. 1996. *Literacy Economy and Society: A critique*. Working Papers on
819 Literacy, No. 1. Montreal: Center for Literacy.
- 820 —. 1999. The Meanings of Literacy. In: *Literacy: An International Handbook*, ed. by
821 D. A. Wagner, R. L. Venezky, and B. L. Street. Boulder, CO: Westview Press.
- 822 SubbaRao, I. V. 2002. *India: LAP Country Case Study*. Working Document Published
823 at <http://www.literacy.org>. Philadelphia: ILI.
- 824 Terry, B. 2003. *The Literacy Assessment and Monitoring Programme (LAMP):*
825 *Background and Approach*. Paper presented at ANLCI Conference in Lyon, France.
- 826 UNESCO. 1990. *Final Report on the World Conference on Education For All: Meeting*
827 *Basic Learning Needs, Jomtien, Thailand*. Paris: UNESCO.
- 828 —. 1996. *Education For All: Mid-decade Report*. Paris: UNESCO.



Monitoring and Evaluation for Practice and Policy

- 829 —. 2000. *The Dakar Framework for Action: Education for All: Meeting our Collective*
830 *Commitments*. Adopted by the World Education Forum, Dakar, Senegal, 26–28 April
831 2000. Paris: UNESCO.
- 832 —. 2005. *EFA Global Monitoring Report 2006, Literacy for Life*. Paris: UNESCO.
- 833 United Nations Statistical Office (UNSO) (Principal Authors Wagner, D. A., and A. B.
834 L. Srivastava). 1989. *Measuring Literacy through Household Surveys* (Doc. No. DP/UN/
835 INT-88-X01/10E). New York: United Nations Statistical Office.
- 836 Venezky, R. L. 1992. *Matching Literacy and Social Policy*. Technical Report.
837 Philadelphia, PA: National Center on Adult Literacy.
- 838 Wagner, Daniel A. 1990. Literacy Assessment in the Third World: An Overview and
839 Proposed Schema for Survey Use. *Comparative Education Review* 33(1): 112–138.
- 840 —. 1993. *Literacy, Culture and Development: Becoming Literate in Morocco*. New
841 York: Cambridge University Press.
- 842 —. 1998. Literacy Retention: Comparisons Across Age, Time and Culture. In: *Global*
843 *Prospects for Education: Development, Culture and Schooling*, ed. by H. Wellman and
844 Scott G. Paris, 229–251. Washington, D.C: American Psychological Association.
- 845 —. 2000. *Literacy and Adult Education*. Global Thematic Review prepared for the
846 U.N. World Education Forum, Dakar, Senegal. Paris: UNESCO.
- 847 —. 2001. Conceptual Dichotomies and the Future of Literacy Work Across Cultures.
848 In: *Literacy and Motivation: Reading Engagement in Individuals and Groups*, ed. by
849 C. Snow and L. Verhoeven. NJ: L. Erlbaum.
- 850 —. 2003. Smaller, Quicker, Cheaper: Alternative Strategies for Literacy Assessment
851 in the UN Literacy Decade. *International Journal of Educational Research* 39(3): 293–
852 309.
- 853 —. 2004a. Literacy in Time, Space: Issues, Concepts, Definitions. In: *Handbook of*
854 *Childrens Literacy*, ed. by T. Nunes and P. Bryant, 499–511. Lancaster, UK: Kluwer
855 Academic Press.
- 856 —. 2004b. Literacy(ies), Culture(s) and Development(s): The Ethnographic Challenge.
857 (A Five-Volume Review). *Reading Research Quarterly* 39(2): 234–241.
- 858 —. (ed.). 2006. *Monitoring and Evaluation of ICT in Education Projects: A Handbook*
859 *for Developing Countries*. Washington, DC: World Bank/InfoDev.
- 860 —. 2008. *Mission Report in Morocco: ALEF Literacy Campaign, Monitoring and*
861 *Evaluation*. Washington, DC: AED/USAID.
- 862 —. 2002. *Towards Guidelines for the Improvement of Literacy Assessment in Developing*
863 *Countries: Conceptual Dimensions Based on the LAP Project*. Philadelphia: ILI/
864 UNESCO.
- 865 Wagner, Daniel A., and Robert Kozma. 2005. *New Technologies for Literacy and Adult*
866 *Education: A Global Perspective*. Paris: UNESCO.
- 867 Wagner, Daniel A., R. L. Venezky, and Brian V. Street (eds.). 1999. *Literacy: An*
868 *International Handbook*. Boulder, CO: Westview Press.
- 869 Ziegahn, Linda. 1992. National Literacy Assessments: Present Practice and Alternatives.
870 *International Journal of Educational Development* 12: 223–232.

871

on my mac



Journal : **11159**

Article No: **9108**

MS Code : **REVI137**

Dispatch : **7-8-2008**

LE

CP

Pages : **22**

TYPESET

DISK

872 **The author**

873 **Daniel A. Wagner** is Professor of Education and Director of the International Literacy
874 Institute, co-founded by UNESCO and the University of Pennsylvania. He is also
875 Director of the National Center on Adult Literacy at the University of Pennsylvania. He
876 has been a visiting fellow at the International Institute of Education Planning in Paris, a
877 visiting professor at the University of Geneva, and a Fulbright Scholar at the University
878 of Paris. His most recent multi-year projects have been in India, South Africa, and
879 Morocco. Dr. Wagner has more than 150 professional publications, and currently
880 specializes in literacy, basic education, and new information technologies.

881 *Contact address:* International Literacy Institute, University of Pennsylvania,
882 3910 Chestnut Street, Philadelphia, PA 19104-3111, USA. E-mail: wagner@literacy.
883 upenn.edu.

884
885
886



Journal : **11159**

Article No: **9108**

MS Code : **REVI137**

Dispatch : **7-8-2008**

LE
 CP

Pages : **22**

TYPESET
 DISK