CONTESTED GROUND: PERFORMANCE ACCOUNTABILITY IN ADULT BASIC EDUCATION

by

Juliet Merrifield
Center for Literacy Studies
University of Tennessee

NCSALL Reports #1
July 1998

The National Center for the Study of Adult Learning and Literacy
Harvard Graduate School of Education
101 Nichols House, Appian Way
Cambridge, MA 02138

NCSALL Reports are funded by the Educational Research and Development Centers Program, Award Number R309B60002, as administered by the Office of Educational Research and Improvement/National Institute of Postsecondary Education, Libraries, and Lifelong Learning, U.S. Department of Education, through contract to Harvard University. The contents of NCSALL Reports do not necessarily represent the positions or policies of the National Institute on Postsecondary Education, Libraries and Lifelong Learning; the Office of Educational Research and Improvement; or the U.S. Department of Education, and you should not assume endorsement by the Federal Government.

TABLE OF CONTENTS

Illustrations:

EXECUTIVE SUMMARY

Adult basic education (ABE) has long been viewed by many educators and policymakers as a tool for addressing social and economic problems. Now, in a context of global economic restructuring, changes in work and employment, and the largest immigration to the USA since the early 1900s, ABE must demonstrate its success in terms of student and societal outcomes. In short, ABE is facing demands to be accountable for its performance.

A focus on results is new for adult education, and potentially transforming. Performance accountability offers the chance to re-focus on what learners accomplish by participating in ABE and to re-orient every aspect of the system to achieve the best results.

It is tempting to rush into developing measurement and reporting systems. Experiences with performance accountability show that this would be a mistake. Agreement on what to measure must be established first.

This policy paper lays out key issues in performance accountability and presents recommendations for policy and action. It is based on the literature from education, government, management, and other fields, and draws on interviews with researchers and adult education leaders at state and national levels. Its recommendations were informed by a discussion with a group of experienced adult education practitioners and researchers.

Why change? What is the problem?

Past policy is no longer a guide for action. The 1966 Adult Education Act, which has governed ABE for more than 30 years, provides little guidance in developing performance accountability. The act charges recipients of funding with enabling adults to "acquire the basic educational skills necessary for literate functioning." Definitions of literate functioning vary and are hotly debated. Over time, views of what literacy means have shifted from academic skills such as decoding text, to functional skills, such as being able to perform certain tasks using literacy skills. Recent research has changed radically our understanding of literacy. Literacy is now described as multiple "literacies" rooted in particular social contexts. This change in definition shifts the focus even further from abstract skills to real-life practices.

This conceptual change requires a re-vision of what "performance" means. When literacy meant mastery of what was taught in schools, performance was testable. When literacy shifted to the notion of functional competency – being able to perform certain tasks using literacy skills – the issue of performance became more complex. Tests had to identify which real-life literacy tasks should be included and which not, with no theory to guide which to choose or how to create scales of difficulty. Test developers had to assume that performance on the test equated with how well the student performed the real-life equivalent task. The recent research on literacy in its social context has been carried out through careful observations of literacy events and activities which shed light on prevailing literacy practices. While it shifts the focus to performance in life, not in test situations, this new research has not yet been incorporated into practice, assessment, or policy.

Agreement on what "literate functioning" means is crucial to accountability. Is literacy a right or a necessity for the good life? Should performance be demonstrated in terms of literacy skills – the earlier view – or literacy practices, the current thinking?

The purpose of literacy is not defined. The Adult Education Act of 1966 had broad social intent. In addition to not strictly defining what literacy was, it did not define the purpose of literacy. If a performance accountability system is to measure only literacy gain, the task of developing such a system would be challenging enough. If a system is to measure the achievement of literacy for a predefined purpose, a lack of clear objectives makes accountability systems even more difficult to develop. The debate in this area centers around whether the purpose of literacy education is individual advancement or community development. Are the desired outcomes productive workers, good citizens, or merely more literate people? Once these questions are answered, the next set of questions arise: who defines whether these outcomes are achieved, and how much responsibility should the literacy program be asked to take for these outcomes?

Stakeholders are not mutually accountable. Another area of concern lies in the mutual responsibility for adult basic education. Many possible stakeholders – learners and teachers, administrators and policymakers, funders, employers, public school personnel, and taxpayers – may be said to have a legitimate concern with the outcomes of adult literacy education. All stakeholders are not, however, equal in terms of access to information

or ability and power to hold the adult education system accountable. Learners, for example, often have limited information and little power to change the system. Congressional representatives stand for taxpayers in exercising accountability over the adult education system which is supported by public money. Legislators are often not held accountable by learners or educators for providing adequate resources and policy guidance to the system.

Capacity is weak. Adult basic education is struggling to create a national accountability system without a national service delivery system. It is difficult to have a management information system when there is no management system.

Research reveals a fragmented and incomplete system with multiple funding sources and reporting formats, diverse institutions, competing objectives, and missing or unreliable performance data. In most states, staff are part-time, and volunteers continue to have an essential role in student services. Per-student funding is low, and most programs are not able to meet client needs for childcare or transportation. While data on performance are fragmentary, what there are suggest that most learners do not stay long, make some initial learning gains, but may not make long-term skill gains.

The capacity to perform – to achieve desired goals – is linked with the capacity to be accountable – to document achievements and measure results. Some states are beginning to use program performance data successfully to improve program services. More reliable accountability data are collected when they are used at the program level to meet program needs.

Measurement tools are not up to the task. Learning is at the heart of ABE, and its measurement is of particular concern to performance accountability. Adult education cannot be accountable to learners or to policymakers without the ability to track learning of individuals, to demonstrate what has been learned, to compare learning across programs, and to judge learning against external standards.

Yet standardized tests, the most widely used tools for measuring learning, have been criticized both by researchers and practitioners because they do not demonstrate what has been learned. They are also incompatible with new research-based conceptions of literacy as social practices rather

than isolated skills. Some programs are using various "authentic assessment" tools, such as portfolios, but so far these cannot compare learning between learners and across programs. Without external standards or criteria, authentic assessment will not meet policy needs.

What change is happening?

New initiatives are addressing performance accountability. The problems of accountability are well recognized and are beginning to be addressed at national and state levels.

Equipped for the Future: The National Institute for Literacy's Equipped for the Future (EFF) project is a broad-based system reform effort that has actively sought input from a wide range of stakeholders. These include stakeholders outside the adult education system (policymakers, employers, and civic leaders) as well as within it (adult learners, teachers, program administrators, and researchers). EFF has developed and validated a set of four purposes for adult education and lifelong learning: to have access to information, to give voice to ideas and opinions, to solve problems, and to be able to continue learning. EFF has related these purposes to the three adult roles of worker, citizen, and family member. With 25 development partners across the country, EFF is now testing a set of common activities and defined skills which will form the basis for future development of standards and performance measures. When fully articulated and validated, these could provide a framework for performance as the basis for an accountability system.

National Outcomes Reporting System: This project is funded by the U.S. Department of Education and has involved many state ABE directors. It is developing a common set of outcomes for adult basic education as a basis for data collection and reporting. However, the process has so far had limited input from the full range of stakeholders.

Other state and national level initiatives: A number of initiatives at state and national levels are bringing stakeholder groups into closer relationships. A National Summit, proposed by the National Center for the Study of Adult Learning and Literacy (NCSALL), in partnership with NIFL and the Department of Education, is an example of one way to ensure stakeholder involvement in deliberations about the future of ABE at the national level. Local and state adult learner organizations are forming, increasing learner engagement around issues like state and federal

funding. A national adult learner organization, VALUE (Voices for Adult Literacy United for Education), has recently been formed and shows promise for building a strong and effective student voice.

State capacity-building: Over the last few years, many states have been focusing on building capacity for program delivery. Some have worked specifically on building capacity for accountability as well. The cases of Arkansas, Connecticut, and Pennsylvania, reviewed in this policy paper, as well as other states like Massachusetts, Iowa, and California, all demonstrate the importance of long-term and systematic investment in training, information, and technical assistance.

Lessons from the business world: High performance or "learning organizations" from the business world may prove useful models for adult education. These emphasize continuous improvement and learning in all parts of the organization, responsiveness to internal and external customers, participation in decisions, and shared responsibility at all levels. Traditional models of accountability are linear – quality control inspectors check widgets at the end of the production line to see if they meet specifications. Mutual accountability engages members of the organization in developing common vision, determining goals and customer expectations, and designing effective means of monitoring, producing, and improving.

What next? A Framework for Action

The policy paper recommends four principles which would enable the field to perform effectively and to be held accountable for performance. This framework for future policy builds on what has been learned about performance accountability in past experiences, and provides guidelines for future action. The principles are:

1. Agree on performance.
2. Build mutual accountability relationships.
3. Develop capacity both to perform and be accountable.
4. Create new tools to measure performance.

1. Agree on performances

Good performance – what needs to be measured – is not a technical question, but inevitably in the realm of values. The challenge is to come to

agreement on performance as a "big tent" which can include the full diversity of purposes.
Experience from business and industry suggests that it is crucial that performance be defined neither too tightly nor too loosely. If performance is defined too tightly, a mis-match between system goals and individual goals may occur. For example, learners' purposes for entering literacy education may be to read to their children, but the system may only measure whether they get a GED. If performance is framed too loosely, no shared mission or common accountability measures can be developed. Getting it right requires a broad-based and inclusive process involving multiple stakeholders.

Lessons from the literature and experience in education and other fields suggest:

- Don't assume the question of what performance means can be skipped over or rushed. Without knowing what is important, measurement becomes an exercise in "gaming the numbers" to satisfy external demands, often with perverse results.
- Involve stakeholders and seek consensus. Without broad public debate it is difficult to frame performance goals which reflect the "big tent."
- Reflect newer understandings of literacy, and connect performance with real life. This is an opportunity for literacy research to connect with and support practice.
- Acknowledge multiple performances. Too narrow or tight a definition of goals will exclude learners and programs or force them to falsify their data.

2. Develop mutual accountability relationships

Reforming accountability in high performance terms requires a switch from one-way, top-down lines of accountability to a mutual web of accountability relationships. To participate, stakeholders need information and the ability or power to hold others accountable. This entails greater transparency in information, increased flows of information, and room at the table for groups who have not been there in the past.

- Bring the full range of stakeholder groups into the process – including teachers and learners who often have not been at the table.

- Provide support for stakeholders who have least access to information and power, including adult learner organizations at national and state levels.
- Increase information flows among and between all stakeholders.
- Develop learning organizations at the program and state levels which would emphasize learning and continuous improvement, shared responsibility, and engagement in monitoring results.

3. Build capacity to perform and be accountable

State experiences with capacity-building reviewed for this policy paper indicate that the two kinds of capacity – to perform and to be accountable – are linked. By developing a learning organization approach in which there are continuous feedback loops, performance data can help programs improve performance.

- Build the capacity to perform. Key elements include increased resources, focusing resources on quality rather than quantity, staff development and training, technical support, use of performance data for continuous improvement.
- Build the capacity to be accountable. Key elements include accountability demands which are commensurate with resources and capacity, engagement of users in developing better measurement tools, staff training and support, timely information loops, rewards for improved performance.

4. Create new tools to measure performance

Accountability systems must meet the different information needs of different stakeholders. To do so, data users and data providers must be in communication so that the most appropriate measurement tools can be applied.

Accountability systems commonly use several types of indicators to track performance over time. Input indicators provide information about the capacity of the system and its programs. Process indicators track participation in programs to see whether different educational approaches produce different results. Output indicators are short-term measures of results, and outcome indicators are long-term measures of outcomes and impacts. No single indicator can suffice to measure performance, especially of an enterprise as complex as adult basic education.

ABE invested a great deal of work into developing indicators of program quality, but much less on performance measures. New approaches and tools for measurement are needed which are linked to performance.

- Develop external standards or criteria against which individual student learning can be measured and through which program performance can be assessed.
- Develop performance assessment tools for measuring learning. These directly assess learners' performance in terms of literacy practices rather than the indirect approaches of standardized tests which "stand for" real-life practices, usually inadequately. Initiatives in performance assessment in countries such as Britain and Australia may provide useful models for measuring and assessing learning.
- Use the full range of potential of research, evaluation, and monitoring technologies to meet the needs of different stakeholders. These approaches to gathering, analyzing, and using information are based on different kinds of data and meet different purposes. Using them in appropriate ways, adult education can develop a dynamic system of information, analysis, and reporting.

Next steps

This framework of principles for action acknowledges that there are no quick answers. To put the principles into action requires consultation with the field and with stakeholders. It will need meetings and taskforces, and it will take time. It requires learning lessons from elsewhere when appropriate, building on current initiatives when they are under way, and creating new tools when none exist. Policymakers have the capacity to set the stage, harness resources, and create a common agenda. Commitment to high performance requires the contributions of many players.

ACKNOWLEDGMENTS

What we bring to a project shapes how we see it and what conclusions we reach. There is no objectivity: all we can do is be as clear about our own biasing factors as possible, so that the reader can take these into account. I bring particular experiences and perspectives to this policy paper. My academic background is as an anthropologist and political scientist. Of greater significance is ten years on the staff of the Highlander Research and Education Center in Tennessee, a nonformal education center with a long history of work for social justice in the southern United States. That work highlighted the importance of accessing information for community action, the ways in which people with little formal education often feel excluded from the democratic process, and the power of action not only for change but also for learning.

This grounding in participatory approaches to research and action shaped my approach as Director of the Center for Literacy Studies (CLS) at The University of Tennessee. My work included action research and more conventional research, along with participatory projects in staff and curriculum development. I was also involved in a performance accountability design and planning project in Tennessee, funded by the National Institute for Literacy. As a state-wide, inter-agency project trying to create a common framework at state and local levels, the project had many challenges. With the power of hindsight, now I would have a much better sense of how to begin.

Many people helped to shape the ideas in this policy paper. Foremost are two former colleagues at CLS: Mary Beth Bingman, the project director for NCSALL work, and Brenda Bell, project director for the CLS work on Equipped for the Future. Olga Ebert provided valuable research assistance which helped overcome the difficulties of writing at a distance. CLS staff members Wil Hawk and Margaret Lindop proofread the manuscript. Mary Ziegler, CLS director and Donal Crosse, CLS staff member, took time to share insights about their Tennessee PMRIS work. NCSALL Director, John Comings, provided guidance and comments. Barbara Garner at World Education contributed to the Executive Summary. Hal Beder, who is preparing a parallel policy paper for NCSALL on outcomes research, added insights along the way.

Several participants in the workshop that was part of the preparation for this policy paper read and commented on drafts: Peggy McGuire, Gail Weinstein, David Rosen, Marilyn Gillespie, Sondra Stein, and Larry Condelli. Members of the Advisory Committee for NCSALL's assessment work provided further comments: Cheryl Keenan, Andres Muro, Carol Kasworm, Loren McGrail, and Glynda Hull. Roberta Pawloski, Garland Hankins, and Patsy Medina also commented on a draft. Heide Spruck Wrigley and Allan Quigley provided a final peer review and critique.

One of the many gaps in my knowledge is of language acquisition theory and ESOL practice. Loren McGrail and Marilyn Gillespie valiantly tried to plug that gap, but the failing remains. This paper is much less grounded than I would like in the ESOL research literature, and more oriented to English literacy and basic skills. The insights from all of the above were gratefully received. The errors that remain are mine alone.

INTRODUCTION

This paper hopes to contribute to the public debate that must answer fundamental questions about accountability in adult education. To whom are adult educators accountable? For what? How shall we demonstrate performance?

To be accountable is to be responsible to others for what we do. In recent years responsibility for achieving results in adult education has been an outgrowth of politicians (and taxpayers) demanding to see their "return on investment." The focus has shifted from process (What does a good quality program look like?) to outcomes (What are the results of participation in programs?). Outcomes can no longer be taken for granted, but must be measured and proved. But just what these outcomes should be is not always clear or agreed, either within the field or outside it.

State and national efforts have begun to construct systems to report on the results of adult basic education.[1] Some are designed to provide data for program improvement, some primarily to report to funders. These efforts include the National Institute for Literacy's Performance Measurement and Reporting Improvement System (PMRIS) projects, the National Governors Association's workforce development projects, the U.S. Department of Education-led program quality indicators initiative, and the National Outcomes Reporting System. Volunteer literacy organizations and some other states have their own performance accountability initiatives.

To measure performance we have to know what success would look like, what the desired outcomes are – in fact, what it means to be literate. Currently, it seems that policymakers, teachers, learners, and the general public hold many and varied ideas about what the outcomes of adult basic education should be. Is the goal for individuals to acquire increased literacy skills, greater self confidence, to participate more in their community, make more effective parents, acquire better jobs? Is the

[1] A note about terms here: adult literacy education and adult basic education are used interchangeably. They include basic literacy skills (reading and writing) and language skills (speaking and listening) as well as numeracy. The legislation which governs a great deal of this in the United States is the Adult Education Act, and for the sake of brevity the term "adult education" is sometimes used here in this sense. Nevertheless, the field of adult education as a whole is broader, addressing all the myriad learning experiences and opportunities encountered in adult life.

benefit for society having higher skilled workers, a more active citizenry, stronger families? Does literacy have value in itself, or only because of its social impacts? Is literacy as firmly associated with social impacts as popularly believed? There are also different understandings of literacy itself: as an autonomous skill or set of skills; as the ability to function in society; as diverse socio-cultural practices rooted in particular contexts. Without clarity and agreement on these fundamental questions, it is difficult to see how performance accountability systems can work.

Getting clear about what we mean by "performance" is not just a technical question, but a question of value. At the heart of performance accountability is what we want adult basic education to be and to become. Only when we are clear about where we want to go can we create mechanisms to show how far we have come. The process of developing that clarity and agreement should start by engaging the field, and must reach beyond it to the broader public.

This policy paper seeks to make a contribution to that debate by examining underlying assumptions, unpacking terms, and identifying ways forward in developing performance accountability in adult basic education. It is written not just for people who think of themselves as policymakers, but for a wider audience who are active and committed to adult basic education, and who want to play a role in shaping its future. Preparation for the policy paper included three main activities:

1. A review of literature from education and other fields. This was a means of identifying major issues to be addressed in the policy paper. It drew on materials developed by the federal government including the U.S. Department of Education, the National Governors Association, the National Institute for Literacy, and several states. It ranged further into recent research and theory about literacy, particularly New Literacy Studies, organizational management, and in particular, performance indicators and measurement in both the national and international arenas.
2. A series of interviews with adult education researchers and leaders at state and national levels about their performance accountability efforts. These were conducted in March and April, 1997, in person and by phone. The 17 individuals included current or former state directors of adult education, people active in national policy, leaders of national volunteer literacy organizations, researchers with extensive experience in the field, and others experienced in state-level performance

management systems. The interviews were not intended to be representative of the field as a whole, but rather to draw on a range of experiences with performance accountability in order to identify major issues and questions.

3. A discussion workshop to reflect on the emerging issues with a group of adult education practitioners and researchers. Participants were selected for their perspectives at national, state and program level. They were not necessarily experienced in performance accountability, but were very experienced in the field – in workplace literacy, family literacy, English for Speakers of Other Languages (ESOL), community-based and institution-based programs.

The paper starts with the context, in terms of societal, governmental, and educational changes. In Chapter Two, some of the key terms are examined: accountability, performance accountability, and performance – with a look at conceptions of literacy itself, and the purpose of literacy education. In Chapter Three, the issue of capacity – both to perform and to be accountable – is discussed, along with specific issues around measurement. Finally, in Chapter Four, a scenario is sketched out for ways forward in terms of debate, action, and additional research.

ONE. FROM CAMPAIGN TO SYSTEM: THE HISTORICAL CONTEXT

The language of adult education has changed from the 1960s, the start of significant governmental support for adult basic education, to the present. Credentials, program quality, reporting, and accountability have assumed higher priority. In broad brush, we can characterize the changes as a shift from the "campaign" mode of operation of the 1960s and 1970s toward the "system" mode of thinking that dominates today's discussions (if not the reality of practice).[2]

Campaigns have some common characteristics: a sense of urgency and crisis; a focus on short-term results; a willingness to push funding to its limits by incorporating volunteers; and a concentration on action rather than accountability – on recruitment and instruction rather than retention and results. Systems, in contrast, focus more on the long haul. They are characterized by professional staffs, an institutional base, funding for capacity-building (teacher training, support and resources), a focus on quality, and the development of accountability mechanisms to measure effectiveness.

Claiming such a shift from campaign to system mode is perhaps simplistic. Some system building activity certainly took place as early as the 1970s, especially in certain states. Even today the system is far from complete: many would argue that we do not yet have a "system" in adult basic education, more a patchwork of different systems, agencies, funding streams and philosophies. Volunteers, part-time staff, and low funding continue.[3] Nevertheless, there have been dramatic changes in emphasis, which have brought the accountability debate to center stage.

None of these changes took place in a vacuum. The seeds of the demands for accountability can be found in changes going on in society during this period. Dividing the history into these threads is somewhat arbitrary, since there have been many cross-cutting influences. However, these changes can be traced in four broad areas:

---

[2] I began thinking about this shift as a result of conversations with Dr. Mary Hamilton of Lancaster University, England, about parallel developments in Britain over the last 25 years. Bob Caswell, President and CEO of Laubach Literacy International has for years talked about different approaches to literacy, from campaigns to institutional systems.

[3] 74% of providers in the NEAEP study reported using volunteers; nearly 60% of programs had no full-time staff, and over 80% of adult education instructors are part-time. Estimates of per-student expenditures range from $160 to $258 (Moore and Stavrianos, 1995).

- societal changes, both economic and social;
- governmental changes, including reinventing government, the increasing adoption of business language and expectations, and a steady stream of legislative changes;
- education system changes, in particular concerns about children's "success" in school (fueled by global economic changes) manifested in "back to basics'" movements, Goals 2000, development of national standards, K-12 reforms;
- changes in conceptions about literacy and the purpose of literacy education.

Societal changes – economic and social

Global economic restructuring has characterized the period from the early 1970s to the late 1990s. Multinational corporations developed the capacity to switch production operations from one country to another and have moved many manufacturing activities from high-wage industrial nations to low-wage regions of the world.[4]

By the early 1980s, fears about jobs moving overseas, and the U.S. becoming a second-rate economic power, found expression in concern about the skills of the workforce. Reports like *Workforce 2000* (Johnston & Packer, 1987) argued that the jobs that would remain in the U.S. require a level of literacy and problem solving ability which had never before been demanded of the blue collar workforce. It made a clear case that adult literacy was not an issue that was going to go away, but one that required ongoing, serious attention if the U.S. was to be ready for global competition. *Workforce 2000* was followed by a series of other reports on a similar theme, including *America's Choice: High Skills or Low Wages?* (National Center on Education and the Economy, 1990) and the Secretary's Commission on Achieving Necessary Skills (SCANS) reports (1991 and 1992).

As concerns about the skills of the workforce grew, preparation for employment became ever more explicitly the primary purpose of education. Voices advocating the broad view of education for citizenship lost ground to a sharper vocational focus in both adult and K-12 education. Paul Miller, in an overview of adult education, said, "This glance across

[4] Foreign Direct Investment (FDI) flows have doubled as a share of global output since the 1970s, and in 1996 reached a record 5% of the world's GDP (UN, 1997).

two centuries suggests the long and inexorable shift that education has made, encouraged and demanded by constituents and funders alike, from education for citizenship toward preparing people for occupational success" (Miller, 1995, p. 46).

The customers of adult education began to be defined as employers, interested in the "product" of skilled employees. The Business Council for Effective Literacy was started in 1983 "to encourage business and industry to join in the fight against adult illiteracy" (Harman, 1985, p. i), with a particular focus on providing information and resources for businesses wishing to upgrade their workforce skills. Japanese management practices gained adherents in the U.S., and firms began to demand that education pay similar attention to quality control, results, and customer demands.

The new orthodoxies about the changing nature of work were summarized in the SCANS reports, which described the characteristics of "high-performance workplaces:"

- Flexible and decentralized production techniques;
- Employee empowerment, by giving employees decision-making responsibility, career paths, and wage progression tied to skills;
- A strong emphasis on "excellence," on continuously improving work performance, and on the kind of management for quality that reduces error and rework, increases customer satisfaction, and cuts costs;
- Continual training to upgrade skills and employees' ability to function effectively in a problem-oriented environment; and
- Increasing integration of tasks through work teams and the identification of workers with their products and services (SCANS, 1992, p. 5).

At the same time that work was becoming more explicitly the primary purpose of adult education, social changes resulting from economic restructuring were being set in motion. As factories closed and moved overseas, long-term unemployment grew, poverty both in inner-city and rural areas increased, as did the temporary "contingent" workforce, with little job stability or security. The gaps between rich and poor widened, crime rates increased. The search for something to blame for these social trends fastened on individuals rather than institutions. "Illiteracy" became for some people a prime cause of social ills, and the purpose for literacy education was to relieve social problems like unemployment, crime, drug abuse (Bush, 1989).

Competing purposes for adult education stemmed from another significant social development: new immigration to the United States, especially from Southeast Asia and Central America (a stream of refugees from Central America and Southeast Asia began to arrive in the U.S. in the late 1970s). The new immigrants' need and demand to learn English fueled the development of ESOL programs, especially in states with high immigration rates, like California, Florida, Texas and New York.[5] Increasingly ESOL programs also had to work with people whose native language literacy was limited. The need to incorporate immigrants and make them legal citizens led to 1988 federal funding for language education, SLIAG (State Legalization Impact Assistance) Grants. More liberal immigration legislation (the Immigration and Nationality Act of 1990) was designed to legalize existing immigrants and discourage new illegal immigration (Gillespie, 1996).

The largest influx of immigrants since the early 1900s continues, generating waiting lists for ESOL classes. Immigrants tend to concentrate in certain states, especially the northeast, southwest and west, and in certain cities – Miami's adult population is now 53% Limited English Proficient (Gillespie, 1996, p. 15). However, immigration is spreading even to small towns and rural areas. Immigration patterns will have significant impact on employment by the turn of the century, when the U.S. Department of Education estimates that immigrants will constitute 29% of new entrants to the workforce (Gillespie, 1996).

These twin themes, of economic restructuring and social restructuring, underpin the other historical trends. The demand to account for results is linked with these economic changes, the growing view of adult education as a means of workforce development, and the acceptance of business as a prime customer for education. The questions from policymakers are: does participating in adult education help people get a job, get off welfare, reduce crime?

Governmental changes

"Reinventing government" initiatives, spurred on by Osborne and Gaebler's book of that name (1993), can be seen partly as a response to the increasing influence of business in government. Business language has been widely adopted in government ("customer," "results," "efficiency,"

[5] A good overview of the research and issues relating to immigration and vocational and workplace ESL is in Gillespie, 1996.

"return on investment"). The Government Performance and Results Act of 1993 requires federal agencies to define clearly their missions and to establish long-term strategic goals, as well as annual goals. Similar accountability language became apparent in 1996 adult education bills (not passed).

State governments have been leaders in performance accountability. States like Oregon and Minnesota began in the 1980s to define long-term objectives to be achieved by the state and measures of progress towards them (see review in National Institute for Literacy, 1995a, and National Governors Association, n.d.). Oregon's early initiative engaged citizens and organizations around the state in consultation about goals and benchmarks. Learning from these efforts has informed other national and state activities in performance reporting.

At the same time that government was being reinvented, the ballooning budget deficit meant that government initiatives launched with great fanfare found themselves with very limited financial resources. That was certainly true of the 1991 National Literacy Act, which promised major changes in the system of resources for the field, through establishment of the National Institute for Literacy (NIFL), state literacy resource centers, funding set-asides for professional development, and demonstration programs. But appropriations were far less than had been authorized.

Taxpayers' revolts, starting with Proposition 13 in California in 1978 and carried to Washington with Ronald Reagan, were fueled by the increasing budget deficits, and contributed to moves to downsize government, and also to end the social support system created in the Great Society legislation of the 1960s.

The 1966 legislation governing adult education was amended both in 1988 and 1991, and ongoing attempts are underway to replace it with a new legislative mandate. However, other legislation became influential over adult education during the late 1980s and 1990s, especially employment training (the Job Training Partnership Act, JTPA, passed in 1982, revised in 1989), the Family Support Act (1988) which initiated the JOBS program of welfare reform, and the new welfare act of 1996. These other initiatives, particularly welfare reform, have shaped the field as much as its own legislative brief, and de facto have established new purposes and rationales for adult literacy education, dominated by preparing low-income adults for work.

Education system changes

Sea changes in economy and society have also shaped education. *A Nation at Risk* (National Commission on Excellence in Education, 1983) prompted concerns about children's success at school, as defined by employers' needs for skilled workers. As the National Education Association's president later suggested, education came to be a weapon in the global competition for economic power:

> The mission of education was to serve the national interest. The destiny of American democracy, it was argued, demanded what revitalized education alone could deliver: technological might in the service of military security, a rejuvenated economy in the service of reclaimed dominance in the international marketplace, and the social and political integration of waves of new immigrants in the service of national harmony. (Futrell, 1989, p. 11)

The "back to basics" movement in K-12 education placed reading, writing, and arithmetic at the center of the educational agenda, and highlighted teaching methods, testing, graduation standards, and technology. The initial wave of educational reform came not from the schoolhouse but from the statehouse: between 1983 and 1985, more than 700 statutes were passed by state legislatures prescribing who should teach what to whom (Futrell, 1989). Throughout successive reform waves, the underlying premise was unquestioned: that education's purpose is primarily economic.

The governors developed national education goals in 1989 – *Goals 2000* (National Education Goals Panel, 1993). A series of annual reports on progress toward these goals were released, despite the fact that many, especially the adult education goal, did not have measures or available data. The goals were mostly so broad that they did not readily lend themselves to measurement, let alone achievement by the year 2000. Slowly, measures are being developed, and national standards for core curriculum areas – math, English, history, science – are intended to specify targets to be achieved. These will need performance indicators to be useful for monitoring and evaluation. Standard-setting in adult education was slower to develop, but in 1993 the National Institute for Literacy initiated the Equipped for the Future project, whose purpose is to develop standards for adult literacy education (Stein, 1995 and 1997).

Anxieties about children's performance in school spilled over into concerns about their parents' abilities, and provided fertile ground for new

developments in "family literacy." In the late 1970s, Tom Sticht (1983) noted that "inter-generational transfer" of literacy meant that adult literacy education's work with parents generated double rewards by also impacting children's success in schools. Other family literacy pioneers, like Denny Taylor (1983) and Elsa Auerbach (1989), recognized the family as an important domain for literacy practices, and developed approaches to working with both parents and children.

By 1985, Sharon Darling had begun the PACE family literacy model in Kentucky, targeting parents and pre-school children, a model which was later enshrined in federal Even Start legislation in 1988. Family literacy as defined by the PACE model became a significant player in adult basic education around the country, with "family" conceived as a parent with pre-school children (although other approaches exist, especially in bilingual family education, and in programs working with grandparents and with older children). The primary impetus for family literacy programs is the success of children in school, with parents in the role of "first teachers," rather than parents' literacy development for its own sake. As such it demonstrates adult education's continuing low priority in the larger education field.

Adult education system changes

Two inter-twined narrative threads have shaped accountability thinking within the field of adult basic education over the last 20 years: the concept of literacy and the purpose of adult literacy education.[6]

Concept of literacy: At the core of accountability lies the concept of what literacy means. That concept changes over time, and there is now a substantial literature on historical changes in the concept of literacy (e.g. Cook, 1977; Stedman & Kaestle, 1987). Early definitions of literacy as reading and writing simple text shifted slowly to seeing literacy as related to "functioning" in society (Stedman & Kaestle, 1987). The 1966 *Adult Education Act* echoes this in its aim of enabling adults "to acquire the basic educational skills necessary for literate functioning." That begs the question of what "literate functioning" means, and how it can be measured.

[6] For a more complete overview of developments in adult literacy practice in recent years, see Fingeret's reviews for ERIC (1984 and 1992), and for a research overview see Beder, 1991 and Sticht, 1988.

The need to count numbers of "illiterates" has been imperative for policy-makers, but fraught with difficulties, in part because it has been a moving target, in part because of the difficulties in defining literacy. Attempts following World War II to gather systematic statistics on literacy worldwide ran into the problem that what literate functioning means is relative, varying from country to country and over time. By 1962, UNESCO acknowledged this relativity by proposing that: "A person is functionally literate when he has acquired the knowledge and skills in reading and writing which enable him to engage effectively in all those activities in which literacy is normally assumed in his cultural group" (as cited in Hunter & Harman, 1979, p. 14).

Two ways to conceptualize the skills needed for literate functioning are as academic, school-based skills, and as functional skills related to the context of daily life outside the schoolroom. While the view of literacy as related to the context of life generally prevails in discussions about literacy, academic school-based skills have commonly shaped the content of literacy education[7] (reflected in the continued use of grade-level tags derived from K-12 education, standardized tests like Test of Adult Basic Education (TABE) and Adult Basic Learning Exam (ABLE), and indeed the GED, all derived from expectations of what children in school can accomplish).

In the last 20 years there have been ongoing U.S. and international efforts to create generalizable measures of functional literacy. The Adult Performance Level (APL) study in the 1970s was one of the earliest such attempts: researchers at the University of Texas "identified 65 specific objectives for adult functioning, identified texts that represented them, and established three levels within each, representing different levels of complexity and competence" (Harman, 1985, p. 5). The APL was heavily criticized for its white, educated, middle class and male interpretation of what constitutes literacy competency (e.g. Griffith & Cervero, 1977). Nevertheless, the spread of the statewide competency-based assessment system, CASAS (Comprehensive Adult Student Assessment System), from California, the state with the largest adult education enrollment, to at least six other states signifies the ongoing demand for generalizable measures of functional literacy.

[7] Functional-context approaches to literacy teaching have flourished particularly in work-related settings (see, e.g. Sticht, 1975).

Perhaps in reaction to the APL attempt to generalize, Hunter and Harman defined literacy in their influential review of the field as:

> the possession of skills perceived as necessary by particular persons and groups to fulfill their own self-determined objectives as family and community members, citizens, consumers, job-holders, and members of social, religious, or other associations of their choosing. (Hunter & Harman, 1979, p. 7)

In this view, the judgment of literacy resides with the individual – not a perspective conducive to measurement and reporting of literacy levels or achievements. The large body of research in New Literacy Studies has further clarified the wide variation in literacy practices in specific social and cultural contexts (see, for example, Barton, 1994a; Barton & Hamilton, 1998; Gee, 1990; Heath, 1983; Lankshear, 1997; Merrifield et al., 1997; Street, 1984 and 1995). Nevertheless, policymakers' need to measure literate functioning continues, most recently in the National Survey of Adult Literacy (NALS). The increasing demands for performance accountability necessitate also measuring the results of adult literacy programs in terms of literacy gains. But competing concepts of literacy – as skills or competencies and as social practices – are at the heart of what performance means (discussed further in Chapter Two) and complicate the issue of measurement (see Chapter Three).

Purpose of education: Public policy at both national and international levels has adopted what Street calls the "autonomous" view of literacy, as a discrete and fixed set of skills, transferable from one context to another, from which economic and social development automatically follow (or at least upon which they are contingent) (Street, 1984). The social impacts of literacy appear to be the guiding purpose for public investment in literacy education.

Over the 30 years from 1966 to 1996, these impacts have steadily narrowed from the apparently broad social intents of the Adult Education Act to a prime focus on employability. This is reflected in the series of reports discussed earlier, stressing the need for workforce skill enhancement, in legislation impacting the field (JTPA, welfare reform), and in the growth of family literacy (targeting parents as teachers of children who are to be the new workforce). New programs have been tailored specifically to workforce skill demands, like the workplace literacy partnership programs established in the 1988 amendments to the Adult Education Act. Although the Secretary's Commission on Achieving

Necessary Skills (SCANS) report (1991) broadened work-related skills to include thinking skills and personal qualities, it reinforced the dominance of employment as the main focus for adult education. Even in "general" adult basic education, the skills needed for work have come to dominate.

While it can be argued that mainstream adult literacy education increasingly starts from employers' definitions of what should be taught, what Fingeret calls "community-oriented programs" start from a different sense of purpose. They start from community needs and issues, and "are more likely to be advocates of social change, facilitating efforts of individuals to address broad community concerns and teaching literacy skills as necessary to assist the larger process of change" (Fingeret, 1984, p. 21). These kinds of participatory literacy programs focus on the wants of literacy learners, and the uses to which they wish to put their developing ease with literacy (see Fingeret & Jurmo, 1989). Students are seen as partners in curriculum development and instruction, and literacy education as a tool not just for advancement but for personal and social transformation (Freire & Macedo, 1987).

Such "ideological" models of literacy, in Street's term, acknowledge that issues of power and access are inherent in literacy practice. Programs may challenge students to "read the world" as well as reading the word: to become more critically aware of their reality not just to learn how to play by the rules, but to question and change the rules.

> Participation in the political process entails not only sufficient functional literacy to operate effectively within existing social and economic systems, but also the ability to make "second order" rational and informed judgments concerning the desirability of social rule systems themselves. "Functional literacy" has, therefore, to embrace not merely knowledge of rules and the ability to follow rules, but also the capacity to think, reason, and judge beyond existing social rules. (de Castell et al., 1986, p. 11)

Given the social and economic changes, it is not surprising that the purpose of adult education has become contested ground in the 1990s. The contrasting views of what literacy education is all about makes accountability systems difficult to design. The General Accounting Office's recent report on ABE remarks, "The State Grant Program lacks clearly defined objectives, the types of skills and knowledge adults need to be considered literate are not clear and, thus, states do not have sufficient direction for measuring results" (GAO, 1995, p. 4). The issues around the

purposes for literacy education are addressed further in Chapters Two and Three, in the context of discussion about performance and capacity.

System development

These social, economic and educational changes provided the impetus for system development in adult education. Some states (like New York, New Jersey, and others) began system-building efforts in adult basic education in the 1970s, although many of these earlier efforts were not sustained. The 1980s saw restricted federal funding for adult literacy education, but by the late 1980s, system building efforts were underway again. Amendments to the Adult Education Act in 1988 added requirements for program evaluation and reporting, as well as staff training. States like Massachusetts, California, Connecticut and others were beginning consciously to strengthen their system for adult basic education. Massachusetts, for example, built a support system to meet learner goals, including staff development and training, capacity building, and increased funding.

Jump Start (Chisman, 1989) was probably the most influential publication of this period, promoting ideas about system development which helped shape the National Literacy Act of 1991. That Act established the National Institute for Literacy (NIFL), along with state literacy resource centers to develop coordination, training, and technical assistance. Other system development efforts included a requirement for states to develop program quality indicators.

However, the distance yet to be traveled in forming an effective system for Adult Basic Education (ABE) was indicated by the report of the national evaluation of ABE (Young et al., 1995) and the GAO's report on measuring program results (GAO, 1995). The latter highlighted the lack of clear goals and objectives, and the difficulties in documenting results. The former struggled with the limited capacity of local programs to collect valid data.

The response has been a rush to measurement, to defining indicators and collecting data. Yet the history of adult education argues that simple solutions have never been very effective, and that the time-honored habit of responding to crisis has been doomed to failure. Rather than instant answers, the moment calls for careful analysis. Instead of quick fixes, concerted attention and sustained, inclusive thinking is needed. More than technical solutions, public debate should address some of the

fundamental questions of accountability and of performance. In the next chapter these two terms – accountability and performance – will be explored.

TWO. UNPACKING TERMS: ACCOUNTABILITY AND PERFORMANCE

Designing performance accountability systems requires answers to some basic questions. What is the purpose of adult education? What is the nature of literacy? What are the relationships between learners, teachers, program administrators, and funders? There is a strong impulse to leap to questions like "what shall we measure?" and "how can we measure this?" But accountability systems are not just technical issues of measurement and testing: they are about what is important to us, what we value, what we aspire to.

Answering questions is made more difficult when the same words are used in different ways, and when underlying assumptions are not explicit. This chapter examines the terms "accountability," "performance accountability," and "performance" and what these terms mean for adult education.

Accountability

In everyday life, accountability means responsibility, being answerable or liable to someone else for one's actions. We cannot use the term without specifying *to whom* and *for what*. Sometimes adult educators feel accountable to learners, sometimes to funders. Accountability looks very different from different places in the system. One's position in the system, particular context and experience, resources and support, all shape to whom one feels accountable and for what. We cannot assume common perspectives on these questions; they have to be negotiated. Boxes 2.1 - 2.3 describe a few of the different perspectives on accountability, even within the particular context of local programs.

Box 2.1. An ESOL teacher:
working part-time, teaches some hours at one community college, a few more hours at a different agency, trying to piece together an income to make ends meet. This teacher often feels like a migrant worker,[8] but she cares strongly about her students and wants to make a difference in their lives. Her relationship with the educational institutions that employ her is marginal – she is only there to teach a class, and her car serves as a mobile resource center, carrying textbooks and teaching materials from place to place. There is little connection with the institution other than a paycheck – little or no staff development, no communication about vision and purpose, no sense of a bigger picture. Reporting accurate numbers up the line to program director, state director and federal Department of Education is far from a priority.

Box 2.2. A teacher in a community-based program:
working part-time, does not make much money, but has a strong identification not only with the students but also with the program.[9] Staff and teachers feel a sense of accountability to students, and have a common vision about the purpose of what they are doing. The program manager, wearing several hats, reminds the other staff about the demands of external players – funders and others – and the need to report the program's successes in terms they can understand. But it may not be easy for such program staff to translate what they see as transformations in student lives into cold data about "results."

Box 2.3. A data clerk:
works in a large, urban ABE program which is part of the local school system. All around the city, teachers enroll new students in their GED classes – at a school, a library, a community center, a church basement. The teachers are part-time, most teach school during the day, and they spend little time at central office. Once a month they are required to turn in their intake forms on new students they have enrolled. The data clerk then enters them into a computer database. The software has built-in features designed to make the data more complete and accurate. Several data items, which are required for reporting purposes (e.g. age of student, employment history), cannot be left blank. But despite constant urging, the teachers keep turning in intake forms with incomplete data. After numerous attempts to contact the teachers who are seldom home during office hours, the data clerk learns the way to an easier life is to make up the missing data. Since the data has no purpose, other than contributing to annual reports which leave the office and are never seen again, what does it matter?

[8] Thanks to Gail Weinstein for this insight, based on her experience with ESOL teachers.
[9] Thanks to Peggy McGuire for this insight, based on her experience with community-based literacy.

A range of possible answers to the question of "accountability to whom" emerged in interviews with adult education leaders, researchers and policy analysts:

- Funders (including public and private funding sources);
- Taxpayers (although these do not have a voice, except through elected officials);
- Employers and business/industry, seeking employable workers;
- Other state agencies (e.g. Workforce Development Boards, employment training agencies, Social Service agencies, etc.), some of whom are also funders of adult basic education, and all of whom place demands on adult basic education to produce a "product" of adults ready for employment training or for work;
- Students in adult education;
- Practitioners, including teachers and local program directors.

The term *stakeholder* is often used to indicate those who have a "stake" or a legitimate concern in an enterprise. Taxpayers, employers, and other agencies are commonly regarded as adult education stakeholders. Adult learners are often not included, although they have the most obvious stake in adult basic education. Describing learners as *participants* or *customers* while using *stakeholder* to refer to those outside the system who have an interest, excludes adult learners from lines of accountability. The entire range of players, including learners and practitioners as well as those outside the learning enterprise, have a legitimate concern with the outcomes of adult literacy education. They are all part of an accountability system.

There has been little discussion in the field about how different groups of stakeholders may have different interests and information needs. Accountability to learners, for example, would involve very different processes from accountability to funders. Sociologist James Coleman argues there are two approaches to accountability:

> One way is from the top down, which is a bureaucratic mode of authority. The other way is from the bottom up – for there to be accountability to parents and children. I think everything that we've seen suggests that the second is a more effective mode of accountability than the first. (cited in Osborne & Gaebler, 1993, p. 181)

A top down model of accountability would start from Congress, or a state legislature, as funder of services to adult learners. In this model, funds flow downward, reporting of results flows upward. Top-down

accountability is not only undemocratic, but ineffective, argue Osborne and Gaebler:

> Public agencies get most of their funding from legislatures, city councils, and elected boards. And most of their "customers" are captive: short of moving, they have few alternatives to the services their governments provide. So managers in the public sector learn to ignore them. The customers public managers aim to please are the executive and the legislature – because that's where they get their funding from. Elected officials, in turn, are driven by their constituents – in most cases, by organized interest groups. So while businesses strive to please customers, government agencies strive to please interest groups. (Osborne & Gaebler, 1993, p. 167)

Bottom-up accountability makes adult education programs accountable to the people they serve. This kind of accountability would involve learners themselves in decisions about learning objectives, and judgments about learning achievements. A purely bottom-up accountability model also has limitations. Accountability to learners has often been an excuse for "anything goes," since each learner is seen as unique, with different learning goals. The diversity of students has led to the argument that there can be no overarching mission or goals, that measuring results is impossible. So although there may be accountability to individuals, system-level accountability is more elusive. For performance accountability, responsibility to students for learning achievement needs to be combined with clarity and rigor in establishing goals and outcomes, documenting achievements, and reporting.

Performance accountability

Performance accountability has to do with demonstrating results. Brizius and Campbell, working with a broader focus on government accountability, say:

> Performance accountability is a means of judging policies and programs by measuring their outcomes or results against agreed-upon standards. A performance accountability system provides the framework for measuring outcomes – not merely processes or workloads – and organizes the information so that it can be used effectively by political leaders, policymakers, and program managers. (Brizius & Campbell, 1991, p. 5)

It contrasts with other approaches to accountability, like higher education, where accountability is through accreditation of institutions. What those

results are, and what it means to report them adequately, are open to different interpretations. The meaning of the term performance accountability is not consistently interpreted by adult educators. In the interviews, many state directors saw performance accountability as synonymous with indicators and measures of program quality. Program quality indicators, though, focus primarily on process, and little on outcomes, so they are not the same as performance accountability indicators.

> When you say performance accountability I think of two things, though one is really a subset of the other. The first is the indicators of program quality, as a guiding force in promoting program accountability, and the second is specific performance accountability for student outcomes. That is really a subset of educational gains [in the program quality indicators] but is in such need of improvement that we need to address it separately. (State ABE Director interview)

Almost everyone interviewed saw the primary drivers behind performance accountability as legislators, at state and federal levels, demanding to know what difference adult education makes in society:

> Ninety percent of the impetus goes back to Congress wanting to be more outcome focused and driven – to see demonstrable change in the community in specific areas. It scared the pants off adult education. It was even money for a while that there might not be an adult education, because we didn't have data that showed we could make a difference – and there was some data that showed we didn't make much difference. (Volunteer literacy leader interview)

It is widely held among adult educators that these legislators are not interested in learning gains or literacy for its own sake, but only interested in social impacts, on the economy, families, the community. "We need to give the legislature something that makes sense in economic terms. Soft ideas are not cutting it, they won't fund it" (State ABE Director).

State directors expect that better reporting of outcomes would strengthen their position in jockeying for continued or increased funding for the field:

> Politically it [performance accountability] is important to achieve more money, federal and state, for our program. We have to show

the taxpayer that they're getting a good bang for their buck. Education is not being looked at favorably. We have to show we

> are accountable for the students, or Congress will not fund us. (State ABE Director)

Among the different purposes of performance accountability discussed in the interviews with leaders, policymakers, and researchers are:

a) Report to funders on "return on investment" – mentioned by most people, and clearly the primary purpose on the minds of state directors right now;
b) Tool to sort out "good" programs from "bad", and to defund "bad" programs – particularly important for state directors, who often find they have less power than they would like to weed out bad programs from good, when state political relationships are at stake;
c) Way of laying out clear expectations for programs in terms of the results they are to achieve, "standards" – again, most important for state directors, as part of their efforts to increase program quality;
d) Tool for program improvement, a way for programs to use data to analyze what works and what doesn't, in terms of getting results;
e) Tool for administrators to make decisions about the use of resources, at state and at local levels;
f) A way for practitioners to know they are making a difference in people's lives.

Not all the purposes can be met at once, and not all require the same measurement, data, or reporting. These different purposes have to be agreed upon and built into any performance accountability system from the beginning. It is much harder to add in other purposes once a system has been developed. One state has spent several years and a great deal of time, energy, and resources on developing a Management Information System whose prime purposes are (a) and (b) from the list above. Trying to now make it also a tool for program improvement (d) is proving difficult, because the MIS does not have the local buy-in needed to use the data for program improvement. "Being a pioneer is not so great: we looked at it from a statewide perspective. If we did it again I would spend more time getting buy-in from locals on using the information" (State ABE Director).

People who work in adult education want to know that their work is effective, that they are "making a difference" in people's lives. And they want their programs to provide better services to learners. "It makes the policy make sense – people feel they need the improvement part of this. People are working in a vacuum of knowledge about what's working" (National policy leader). Adult educators, especially at state level, make

a link between performance accountability and program improvement. "There are things that we're accomplishing and that we need to market. There's a lot that's not working well, and we need to know about it so that we can fix it" (State ABE Director).

A focus on educational improvement challenges a powerful, but outmoded, metaphor for performance accountability borrowed from industry: the production line. The dominant metaphor in measuring results portrays adult education as a production process, with adult learners rolling off the end of the line, equipped with certain skills and knowledge which can be tested and reported in the same way that businesses make sure that widgets coming off the production line meet specifications.

However, quality control at the end of the production line is an old-fashioned concept in the business world. Total Quality Management (TQM) and related approaches have become commonplace, and monitoring outcomes is only one consideration in quality assurance: "TQM views outcomes assessment as having a place in determining quality; however, its view is much more comprehensive and process-oriented" (Stagg, 1992, p. 16). As Stagg outlines, the goals of TQM are to mobilize everyone in an institution to "manage work processes; exceed customer expectations; ensure a systematic approach; measure for continuous improvement, and become involved in the entire process" (Stagg, 1992, p. 17).

So-called "high performance" workplaces build in processes at each stage of production to monitor and improve performance. Continuous improvement involves workers in monitoring inputs and outputs, how a process is working, assessing quality, and evaluating production. TQM is concerned with accountability, but above all with quality. As Stein points out in a report for the Association for Community Based Education,

> TQM's approach to quality is based on the recognition that achieving quality is not magic: rather, it is a direct result of the conditions, the processes and structures that make up the "production process." Therefore, by paying careful attention to each step in the process, and analyzing it to see how it facilitates or impedes the process, contributes to or interferes with quality, an organization can have a powerful impact on increasing quality. (Stein, 1993, p. 3)

Organizational development in the business world is now evolving beyond TQM to integrated concepts like the *learning organization*: "A Learning Company is an organization that facilitates the learning of all its members and continuously transforms itself" (Pedler et al., 1991). Learning, not just by individuals, but by the organization as a whole, is a way of responding to changing environments and multiple demands. Policy and strategy formation are consciously formulated as a learning process. Through cooperating in a common purpose, individuals and organizations develop.

In adult education, we may have a lot to learn from these business approaches which move learning out of the classroom into the organization. Performance accountability systems could be part of a transformation to a flexible and responsive adult education system. But we need to ask whether we expect performance accountability to be about better reporting or better programs (and what is the connection between these)? Is it to justify more funding or to increase program effectiveness? Is the primary audience external (funders, taxpayers) or internal (students, program staff)? How central is program improvement? These different questions are not always clearly delineated and may not be easily compatible.

Performance accountability initiatives: During the 1990s, a number of projects have focused on adult basic education system development and accountability. These include federal, state, and private sector development of program quality indicators, NIFL's PMRIS and Equipped for the Future projects, and the U.S. Department of Education's development of a national outcomes reporting system. They represent a range of ways of answering the question of what performance accountability means.

In response to requirements of the National Literacy Act of 1991, states began to develop program quality indicators, starting in 1992. Some states adopted the national model indicators developed by the U.S. Department of Education (1992). Others went through an extensive process of consultation to develop their own. By 1996, the quality indicator system had "become the guiding framework for states in their efforts to define program quality and to hold programs accountable" (Condelli, 1996, p. 14). However, these are state-level, not national frameworks, and there are no requirements for states to report their indicator measures or standards to the federal government.

Most states' quality indicators focus on process, and only marginally on outcomes (primarily learning gains). In private sector efforts, the Association for Community Based Education and Laubach Literacy Action (LLA) developed their own program quality standards and indicators (LLA, 1996; Stein, 1993). Like the state and federal program quality indicators, these also have limited emphasis on outcomes. For example, LLA's national quality standards cover governance, program management, program operations, and volunteer development. The focus is on how a program should be run, not what the results should be. Program quality efforts are just a first step toward performance accountability.

In 1993, NIFL began a project to create interagency Performance Measurement, Reporting and Improvement Systems (PMRIS) in five pilot states (NIFL, 1995b). The purpose of the competitive grant program was to help state and local leaders create interagency systems that would enhance their ability to monitor results and improve their adult literacy programs. For the first time, a systematic attempt was made to link literacy outcomes to broader state policy outcomes, and to develop measures which could be used to improve program services.

In the PMRIS pilot states, the intention was to make some significant shifts:

- from individual and separate program planning to collaborative and strategic planning across the agencies with responsibilities for adult education and training;
- from measuring inputs and processes to measuring results;
- from periodic, sporadic, program evaluations to continuous improvement.

Given the magnitude of the paradigm shift required for such systemic changes, it is not surprising that the five states selected for PMRIS grants (Kentucky, Tennessee, Hawaii, New York, and Alabama) experienced mixed results within the timeframe of the grants (Swadley & Ziolkowski, 1996). The four states that continued into the second year of the grant (Alabama chose instead to work on an MIS for ABE alone) made substantial progress in development of a more collaborative and strategic approach to planning. Building a system for measuring results rather than inputs, process, or outputs was more challenging and was incomplete at the end of the grant period. All of the states found it difficult to develop broad agreement on outcomes which could be measured. Using the data for continuous improvement at the program level was still under

development. All four states did learn much about the complexities of system change, and their lessons are useful for other performance accountability system-building (Swadley & Ziolkowski, 1996; see also Condelli & Kutner, 1997).

Overlapping with the PMRIS project, NIFL also began development of the Equipped for the Future (EFF) project in 1993, an ambitious system reform initiative to:

> develop a consensus of support for the very real changes students are trying to make in their lives, identify the essential skills and knowledge they need to move successfully into the future, and build a reliable system that can make their hopes a new reality. (Stein, 1997, p. v)

Equipped for the Future focuses on developing a vision and direction for literacy education, and a framework for what it needs to achieve. The EFF initiative will be described further in Chapter Four.

During the last few years, the state directors of adult basic education, recognizing the demands for accountability being placed on the field by external sources, pressured the U.S. Department of Education to focus resources on this issue. In a series of meetings held between 1995 and 1997, the state ABE directors met with federal officials to work on design for a framework for a national outcomes reporting system. (Condelli & Kutner, 1997) In March, 1996, the state directors adopted the following resolution:

> We recommend a collaboratively funded and managed project to analyze and synthesize accountability systems that have been developed nationally and in separate states that incorporate adult education outputs and outcomes. The project will continue the next steps of work begun here by state directors to draft protocols, determine how data would be collected and how reliability could be optimized. The project will involve state directors of adult education and other stakeholders in setting project policy and project operation. (Condelli & Kutner, 1997, p. 1)

In response to this resolution, the Office of Vocational and Adult Education (OVAE) of the U.S. Department of Education created the National Outcome Reporting System Project, with technical support from federal contractors, Pelavin Research Institute, still under development at the time of writing.

In an early meeting, the state directors agreed on seven categories of outcome measures for adult education to be used to demonstrate the program's effectiveness:

- economic impact,
- credentials,
- learning gains,
- family impact,
- further education and training,
- community impact,
- customer satisfaction. (Condelli & Kutner, 1997, p. 3)

As the reporting system is developed, it is intended that "all states will report a uniform set of quantitative measures using an individual student record system at the local level" (Condelli & Kutner, 1997). However, some basic questions are yet to be decided, including the policy goals which adult education supports, the purpose of the reporting system, the outcomes, indicators, and the measures to be used to collect data and report.

Performance

The key to designing accountability is getting clear about goals, about what "good performance" would look like. This is not easy. The purposes of literacy education are contested ground. The field as a whole is not agreed about what literacy is, or what would constitute good performance. Is it skills or practices, individual advancement or community development, productive workers or good citizens? Is literacy a right or a requirement for the good life? Different stakeholders – learners, teachers, funders, employers – have different expectations. If learners come to a program to learn to read to their children, and policy makers want them to get a job, we have a recipe for mis-match of goals and for failure rather than success. Without clarity and agreement on these questions, performance accountability cannot succeed. We cannot assume that all the goals fit together, that all are equally valuable, that all can be met, especially with limited funding. But what is the process for reaching agreement on performance?

An Australian literacy researcher, Delia Bradshaw, has noted the shift in terms between what most people mean by performing and what "performance" has come to mean to government administrators:

> ... to per-form means literally to "bring into form," to manifest the subtle world of ideas, spirit, creativity and inspiration in the tangible

> world of form. Performing then becomes a powerful, multi-stranded activity that entwines the threads of informing, reforming, and transforming into the one strong, stress-resistant braid.
> [For bureaucrats today, though] ... "performance" means meeting certain standards, standards that must conform to the dictates of a competition-driven economy. Performance then becomes the setting, classifying, enforcing, monitoring, measuring and rating of these standards. (Bradshaw, 1996, p. 56)

When we ask just what good performance is, we are inevitably in the realm of values. While measurement of performance may be a technical matter, the question of what needs to be measured, what we believe to be good performance, is shaped by notions about literacy. Concepts of literacy have been changing over the last 30 years, from a school-based view of literacy, to a scale of functional skills or competencies, to social-contextual concepts of multiple "literacies." In particular, the concepts of literacy as competency and social-contextual literacy are widely held, and both challenge in different ways traditional school-based approaches to performance.

Literacy as competency: When literacy meant what was taught in schools, it was relatively unproblematic. It may not have had much to do with the reality of life for most people, but it could be tested, using the same tests that school children took. The developing concept of literacy "competency" shifted the focus away from decontextualized school-like skills to the application of reading, writing, and math in everyday life settings (and especially in workplaces). The Adult Performance Level (APL) study of the early 1970s began that shift, albeit surrounded by a great deal of controversy (see Chapter One).

Other competency-based approaches to understanding literacy continued and flourished, trying to avoid the criticisms of APL. CASAS (Comprehensive Adult Student Assessment System) is the best known. Begun in 1980 in California to provide an approach to learning and assessment based on tasks related to everyday life, CASAS identified a core of competencies and learning outcomes, and developed assessments of these. Since then, CASAS has grown nationwide, and in at least six states is part of the state-mandated assessment system (Kutner, Webb & Matheson, 1996).

ESOL has also developed "communicative competency" approaches, focusing on purposeful communication, what learners can do with

language, as the core for instruction. The Basic English Skills Test (BEST) and the Mainstream English Language Training (MELT) curriculum took a competency-based approach to both skills development and assessment (Center for Applied Linguistics, 1983, 1989).

The increasing dominance of this view of literacy as "competence" in "real-life tasks"[10] is demonstrated in the National Adult Literacy Survey (NALS). The NALS aim was "to profile the English literacy of adults in the United States based on their performance across a wide array of tasks that reflect the types of materials and demands they encounter in their daily lives" (National Center for Education Statistics, 1994, p. xiii).

Although it replaced a single scale of literacy competence with three scales (prose, document and quantitative), NALS, like other competency-based assessments, has several underlying assumptions. It assumes that literacy is a technical skill which is transferable to multiple contexts, and that a single set of competencies can be defined. It assumes that competencies can be measured in formal tests, and that there is an equivalence between how well a person completes the pencil and paper test and his or her ability to perform a task in real life. And finally, it assumes that the tasks chosen for the test are both important to everyone and can be accomplished in only one way (see Box 2.4 for challenges to these assumptions).

Functional and competency approaches to literacy assume that knowing how well someone reads tells us something about how they carry on the rest of their lives. There is not a great deal of evidence for this. Sticht's work in the military and in workplaces suggests that reading tests have limited relevance to effectiveness on the job (Sticht, 1988). Whether the deficiency is in the reading test or that literacy is not closely related to life competence, is yet to be determined.

One of the problems with competency approaches to assessment is that there are no *a priori* reasons for choosing one task over another. There is no theory to guide scaling the difficulty of tasks. De Castell and Lukes criticized some time ago the "systematic carving up of the universe of literate behavior into ever more specific kinds of competence" (1986, p. 9). They suggest a need for "second-order competences" which would enable individuals to select, apply and modify an existing reading

[10] The quotation marks signify the problems in defining both competence and real life tasks.

competence to a new task. Concepts like Bourdieu's "expanded competence" (when to speak, keep silent, speak in this or that style) need further research.[11]

Box 2.4. Reading between the lines: literacy as competency

Reading a bus timetable is an old favorite of competency-based assessment. Behind the task of reading a fictitious bus schedule and answering questions about it for assessment purposes are various assumptions:

- that bus timetables are important in everyday life (even for people who live in rural areas perhaps 50 miles from the nearest bus route);
- that people find out which bus to take and when to catch it from reading the schedule (though many fluent readers would use alternative strategies like asking someone);
- that the task of reading and answering questions in the test is the same as the task of using a real bus schedule in ordinary life (although the contexts are very different, and the transferability not researched);
- and that to use a public transportation system effectively requires the ability to read and decode a bus timetable, and vice versa, that if you can read a bus timetable you can figure out public transportation (again not researched).

These assumptions are seldom directly addressed in competency-based assessment.

Perhaps the central issue is who decides what is competence, and who chooses the tasks. Back in 1979 Hunter and Harman argued: "Who but the person or group involved can really describe what 'effective functioning in one's own cultural group' really means?" (1979, p. 19). That question still needs to be asked.

Literacy as social and cultural practices: New Literacy Studies research explores how literacy is used within social groups. Much of the impetus has come from outside the field of education – from

[11] Bourdieu says what is problematic about language and literacy is not so much being able to produce grammatically correct, coherent sentences, but "the possibility of using an infinite number of sentences in an infinite number of situations, coherently and pertinently" (Quoted in de Castell et al., 1986, p. 10).

anthropologists like Brian Street and Shirley Brice Heath, psychologists like Sylvia Scribner, cognitive scientists and linguistics researchers like Jean Lave, Barbara Rogoff, James Gee, and Colin Lankshear. The social practices of literacy from North Carolina to Iran and South Africa, from Philadelphia dairies to California's Silicon Valley, have shaped our understanding of how literacy works (see, among others, Barton, 1994a; Hamilton et al., 1994; Hull, 1997; Lankshear, 1997; Prinsloo, 1997; Street, 1984, 1995; Szwed, 1981).

The concept of literacy which is emerging from this painstaking research is very different from the one-dimensional scale that holds sway in public policy. Instead of seeing literacy as a matter of technical skill, which once acquired can be applied to many different tasks (Street's "autonomous" view of literacy), the new view sees literacy and language as embedded in social context. Home, work and school have different literacies, as do different communities and social groups, like Heath's white working class community of Roadville and black working class community of Trackton (Heath, 1983). *Reading* has no meaning unless we say who is reading what, in what setting, and for what purpose – we have to separate the medium (text) from the message (meaning) (Street, 1984, p. 221). Street calls this perspective on literacy "ideological" because it is "implicated in power relations and embedded in specific cultural meanings and practices'' (Street, 1995, p. 1).

Box 2.5. Key terms in New Literacy Studies

Literacy practices – "General cultural ways of utilizing literacy" (Barton, 1994b, p. 5) that people draw upon in the varied contexts in which they live their lives – school, work, home, social groups, neighborhoods. They include not only behavior but also meanings, values, and social relationships. Literacy practices are shaped by power relationships – some are more prestigious and valued than others.

Literacy events – "The particular activities in which literacy has a role" (Barton, 1994b, p. 5) which draw on the general literacy practices in that domain, but are directly observable. Literacy events might include reading a newspaper, writing a grocery list, writing a letter to a friend (from the domain of family life), reading a company memo or a payslip, writing an accident report (from the domain of work life), writing in a workbook, reading a textbook, writing a term paper (from the domain of school).

Domains – the broad contexts of life in which we operate. We all engage in the home domain, some in work, some in school, some in a variety of social groups (Rotary or trade unions, neighborhood associations or political parties). Each domain is shaped in turn by the broader culture and by class, gender, ethnicity, and regional variations. These differences give rise to the concept of **multi-literacies** to indicate the different forms of literacy which characterize these different contexts.

Lankshear has described the shift in thinking about literacy:

> Notions of reading and writing as specific (cognitive) abilities or sets of skills based on an identifiable technology (e.g. alphabetic script) held sway within educational theory and practice, almost to the point of having a monopoly, until the 1970s. Since then greater theoretical space has been usurped by conceptions of reading and writing which stress their inherently social character and embeddedness in larger social practices. (Lankshear, 1997, p. 2)

As James Gee has argued, the change is from the view of literacy as "something we do with our heads" to seeing literacy as "something to do with social, institutional, and cultural relationships" (Gee et al., 1996, p. 1). Parallel research in language acquisition, and especially ESOL, has taken a similar social-contextual approach (Auerbach, 1989; Macedo, 1994; Wallerstein, 1983). Several key terms are associated with this new research in literacy (see Box 2.5).

Over the last decade, these socio-cultural theories of literacy and language have begun to be applied in literacy education. Teaching shaped by such theory starts with learners' own literacy practices and events, and uses the learner's own purposes for literacy to create learning opportunities.[12] The aim is not simply skills acquisition, but making meaning and critical understanding of how literacy is used in social contexts. As Auerbach argues, "Literacy is meaningful to students to the extent that it relates to daily realities and helps them to act on them" (Auerbach, 1989, p. 166). The focus of teaching is on meaning rather than on mechanical skills. Many of us can "read" texts we cannot understand (tax forms, perhaps, or insurance documents, or scientific journals) – we can read the words, but not the world they represent.

However, while the theories of New Literacy Studies are being applied in teaching, they have had much less currency at the level of educational systems and policies – institutions, funding, and accountability. Such a shift in the understanding of literacy means that "performance" is defined differently and requires a different approach to accountability.

Perspectives on performance

The legislation which has regulated the field of adult education for over 30 years has given us a broad, and rather loose, vision of adult education.

As recently re-stated by the General Accounting Office, the purpose of the Adult Education Act is to:

- improve educational opportunities for adults who lack literacy skills necessary for effective citizenship and productive employment;
- expand and improve the current adult education delivery system; and
- encourage the establishment of adult education programs for adults to (1) acquire basic skills needed for literate functioning, (2) acquire basic education needed to benefit from job training and obtain and keep productive employment, and (3) continue their education to at least the secondary school level. (GAO, 1995, p. 14)

[12] For example, in Auerbach, 1989; Auerbach & McGrail, 1991; Gillespie, 1990 and 1996; the Adventures in Assessment series from SABES; McGrail, 1995; Fingeret, 1993; Literacy South, 1997.

Within this loose vision, many competing visions of literacy education have been able to develop. Many programs operate without a clearly articulated vision of what literate functioning means, or a set of clear and consistent goals about "effective citizenship" and "productive employment." It is not clear what literacy skills are actually needed for effective citizenship and productive employment. Without a common framework, it is difficult to demonstrate, or achieve, success. A principal criticism in the GAO report on adult education programs centered on this lack of clarity.

> Evaluating program results depends on clear program objectives as well as criteria for measuring the achievement of those objectives. The broad objectives of the State Grant Program give the states the flexibility to set their own priorities but, some argue, they do not provide states with sufficient direction for measuring results... Several experts and program officials told us that the State Grant Program lacks a coherent vision of the skills and knowledge adults need to be considered literate. (GAO, 1995, p. 23)

One's perspective on what is good performance in adult education, and what should be measured, depends on one's context and position. Learners' perspectives on what is a successful program may not be the same as policymakers' perspectives. Learners may want a program that treats them with respect, allows them to feel successful, provides them with the learning opportunities they want, and supports the results that are important to them, whether they are a credential or the ability to read to their children. Policymakers may not care about any of the process, but want a program that gets people into jobs. Educators, rooted in the kindergarten-through-higher-education tradition, may care most about credentials.

The concepts of performance among the people interviewed for this policy paper vary considerably. Some have a narrow focus on "basic skills" and improving literacy, while others have a broader focus on the purposes to which literacy is put in society. Some see employability as the primary goal for improving basic skills, while others include personal growth and self confidence, family impacts and involvement in the community. Some think adult education should respond to the full range of individual student goals, while others think there should be a finer focus, serving fewer students with greater impact in limited areas. One person suggested, "We treat learners as if they're homogenous, but they're not. We need to

customize services, work with sub-populations" (from an interview). Some believe that individual learners' goals are fully compatible with the broad social purposes conceived for literacy. Others suggest there may be tension between the two.

Most seem much clearer about the more distant outcomes desired of literacy education (in terms of work, family, community involvement) than about what it means to be literate. If literacy is defined as autonomous skills that are transferable to any context, then performance can be conceived in terms of mastering grammar rules and vocabulary, word families and decoding text. The traditional methods to document such learning – standardized tests – will be seen as appropriate.

When literacy is defined as social practices, rooted in context, intertwined with social relationships and power, constantly changing and being changed, then what is important is what students do with what they learn. Standardized tests are not very useful as guides of this kind of performance, precisely because they do not measure what students can do with their knowledge, only whether they perform well on the test.

Performance accountability demands careful setting of goals. Again the business world offers some lessons: Peters and Waterman have argued that excellent companies possess "simultaneous loose-tight properties." Behn says:

> By this, they mean that these firms are very "tight" about the objectives they are attempting to achieve while simultaneously being very "loose" about how to achieve those objectives. Too often, however, government agencies are forced to operate with simultaneous tight-loose properties. The legislature, being unable to agree on exactly what the agency should accomplish, enacts legislation with very vague purposes. Then, however, it imposes very strict rules under which the agency must pursue those vague purposes. This of course leaves each individual legislator free to complain when the agency fails to achieve the specific purpose that he or she thought was embodied in that vague legislation – and to complain when the agency violates any of the strict rules. (1993, p. 2)

If performance goals are framed too tightly, without getting the input of all the players, there is a risk of conflict between goals framed at the system level by policymakers and those framed at the program level by

practitioners and/or learners. If performance goals are framed too loosely, no common mission and purpose shapes the work, and when "anything goes" accountability suffers. The challenge is to come to a common agreement that fits the theories and research, fits society's aims, and fits the practice. That agreement can only be the result of a broad-based public debate, and this will be discussed further in Chapter Four.

For effective accountability, the system needs both clarity about the ultimate goals and flexibility in how to reach them. It also needs the capacity to achieve the goals. The next chapter addresses issues about the capacity of the system to be accountable, as well as the development of indicators and measures, and some particular concerns about measuring learning.

THREE. CAPACITY TO COUNT AND TO PERFORM

Adult basic education is struggling to create a national accountability system without having created a national service delivery system. Accountability requires not only clarity about the multiple purposes for adult education, but also capacity among the diverse institutions which provide educational services. As Robert Behn points out,

> The publication of a performance report card can motivate an agency, but it does nothing to change that agency's capabilities.... To make a difference, the monitoring of performance must be accompanied by some assistance in creating new, expanded organizational capacity. (Behn, 1993, p. 13)

Creating capacity implies two things: the capacity to perform – to achieve the goals set – and the capacity to be accountable – to document what has been achieved.

Capacity to perform

Efforts have been ongoing for years to strengthen the capacity of the adult basic education system to perform. Amendments to the Adult Education Act in 1988 were designed to strengthen evaluation and accountability structures. Further amendments in 1991 strengthened staff development and technical assistance. Nevertheless, adult basic education is a complex and incomplete system characterized by diverse and multiple funding sources, institutional arrangements, goals and objectives, reporting formats, and learning opportunities. Osborne and Gaebler refer to the "funding streams" in adult education and employment training in Michigan as "puddles."

> It was, in effect, a nonsystem: 70 categorical pots of money with no coherence, no user-friendliness, little performance data, and little management. People in government often talk about "funding streams." But these were puddles: funds could not flow from one to another as needs shifted. They were driven by legislation and organized for the convenience of public agencies – not driven by demand and organized for the convenience of customers. (1993, p. 190)

How well adult basic education performs right now is not firmly established. The GAO report on adult basic education concludes:

> The program has had difficulty ensuring accountability for results – that is, being able to clearly or accurately say what programs funds

> have accomplished ... the data the Department receives are of questionable value. Because state and local client data are missing or inaccurate, attempts to make the program accountable may be compromised. (GAO, 1995, p. 33)

What we do know is problematic. The national evaluation of adult education (NEAEP) found that most participants stay in the program a very short time: adult secondary education participants receive on average 28 hours of instruction, ABE 35 hours, and ESOL 113 hours (Young et al., 1995). Except for ESOL, no direct relation was found between persistence (total hours of instruction) and test score gains (Fitzgerald & Young, 1997).

Stites re-analyzed the small sub-sample of the NEAEP population for which there was good pre-test and post-test data, and found that less than half of the learners in the group showed strong evidence of having improved their levels of reading skills in the period in which they received instruction (Stites, 1996). Sticht and Armstrong's review of test reports from a wide variety of programs across the nation found that adult literacy programs consistently seem able to increase "grade level" test scores by 0.5 to 1.5, but that in the limited longitudinal data that exists, the rate of gain slowed or stopped after the first year of program participation (Sticht & Armstrong, 1994).

What data we have suggest that, on average, learners don't stay long, may make initial limited learning gains, but not a lot of long-term literacy skill gains. Clearly the average disguises many learners who do persist, and who make substantial learning gains – but also those who drop out quickly and make little or no learning gain.

Those who gain the GED credential do seem to realize long-term impacts on employment and earnings (see summary of research in Murnane & Bickerton, 1997). The problem for ABE is with those who do not gain the credential, and who enter – and probably leave – with limited literacy skills.

The field's capacity to perform well is challenged by many factors. The *Jump Start* report says, "Overall, the field is intellectually, institutionally, and politically weak and fragmented" (Chisman, 1989, p. 5). Not only is the knowledge base fragmentary and unsystematic, but the institutional base is also a "jumbled system" (Chisman, 1989, p. 9). It is also resource-poor. While in some respects the picture has improved since 1989, with

both national and state level system-development efforts, in most states, most staff are part-time; per-student funding is low; volunteers continue to have an essential role in student services; and most programs are not able to meet other client needs such as for transportation or childcare (Moore & Stavrianos, 1995, pp. 12-13).

Capacity for accountability

There is evidence that the capacity to measure performance in adult education is also severely limited. Recent reports have highlighted issues in capacity to collect, let alone to use, valid and reliable data about performance. These include:

- Absence of important data, even basic data on how many students enroll: the NEAEP found that: "Nor did many programs have any precise idea of the number of adults newly enrolled each year or of the number of different individuals enrolled at any given time or over the period of a program year" (Young et al., 1995, p. 7).
- Difficulties in collecting valid data: Pelavin Associates found that "local programs often lack staff and other resources to collect data, which frequently results in incomplete or low quality data" (Condelli, 1994, p. 10).
- Double counting and under-counting: the GAO noted "serious problems with the quality of the statistical reports, some of which are based on double counting or under-counting of students in adult education programs"' (GAO, 1995, p. 26).
- Programs do not understand the purpose of the data they do collect: Pelavin noted "local programs' lack of understanding of the purpose of data collected and their inability to access the data they collect" (Condelli, 1994, p. 10).
- There are doubts about whether the information collected does measure real performance: the GAO reports that "some [local staff] said they thought that the information they are required to report does not accurately reflect the accomplishments of their adult education students" (GAO, 1995, p. 27).
- Staff turnover severely hampers program capacity to gather and use good information: the national evaluation of ABE programs found that "within the first 6 months of data collection ... 16 percent of program
-

directors trained in the requirements of the study had departed" (Young et al., 1995, p. 7).

The lack of capacity to measure performance is a backdrop to any efforts to develop accountability systems. Without the provision of resources and support, the demand for performance data is likely to be misplaced, because the management systems are not in place either to report data or use it. As one person interviewed said:

> We have a part-time fragmented program delivery system. We do very little professional training on assessment and evaluation and the link between instruction and assessment. That is critical or none of this will work. None of the rhetoric will be useful unless you include a strong staff development component. Performance accountability has to be useful to instruction, to program improvement, in allocating resources. Programs don't know how to do it. (from an interview)

When asked to report numbers, literacy programs will indeed report numbers – but when they see no purpose in the numbers, do not use them themselves, never see reports based on the numbers, and place their own priorities on providing learning opportunities, there is little incentive to make the numbers accurate. Programs which are using such data themselves to improve learning opportunities will make sure that the data are valid, reliable, and complete. The two kinds of capacity – to perform and to be accountable – are linked.

Building capacity for accountability: state experiences

A number of states already have begun to develop performance accountability systems, creating a substantial body of experience. Performance accountability projects in Oregon and Texas, the four PMRIS states and six NGA states, as well as other states like Iowa and California, have been recently reviewed elsewhere (see Condelli & Kutner, 1997; Kutner et al., 1993, 1996; NGA (nd); NIFL, 1995a; Swadley & Ziolkowski, 1996). Three other states' approaches to building capacity for accountability are described here: Connecticut, Pennsylvania, and Arkansas. These accounts are based on interviews and other documentation. They illustrate some of what has been learned about capacity and measurement.

BOX 3.1. Connecticut

has developed a reporting system whose primary focus is data management for the state level.[13] In Connecticut, ABE is state-driven, rather than federally-driven, and it has been system-building for a long time. In 1983, a new state law required local school districts to offer adult education, which includes citizenship education, ESOL, secondary school completion (through local high school diploma, external diploma or GED), and adult basic education. The law required state funds to be matched by local cash funds, on a sliding scale based on poverty level in the community. No other state requires local school districts to contribute cash rather than in-kind contributions to the ABE system. Not long after this law was enacted, Connecticut initiated a competency-based instruction and assessment system, using CASAS. The Connecticut Adult Performance Program (CAPP) took 5 years to be implemented and included the specification of competency-based instruction and assessment and the creation of a support system including staff development and technical assistance (for more details see Alamprese, 1993). *"CASAS gave us the first tools ... we all report on a single standard so there can be some uniformity and understanding across the state."* (Roberta Pawloski, state ABE director).

As part of CAPP, Connecticut has been developing a Management Information System (MIS) since 1991. The primary driver was the state-level need for accurate information about what was happening in local programs. Expectations of federal demands for performance data was a secondary driver, and state staff pointed out to local program directors that they themselves also needed such data in reporting to local school boards. The process of MIS development has been intense and time consuming – *"a process not an event"* says Pawloski.

The MIS has had significant impact at the state level. It has enabled the state to report on changes in its target population and services over the last ten years, and to conclude that ABE is serving its target population and that their skill levels are less than 10 years ago. This helps make the case for increased funding, and it assures other state agencies like the Departments of Labor and Social Services that ABE has a system in place and is accountable.

At the level of local programs, the impact of the system has been slower to emerge. Pawloski says that "accountability is something that scares program directors – they want to live in anonymity, but at the same time they say that no one respects and acknowledges us." Uneven access to computers and resistance to computers during this period added problems. She says that the MIS is beginning to have an impact at the local level, both in helping at least some local program directors make their case for funding requests, and in some cases in program quality: "They can track individual teachers by score gains, attendance records. It is beginning to be a tool for them*."*

Although there was an advisory group which included local program directors and other agency staff, with the power of hindsight, Pawloski says that if they were developing the system again she would get more buy-in from local programs from the beginning on what data they need and how they would use it. "We have been remiss on how to use the data [at the local level]: adult educators are not experts on data analysis and how to use data." They are now offering more training for local staff on using the system. She stresses the need for the system to be flexible in responding to changes in the delivery system and funding streams, and for ongoing professional development and support for local programs.

[13] This description is based on an interview with Roberta Pawloski, Connecticut ABE Director, on 4 April 1997, and the final evaluation report on CAPP (Alamprese, 1993).

Issues that emerge from Connecticut's experience include:

- if local programs are not involved in system design, they are likely to resist it;
- a significant investment in technology at the local level is needed;
- extensive professional development support is needed, not only at the introductory stages, but ongoing (because of constant staff turnover).

BOX 3.2. Arkansas

has approached accountability through the development of performance funding.[14] Like Connecticut, ABE in Arkansas is driven more from the state level than the federal level as state funding greatly outweighs federal funding. According to Garland Hankins, state ABE director, performance funding is a response to demands from the governor's office and state legislature for increased accountability. The current funding formula has a performance element: each program gets allocated base funding, then additional funds, 50 percent of which are allocated on the basis of the literacy rate in the community and 50% on the program's performance in terms of learning gains. The state is engaged in a process "under the gun" of the state legislature, to identify better measures of "performance," and when those are in place all funding will be allocated 50% on the literacy rate and 50% on performance.

In addition to this funding formula, Arkansas has in place "effective and efficient" criteria for funding. The state MIS collects data on administration costs, monitoring and operations costs, and three levels of learning gains. Programs must score 75 points based on these criteria to be deemed "effective and efficient." If they score below 75 points, they must work with the state program managers to prepare an improvement plan in order to receive the next year's funding. In the second year if they still do not reach the "effective and efficient" level, administration of the services is moved to another provider.

The purpose of the benchmarks and performance standards now under development is to give legislators something that makes sense to them in economic terms: Hankins says: "Soft ideas are not cutting it, they won't fund it." A committee from the state Board of Education and including five practitioners is working on the new standards, and is looking at proposals for each county to produce a plan based on local needs, using broad state-set criteria to establish their own benchmarks.

As in the other states that have embarked on accountability projects, the reaction of local programs has been the concern, "can we really produce?" Not everyone is happy at this stage, but after several years of talking about the need for change and the need for accountability, Hankins feels most programs have accepted "that something has to change." Leadership institutes and staff development have targeted support for the accountability work.

The biggest challenge now is "getting indicators that really indicate performance." There is some tension between the demands of policymakers and the expectations of adult educators, Hankins says,

> Policymakers don't understand adult education, they focus on return on investment, they don't understand softer life skills, it's hard to sell to them. There is more to the story than jobs and income. Sometimes self esteem, confidence building, motivation is the best you can do for a student. But you have to tie it to something stronger to satisfy the legislators.

[14] This description based on an interview with Garland Hankins, Arkansas ABE Director, on 1 April, 1997.

Arkansas' experience with performance funding also highlights some issues:

- determining what is good "performance" in terms that satisfy both practitioners and legislators is very challenging;
- getting local buy-in can be difficult, when some programs will see performance funding as punitive, not necessarily fair;
- there is a concern that programs will lose their focus on student needs and wants in order to respond to policymakers demands. Hankins says, "I have to hope that the folks at the local level are student-driven, because I can't be."

BOX 3.3: Pennsylvania

is developing an accountability system whose primary focus is improved program evaluation, rather than an initial emphasis on reporting aggregate data.[15] State ABE Director, Cheryl Keenan, says it was deliberately not intended to begin with a state-level MIS with measures and standards, "because all that would tell me is that I'm not getting good learner outcomes." Instead, the effort was conceived as essentially about system change, working both from the top down and from the bottom up, with both program improvement and accountability to result.

Project Equal, launched in 1994, is intended to "build the capacity of local adult education providers to collect and use data about their learners and their programs." The purpose is program improvement: "to become more effective in meeting the education needs of our adult learners" (Pennsylvania Bureau of Adult Basic and Literacy Education, n.d). At the same time, system-building from the top has included performance standards, program and staff development, and state funding strategies.

Keenan outlined the working assumptions behind the project:

- program services need to be improved at all three levels – learners, staff and program.
- assessment plays a key role in quality service delivery, when integrated with instruction.
- using data to analyze program issues helps to determine solutions, and participatory decision-making involving staff and administrators is optimal.

At the local level, the project has been developing through a "diffusion" approach, in which pilot sites (innovative local programs) have been selected (10 each in Years 2 and 3). The pilot sites were asked to set up a process to look at some data about their programs and plan program improvement measures. They also reported pre- and post-testing of learning gains on sample populations of learners, and this data provided an initial basis for creating performance standards. The pilots revealed the need for intensive staff development around data collection, reporting and analysis.

At the same time, efforts are underway to define performance standards at the state level:

> They are based on results obtained in actual practice by programs capable of valid and reliable data collection and analysis over a significant period of time. Program performance standards express how well or how proficiently a program is doing its job of educating adults in the areas addressed by the standards (Pennsylvania Bureau of Adult Basic and Literacy Education, 1997).

By 1997, Pennsylvania had developed draft performance standards that have been disseminated for program feedback. Keenan says, "We want people to understand that the standards are one tool of a range to improve programs ... as we develop capacity we will tie them to funding, but not yet." The standards themselves are preliminary, and no-one is satisfied with them yet. More important in these early stages, she says, is getting programs to value and use program data. "If local people don't value the data and use it locally, they will give you any number you want. You have to invest practitioners in valuing data – if you can do that, then accountability is a breeze."

In the Pennsylvania view, continuous improvement requires symbiosis between practitioners, programs and the state, feeding data up and down between them. Everyone at all levels needs to understand and appreciate the usefulness of the data for the system to work. This approach needs extensive staff development and participation from the local level at every stage. As Keenan says, "We can't sit at the state level and say we're going to do system change at the local level." But the state can be a "change agent," and use its resources (funding, staff development resources) to support system change.

[15] This description is based on interviews with Cheryl Keenan, Pennsylvania ABLE Director, and research consultant Judy Alamprese, and on reports on Project Equal, including Pennsylvania Bureau of Adult Basic and Literacy Education, 1997.

The Pennsylvania experience highlights a number of issues for performance accountability:

- a program improvement approach to accountability requires practitioner participation from the beginning – local purposes as well as state and national purposes have to drive data collection;
- developing standards from the bottom up is a slow process (but because they are based on actual and reliable learner data they may be more realistic);
- any system change effort needs to be both bottom up and top down – every level is connected;
- change agents play a crucial role – encouraging and channeling resources to support new initiatives.

The three states discussed here have taken very different approaches to the same problem of accountability: Connecticut's Management Information System (MIS) is designed to give the state agency detailed information about the performance of local ABE programs. Arkansas' "performance funding" system rewards programs with additional funding for good performance and withdraws some of the funding for programs which fail to perform to expected standards. Pennsylvania is in the midst of developing accountability closely integrated with program improvement.

The experiences of these three states, and their varying answers to the central questions about performance and capacity, demonstrate the range and variability of the field of adult basic education. They may lend weight to the argument by some that there is not really a "system" of adult basic education in the United States, but rather multiple systems, multiple goals, and multiple products. Whether this is good (more responsive to local and learner needs) or bad (haphazard, uneven quality, unaccountable) depends on the eye of the beholder.

Despite the differences, though, there are some commonalities. All three states found that the process of developing accountability systems was slow and time consuming. All three states now place emphasis on local program participation in system planning and design. All three states are ultimately looking to link, perhaps integrate, program improvement efforts with accountability efforts. All three states have had difficulty with getting accurate data from the local level, with establishing accurate and reasonable indicators of performance, and with training and supporting local programs to analyze and use data for their own purposes. All three

states emphasize the importance of sustained and systematic professional development and technical support.

Measuring performance

The capacity to be accountable is not only about programs' ability to collect data, but also about the quality and value of that data. To know what is important to measure requires a decision on what is important. Without a clear vision of this, measurement can become a sterile exercise to "come up with the numbers" to satisfy external agencies. Measurement has to shed light on what matters: the outcomes.

What is counted usually becomes "what counts." This is true in education as in business. A newspaper reporter writing about businesses changing from accounting-based performance measures in favor of shareholder value, says:

> This is much more than a debate about measuring performance. How companies measure value determines how they are run. Hanson measured the performance of all its businesses in terms of return on capital employed. It worked – but only for a time. When the world and the stockmarket changed, Hanson didn't, and self-destructed on its own performance metric (Caulkin, 1997).

Measures of all kinds are used to sort people into categories, reward some and punish others. Some do not do what they are intended to. Sticht describes what happened when the military miscalibrated its aptitude battery in the period 1976-1980, thus permitting "over 300,000 lower-aptitude people who would have been rejected into the military without the services being aware that they were getting 'functionally illiterate' personnel" (Sticht, 1988, p. 69). What happened? Nothing. Data shows that in terms of completing military training, and receiving satisfactory job ratings, 80% of those 300,000 "low aptitude" people performed 80-95% as well as average-aptitude personnel. As Sticht remarks, "These data suggest that great caution ought to be exercised in declaring people functionally incompetent because of their performance on literacy or other types of aptitude tests" (1988, p. 7).

Measuring the wrong things is a problem in many government endeavors. In Britain, for example, the New Economics Foundation (NEF) points out the inadequacies of the dominant economic indicator – gross domestic product (GDP). While the GDP has almost doubled in Britain since the early 1970s, the quality of life (as measured by sustainable economic

indicators) has declined. Violent crime has quadrupled, the number of workless households has tripled, car traffic has almost doubled, and the concentrations of climate-changing gas in the atmosphere have been growing (NEF, 1997). Economic planning, they conclude, needs to be steered by much broader indicators than GDP.

In adult education, a similarly one-sided picture would emerge if indicators measured only how many adults enrolled in a basic skills program, rather than how long they stayed. Or if indicators measured only academic credentials, not if students had different achievements, such as passing a commercial drivers license test. Indeed, these are precisely what ABE has been measuring. Annual reporting to the federal government has focused on how many students are recruited and stay for 12 hours or more, but not on whether they learn. So recruitment has been more important than how long learners stay, and academic credentials like the GED have been valued more than practical applications of skill and knowledge.

Measuring the wrong things is only part of the problem; distortion effects are another. Distortions occur when agencies achieve their performance indicators but not their underlying mission. JTPA "creaming" is perhaps the best-known distortion example in employment training. Early performance indicators focused on getting clients into jobs quickly and cheaply. They encouraged programs to recruit clients who were already well-qualified so that they could meet their performance requirements (see discussion of distortions and unintended consequences in Condelli and Kutner, 1992). There are also examples of distortions in education: if teachers "teach to the test," more students may gain passing grades, but not necessarily a good education. The percentage of GED test passers would improve if less able students were discouraged from taking the test – but no more people would necessarily pass.

When measurement for its own sake becomes important, programs quickly learn to "game the numbers." To avoid a numbers game, you have to measure what you value, and value what you measure. There is nothing magical about measurement itself. Right now, in most adult education programs, there is a mismatch between what gets measured and what is valued. As the New York PMRIS project found:

> ...much of the information gathered by programs is not reported to anyone beyond program level. This is interpreted by local programs to mean that this information is not valued by other levels

> of the system. ...mandated data is not shared with the program managers in any direct way. This is interpreted by local programs to mean that the important information is kept from program level. (Toms, 1995, p. 5)

The core of good performance is to know where you're going, what it takes to get there, and to have the capacity to achieve it. Then measurement can tell you if the process is working or not. The literature on measurement shows that programs should be collecting data that they use, and whose value they understand and appreciate. This could impact both the quality of services, and the quality of the data.

What to measure

Measures can be developed for different aspects of program performance. Box 3.4 defines the key terms that will be used in this policy paper. Measuring more than one kind of indicator is important to give a full view of performance and to provide tools to improve the system. Accountability systems need input indicators to understand the capacity of the system, process indicators to understand whether different approaches produce different outcomes, output indicators as short-term and immediate measures of performance, and outcome indicators because in the long-run these are what matter to society. It is not enough to know if a program has a great process, if we do not know that these processes produce the outcomes. No single indicator can suffice to measure performance, particularly of as complex an enterprise as adult education.

Box 3.4. Indicators, measures, and data

An **indicator** is a symbol that stands for an aspect of reality. Fever is an indicator of illness, for example – it is not the illness itself, and it is not the only indicator of illness (you can be ill without a fever), but it is quite a good indicator in that if you have a fever you are likely to be ill.

A **measure** is a way of operationalizing that indicator. Calibration of a thermometer is a common measure of fever, for example, and there are alternative measures which don't require a thermometer, like a mother's hand on the forehead.

Data collection consists of specific application of the measure – taking temperature with a thermometer for example – and recording the finding.

Performance indicators include different kinds of indicators which reveal different aspects of reality, for different purposes:

- **Input indicators** – measure what conditions are available for successful performance, and might include, for example, funding levels, infrastructure indicators like building space, educational resources like books and materials, teacher training, availability of counseling services, and so on, as well as what learners themselves bring in terms of goals, expectations, and experience.
- **Process indicators** – measure what kinds of organizational and educational processes are present to support performance, and might include, for example, recruitment of hard-to-serve groups, availability of diverse teaching approaches (groups, tutoring, computer-aided instruction), the use of learning contracts, participation by different players in decisions about program design, and so on.
- **Output indicators** – measure the immediate results of the services provided, and might include, for example, customer satisfaction, progress in learning, transition of learners to other education and training programs, certificates of achievement.
- **Outcome indicators** – measure the longer-term impacts (or results) of education on individuals and communities, and might include, for example, employment and wage indicators, children's educational success, civic participation. Credentials, like the GED, whose association with such impacts have been documented, could be surrogate outcome indicators.

Experience from government reform and other fields reinforces the view that a variety of indicators is needed, and that they must be constantly

reviewed (Osborne & Gaebler, 1993, p. 355). Behn points out that all monitoring is done in the short run, but most results we are interested in are long-term. "Consequently, even when outputs and outcomes are closely linked, it is important to monitor a number of different performance indicators" (Behn, 1993, p. 2). Like a CAT scan in medicine, one snapshot does not suffice to give a good representation of reality: we need a series of snapshots, taken from different angles. Better still, like medical PET scanning, we need to develop an ability to show movement over time.

Indicators pick up on expected results, but do not show the unexpected.[16] In addition to routine monitoring of selected indicators, accountability systems need processes that allow the unexpected to be revealed – whether the unexpected is distortion effects of measuring, or positive achievements of learners.

Indicators not only need to be snapshots over time and of different levels, but also have to reflect the multiple, complex realities and viewpoints of different players. Because there are different perspectives on performance, as discussed in Chapter Two, the process of choosing effective indicators involves different players. States that have embarked on developing accountability systems have found that broad participation in design and development produces more effective systems (see Chapter Two). Including learners and teachers in the process incorporates their close knowledge of the reality of the learning environment, allows better alignment between system goals and learners' own goals, and creates indicators that are more likely to be accurately monitored because they are better understood and better supported.

Measuring learning

Because learning is at the heart of adult education, the measurement of learning is central to performance accountability. There are particular issues related to the measurement of learning which must be addressed in any performance accountability system for adult education. In particular, system developers need to evaluate the quality and value of standardized test data, the dominant measurement of learning in the field, and potential alternatives.

[16] Irene Guijt, personal communication, March 1997.

Standardized tests: Standardized tests perform a function that is vital for the adult education system. They enable "learning" to be compared across learners and across programs. Without this capacity, it is difficult to see how the field could be accountable to anyone beyond the individual learner (and in a sense, perhaps not even to individuals, since many will want to know how their skills compare with others). For this reason, in recent years, adult basic education has been committed to using standardized tests as its primary instrument for measuring learning.

The commitment to standardized tests is despite considerable dissatisfaction at both research and program levels. The GAO report on adult basic education notes, "The research literature raises questions about the validity of standardized tests used to measure adult literacy, and local program staff have questioned the appropriateness of using these assessments to measure program results" (GAO, 1995, p. 24). In a review of standardized tests and their uses in adult education, Mislevy raises questions about the adequacy of standardized tests in achieving their main purpose – assessing learning across programs and across populations. He notes that adult education programs vary considerably with respect to the nature and level of skills they emphasize, and the kinds of students with whom they work. Programs use tests for a broad variety of diagnostic, instructional, and evaluative purposes. These tests vary widely with respect to contexts, formats, and mixes of skills they tap. He concludes: "No single score can give a full picture of the range of skills that are important to all the different students in different ABE programs" (Mislevy, 1995, p. ii). For this reason, multiple assessments are needed to provide a fuller picture of student competencies.

Once again, defining what is meant by performance proves crucial: Mislevy points out that "Any assessment starts from an 'operational definition' of competence" (Mislevy, 1995, p. 5). Without that definition, investigations are costly and contested. Venezky suggests that different tests measure different things (Venezky, 1992). The "functional literacy" instruments widely used with adults (such as the NALS) are incompatible with the national surveys of elementary and secondary level literacy based on basic reading, writing and mathematical skills definition, like the National Assessment of Educational Progress (NAEP). Both are incompatible with the grade-level tests of adult basic skills (like TABE).

> The incompatibility between basic skills tests and national literacy surveys places a special burden on policymakers. The NALS results, which will be available within a year, will not be compatible with scores reported by literacy programs that use basic skills tests,

> and while NALS scores could be converted to a crude grade-level equivalency, this would have little validity in that most of the NALS tasks are not taught extensively in school. (Venezky, 1992, p. 4)

There are many problematic issues in the use of standardized tests to measure learning. Some revolve around the context of testing in many programs – lack of consistent time limits for tests, using the wrong level of test so creating floor or ceiling effects, failing to administer the post-test after a fixed number of instructional hours (Van Horn et al., 1996, p. 1.7-1.8).

Other issues concern the content of the test and how closely it is related to instruction. Just as literacy itself is context-related, so also are assessment tools. "Once we recognize that measurement models are at best gross summaries of aspects of students' thinking and problem-solving, we are obliged to identify the contexts that circumscribe their usefulness" (Mislevy, 1995, p. 5). But this creates a new problem: "Programs encounter a trade-off: The better a test focuses specifically on information tailored to its objectives and participants, the less likely it is to overlap the information provided by tests tailored to other programs. Yet such tests are invaluable in evaluating its successes" (Mislevy, 1995, p. 4). The best tests are quite specific as to time, place, program, and population – but these cannot be compared across programs. Any attempt to compare across programs is necessarily less accurate.

There are also serious doubts about whether any of the tests really measure literacy practices in real life terms (as opposed to proficiency in test-taking) – or what they do measure. Functional literacy tests like CASAS and NALS, Venezky says, "are not derived from theoretical models of skill ability ... Considerable research remains to be done on the interaction between text and task difficulty and on the skills that are involved in functional literacy tests" (Venezky, 1992, p. 4). Regie Stites, who re-analyzed NEAEP data to look at learning gains and their retention over time says, "I am not confident that the TABE or CASAS or any other assessment instruments currently in widespread use are adequate measures of the skills, knowledge, and attitudes that enable full engagement in literacy practices and lifelong learning" (Stites, 1996).

Issues of assessment bias further confuse the picture. Venezky points out the difficulty in many functional tests of disentangling reading from math and problem-solving skills (1992, p. 4). Lam questions:

> ...to what extent is the assessment task measuring the same construct and hence has similar meaning for different populations?
>
> For example, ability to read and write is a biasing factor in measuring mathematics skills because it is irrelevant to mathematics skills and it affects Limited English Proficient (LEP) and native English speaking students' performance differently on a mathematics test. (Lam, 1995)

The difficulty is that the field needs the ability both to compare learning across programs and also to demonstrate adequately what has been learned. While standardized tests allow us to compare test results, practitioners are often concerned that they do not demonstrate what has been learned. For example, Burt and Keenan (1995), reviewing learner assessment in ESL:

> The use of commercially available tests with adult learners is problematic because these tools may not adequately assess individual learner strengths and weaknesses especially at the lowest level of literacy skills. Such tests do not necessarily measure what has been learned in class, nor address learner goals. (Burt & Keenan, 1995, p. 1)

They also point out some particular issues in measuring learning of ESOL learners:

> Some testing issues are unique to ESL learners. It is not always clear whether ESL learners have trouble with selected test items because of difficulties with reading, with the vocabulary, or with the cultural notions underlying the test items. Another problem may be that some low-literate ESL learners are unfamiliar with classroom conventions such as test taking. (Burt & Keenan, 1995, p. 1)

The only thing we can be sure we are measuring with standardized tests is the ability to perform on standardized tests. This is not completely worthless – anyone wanting a credential like the GED, in order to enter further education, needs to be able to perform well on tests. But we should not assume that they have much to do with how people engage in literacy in everyday life. Tom Sticht's cautionary tale about armed forces literacy testing should alert us to the dangers of making judgments about how well people will perform in real life settings on the basis of a test, literacy or otherwise.

Performance assessment: In response to the need for assessment tools that more authentically demonstrate learning in real life terms, there has been increased interest in alternative approaches, such as portfolios, demonstrations, narrative and ethnographic approaches to learning evaluation.[17] In particular, performance-based assessment tools have been developed in K-12 education to evaluate how students actually use knowledge and skills.

> There are several reasons for the increased interest in performance-based assessments. .. there is an increased awareness that students can leave school without being able to apply their content knowledge outside of the classroom, partly because tests do not ask students to demonstrate use of their knowledge or skills, but merely to recite them on selected response tests. (Grummon, 1997, p. 4)

Performance assessment has been defined by the Office of Technology Assessment as "testing that requires a student to create an answer or a product that demonstrates his or her knowledge or skills" (Office of Technology Assessment, US Congress, 1992, cited by Rudner & Boston, 1994). As Rudner and Boston point out, designing performance assessment "requires examining the purposes of education, identifying skills we want students to master, and empowering teachers" (1994, p. 2).

Shavelson et al. (1992) conducted research to evaluate the reliability and validity of science performance assessments with 5th and 8th graders. The authors compared hands-on assessment (in which students were observed conducting real investigations) with various surrogates, including computer simulations, notebooks and multiple choice measures (akin to the form of standardized tests). Results of the multiple choice tests were the furthest from those of hands-on assessment of any of the surrogate measures (p. 26). They found that "measures of science achievement are highly sensitive to the method used to measure performance. ... Each method provides different insight into what students know and can do" (p. 26).

However, authentic assessment has limited application in adult basic education because it does not provide policymakers and administrators the opportunity to compare learning across learners and across programs. Without external standards or criteria against which individual learning can

[17] For a discussion of these approaches in adult basic education, see Lytle & Wolfe, 1989.

be judged, alternative assessment will not meet policy needs. Standardized tests do this, and for now, continue to be used in uneasy partnership with various explorations of portfolios and related methods (see, for example, Fingeret, 1993; Literacy South, 1997).

In the maelstrom of confusion about how to measure learning is an opportunity: new research could break through the barriers of an approach that is widely disliked and create new forms of assessment that are firmly based in new understandings of the nature of literacy and cognitive learning. We need systematic ways of documenting achievement and performance that are linked to real learning and to what is taught, and which also provide cross-program comparison.

FOUR. BUILDING FOR THE FUTURE

The discussion so far suggests that developing performance accountability is not just technically challenging, but also value-challenging. The way forward must involve the field in debate and action that address both accountability and performance. To create accountability systems, the adult basic education field needs to:

- define performance – what literacy education should achieve, for individuals, for communities, and for society;
- develop mutual accountability relationships at all levels of the system, from local program to national level;
- build the capacity of the field to be accountable by harnessing existing resources and providing new ones for technical assistance, professional development, support, and information;
- design new accountability technologies to measure performance, report on results, and provide the information tools needed for program improvement.

These are not separate, stand-alone activities: each requires and supports the others. Together the work could not only create accountability, but also result in an ABE system that can meet the multiple performance expectations placed on it, and demonstrate the value of its public investment.

Defining "Performances" – What does success look like?

Chapter Two suggested that accountability systems work only if there is agreement on what should be achieved – what "success" looks like. The concept of literacy lies at the heart of defining success for adult literacy and basic education programs. Traditional definitions of literacy as a bilateral condition (literate/illiterate), or of literacy as a single scale (from illiterate to literate) are challenged by more recent research on literacy in its social context. These should stimulate the field to rethink performance in terms of literacy practices rather than literacy skills, of application and use rather than classroom achievement.

The new theoretical and research work explores "multiliteracies," each linked with different social and cultural contexts. It follows from these new understandings of literacy that there are multiple purposes and uses of literacy and multiple goals and expectations for literacy education. What does this mean for a definition of success?

When reality has many faces, to decide on a single portrait of performance would reduce complexity to a least common denominator that pleases no one. Instead accountability could be approached through a concept of "performances" – multiple purposes and expectations that must be negotiated among multiple players. Adult educators have often claimed that every learner is different. The concept of performances, however, suggests that while there is diversity, there are also commonalities. Learners do not all have the same purposes and expectations, but they share some common perspectives. An overarching accountability system needs to incorporate different concepts of performance, but to be manageable, it must look for commonalities as well as differences.

Several assumptions support the concept of multiple performances for adult basic education:

- Research on literacy practices in everyday life reveal "multiliteracies" and these must underpin definitions of performance for accountability.
- Different learners have varied goals and diverse interests. There is no one kind of learner, no one reason for working on literacy skills, and many definitions of success.
- Multiple performances must be linked in a framework and overarching sense of purpose within which most people can place themselves.
- Negotiating multiple performances requires input from many stakeholders to ensure that the full range of diversity is represented.
- At the same time, lessons from the business world suggest that performances be defined neither too tightly nor too loosely – if they are too loose there is no shared mission; if they are too tight there is a risk of mismatch between system goals and individual goals.

A common framework is needed, within which different performances can be nested. Programs might specialize. Some might concentrate on economic goals, preparing learners for employment. Some might concentrate on family literacy. Some might concentrate on community involvement. All together would address the full range of learners' purposes. These specialized purposes (or definitions of performance) would then be linked with specific performance indicators and measurement which would track their performance separately. If the same variables were to be tracked across the entire learner population even though they are pursuing different goals, then the results would wash out – no gains would show up for anyone.

Creating such a framework not only requires the participation of players at different positions in the system, but also involves regular checking that the impacts are what society wants and needs. Holding programs accountable for training welders, when there is no demand for welders in the job market, is a waste of time and resources. The system of accountability needs constant retuning to match social needs, learner wants, and program capability.

Just as there are multiple performances, so there are multiple levels of accountability – from program to system. The social impacts of literacy education are legitimate concerns of system level accountability. But an individual program cannot be held accountable for the social impacts on individual learners because there are too many intervening variables – the state of the local economy, job market and unemployment rates, the availability of entry level jobs, the demographics of the learner and his/her personal characteristics, and so on. As a whole, the system might be held accountable for employment outcomes only if employment purposes have been clearly articulated in policy and appropriate resources have been provided at the program level to support the work.

System level accountability and program level accountability are different but must be synchronized, or both will fail to meet their purposes. If system accountability is concerned with employment, for example, and programs do nothing to focus on employment outcomes, then they may not get the results that funders desire. The overall vision, purposes and practice of adult education need to be part of a common framework that links “big picture” system goals and “small picture” classroom activity. Unless they fit together, the system cannot meet its goals.

Some efforts are already underway to define performance frameworks at the national level. The National Outcomes Reporting System is being developed by the US Department of Education. It is defining a common set of ‘outcomes’ for adult basic education as the basis for data collection and reporting. State ABE directors have been closely involved in the development. However, the process has so far had limited input from other stakeholders.

The Equipped for the Future project of the National Institute for Literacy has been referred to already in this report. This broad-based system reform effort has actively sought input from a wide range of stakeholders. It has developed and validated four purposes for adult learning and a set of "role maps" for each of the adult roles of citizen, worker, and

parent/family member. These role maps lay out what adults in these roles do and what they need to know for effective role performance. With 25 development partners in 17 states, EFF is now testing draft content standards. Performance standards will be based on the final version of these. When fully articulated and validated, these could provide a framework for performance as the basis for an accountability system.

It matters who participates in setting the vision and goals for the field. If learners are not part of the decision process, and their own goals are not the same as the system goals, then either they will not enroll or will drop out of programs, or the programs will not show successful results. It is crucial that system goals are aligned with student goals and for everyone to be clear about what they are. Then teachers and programs can work on meeting the goals, and success is possible.

Developing mutual accountability relationships: Who is accountable to whom?

It was argued in Chapter Two that traditional approaches to quality control derived from "Taylorist" manufacturing are not very useful for adult education. Quality control checks on results at the end of the “production line” may highlight problems, but only when it is too late. Newer business ideas about quality emphasize continuous improvement and learning, responsiveness to internal and external customers, and participation in decisions and mutual responsibility at all levels of the organization.

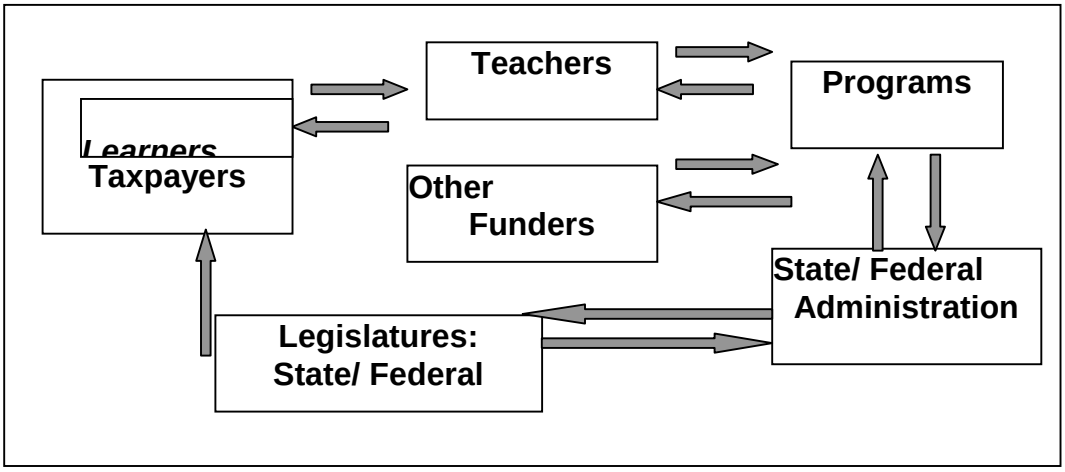
Underpinning these “high performance” or learning organizations is a system that includes everyone in mutual accountability relationships. Accountability is no longer the simple matter of an inspector checking the widgets at the end of the conveyor belt to see if they meet specifications. Mutual accountability engages members of the organization in creating a common vision, determining goals and customer expectations, and designing effective means of monitoring as well as producing in order to learn how to do it better.

In a mutual accountability system, every “player” would be both accountable to other players and held accountable by them (see Box 4.1). Teachers, for example, would be held accountable by learners for providing learning opportunities that meet their needs. Teachers would hold program directors and funders accountable for providing the resources they need to meet learner needs – which might include materials, space, training, pay for lesson planning and assessment.

Spelling out relationships of mutual accountability would reveal some which are overlooked in conventional accountability systems. Congress, for example, would hold adult education programs accountable for providing effective and efficient services. But Congress would also be held accountable by programs, by learners, and by voters for identifying a social need, passing appropriate guiding legislation, and providing the resources needed to create a strong adult education system.

Learners would hold their teachers accountable. But programs would also hold learners accountable for taking learning seriously, for making an effort to participate fully. Businesses that expect adult education to provide them with workers equipped with basic skills might be expected in turn to provide jobs for those workers, or to continue a workplace basic skills program when the grant runs out. Right now, many workplace educators are frustrated because they establish a partnership with business and often, partway through the project, the business partner disappears – they downsize, or are taken over, or corporate office requires them to switch attention elsewhere. Mutual accountability would require all the partners to honor their "contract."

Box 4.1 Some mutual accountability relationships



An accountability system based in the concept of mutuality has several characteristics:

- it is negotiated between the stakeholders in a process that engages all the players in clarifying expectations, designing indicators of success, negotiating information flows, and building capacity;

- each responsibility is matched with an equal, enabling right;
- every player knows clearly and agrees to what is expected of them;

- every player has the capacity to hold others accountable;
- efficient and effective information flows enable all players to hold and be held accountable and act to improve services.

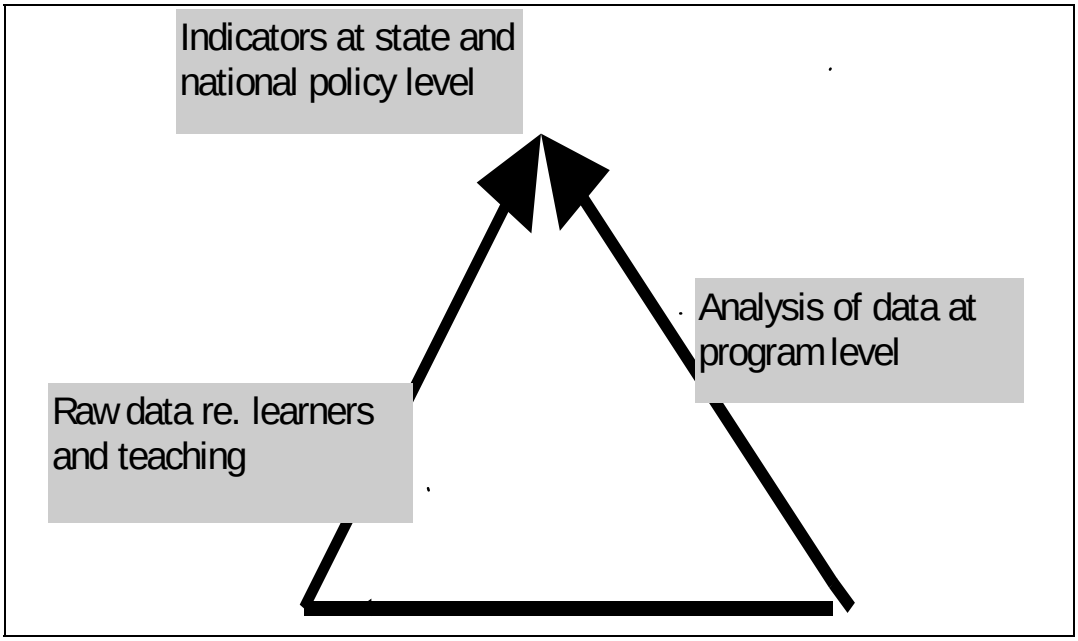
A number of initiatives at state and national levels are working to bring stakeholder groups into closer relationships, the first step in mutual accountability. The National Summit proposed by NCSALL in partnership with NIFL and the Department of Education is an example of the kind of work to be done at the national level to ensure stakeholder involvement in deliberations about the future of ABE. There are now many examples of adult learner organizations at the state and local level that are working to address the inequalities in power, access to information, and voice of adult learners in the system. A new national learner organization, VALUE (Voices for Adult Literacy United for Education), has recently been formed and shows promise for building a strong and effective learner voice. Other stakeholders often are represented in state adult education councils, including business and community organizations.

Inequalities of power and influence affect the capacity of the system. How do learners become real stakeholders? Simply providing learners with information does not necessarily engage them in mutual accountability if in fact their power is very limited and their ability to effect change is confined to dropping out – understanding and communication do not eliminate power differentials. Learners become structurally part of the system of accountability when they have real power to make choices (and not just to vote with their feet). Some community based programs are beginning learner participation in management, with learner representatives sitting on boards and being involved in management decisions about the program. That is still unusual in ABE. Learning organizations in adult education have been even slower to develop, but are being addressed by, among others, Laubach Literacy Action in its own organization and in management training with local programs.

Perhaps a central issue in mutual accountability is the question of information. Without adequate access to information, stakeholders cannot hold others accountable. To create mutual accountability we have to change the map of information flows that has been traditional in the field. The usual flow of information in adult education can be portrayed as a pyramid (see Box 4.2). Information is collected at the base and increasingly summarized for the purposes of different levels on the way up (from program to community, state, and national levels). In the pyramid, almost all information is collected at the classroom level. This raw data is about individual learners (their goals, existing skill levels, learning gains), and about the teaching-learning process (attendance rates, curriculum,

assessment). While teachers could use such data in designing teaching, it may not be readily available to them. Information flows only one way: up the system to the state and national levels.

BOX 4.2 Pyramid data systems



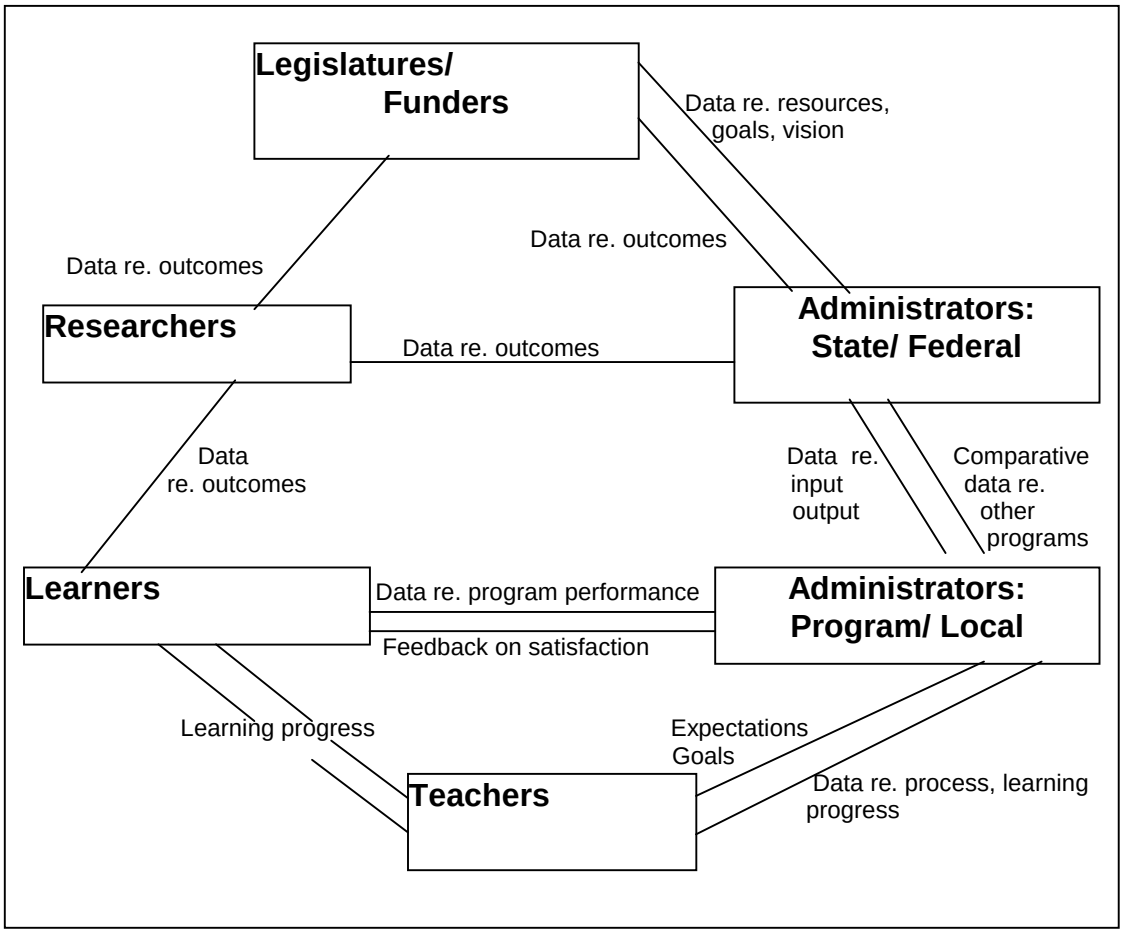
In this model of data collection, raw data is processed and synthesized in various ways above the classroom level. A short way up the pyramid, local program administrators need certain kinds of analysis of the raw data in order to plan the use of resources, order materials, evaluate teachers, and assess whether the program is meeting its targets. Further up the pyramid, state or regional administrators require a different kind of synthesis of the raw data to answer their questions about effectiveness, efficiency, and productivity of local programs. At the national level, policymakers require a more condensed set of information, a set of indicators which may focus on both program quality and interpretations of impact.

This simplistic model may not fit the needs of a performance accountability system that takes into account different performances and

purposes. Mutual accountability in a learning organization demands a more complex flow of information depending on the responsibilities of players. An alternative model would suggest information flowing around

the system, up, down and across it, among and between different players. A simple version of this is suggested in Box 4.3.

Box 4.3. Information flows for mutual accountability



For accountability to work, information must flow up, down, and across the system. Right now, many of the needed communication links are not in place, or poorly constructed. The information needs of different players in the system can be met in different ways. Understanding the purpose for which information is needed is essential, so data users and data providers need to communicate. The capacity of the system to collect, analyze and use data must be increased.

Building capacity for accountability: Systems that learn

A system that supports continuous improvement is a learning system. The process of learning creates capacity at all levels, but to engage in that system at all, capacity has to be built. Perhaps this sounds contradictory – participation builds capacity, but capacity must be built in order to

participate. It fits with what we understand about adult learning, however: individuals cannot learn when they are hungry, when basic survival demands all their attention. Organizations cannot learn when basic resources are denied, or when power relationships create significant inequalities, or barriers prevent application of what is learned. So capacity is required to engage in learning at all, and engaging in learning builds capacity further. If one of the objectives of performance accountability is to improve performance, then there has to be investment in the capacity of local programs to use data to monitor how well they are doing and to improve services.

It is difficult to have a Management Information System when there is no management system. Building a national accountability system requires that we build the capacity of the adult education system at all levels to collect, analyze, and use information for accountability and improvement. Learners need to have information about the programs in which they enroll to be sure their goals will be met. Teachers need resources and tools to be informed about learner goals. Local program managers need data to use in program planning. State program administrators need reliable data about student participation and progress.

Accountability demands reliable and accurate data, so capacity must be built within the system to collect, interpret, and use data. One way to ensure reliable data is for those who collect the data to have a need to use them. If they collect information that is not for their use they need to understand its value. Building capacity to collect, analyze and use information requires staff development, technology, and technical support.

It also requires that we address the workforce and the resources provided to adult education. In most states, teachers are part-time, low-paid, with no job security, and no union (see Condelli, 1994; Young et al., 1995). They do not always take part in program design and management or even in broad vision-setting. For teachers to become real stakeholders, they must have a real stake in the success of their program, and this means a greater role in decision-making. The contingent workforce that presently is the norm in many states cannot be effective stakeholders.

Many states have been working to build capacity for program delivery – the capacity to perform. The experiences of states like Arkansas, Connecticut and Pennsylvania (reviewed in Chapter 3) – as well as other states like Massachusetts, California, and Iowa – all demonstrate the importance of long-term, systematic investment in training, information systems, and technical support. Many of the same actions are needed to

build capacity for accountability – staff training and support, use of information for continuous improvement, appropriate resources.

Rewards for improved performance can be part of an accountability system. Performance-based funding has not yet been developed in most states (Arkansas is an exception), for only when capacity has been built to achieve expected performance will it be effective to tie funding to it.[18] When "performance" is not agreed, lines of accountability are shadowy, and resources at the program level are minimal, performance data will be shaky and unreliable, and performance funding misguided.

However, as a step in the right direction, rewards for collecting and analyzing information can be provided along with encouragement to use it for improving program services. In Pennsylvania for example, a new block of funding is tied to an agreement from programs to do rigorous data collection and analysis. Program managers and teachers will participate in professional development around using data for decision-making (see Chapter 3).

Capacity to be accountable means that resources have to be commensurate with accountability expectations. In a posting to the NLA list-serve recently, David Rosen proposed a sliding scale of resources matched to levels of accountability:

> A program which is funded at $200 per student per year should be held accountable to keep records of the students served, and how the money was spent...at this level of funding, whatever else we ask for may jeopardize the program's ability to succeed.
>
> A program funded at $1,000 per student per year should have a higher intensity and duration of services, a good retention rate, a solid plan for staff and program development, and should meet many other indicators of program quality. It should be held accountable for these. The goals of the program and students should match. And the program should have some evidence of student gain/progress/ accomplishment in attaining these goals.
>
> At $5,000 per student per year or more, I would expect to see learning gains measured with valid and reliable instruments (not necessarily or only standardized tests). At this level we could expect to see some research on these programs' outcomes and impact on learners' and their families' lives...we could expect not only individual

[18] Arkansas is one of the few states to tie performance to funding (see Chapter 3). Arkansas puts substantial state funding into adult education, most teachers and programs are full-time, so the state has more capacity for performance funding.

> program accountability but system accountability, and evidence that investment was succeeding. (Rosen, 1997)

The capacity to perform requires:
- adequate funding for staff, programs, materials;
- staff development;
- appropriate curricula and materials;
- technical support;
- using performance data for improvement.

Building capacity for accountability requires:
- accountability demands that match resources provided;
- staff training and technical assistance on data collection, analysis and use;
- information technology to support the collection, analysis and use of performance data;
- engagement of data users in designing useful measurement tools;
- rewards for improved performance.

Designing accountability technologies – Learning to measure

Adult educators all use measurement – to report to funders how many people participated in programs, to test learners to see how well they read, to check attendance in classes, to identify the extent of literacy "need" in the community. Performance accountability requires that the numbers collected are useful and used.

Adult educators are asked by different stakeholders for different kinds of data – funders may ask how many learners persist in programs for more than a few hours; learners may ask how long will it take to get a GED; employers may ask what will be the effect of an in-house education program on company productivity. The tools to answer these questions well are not readily available. Too often, adult educators have to say "It depends." Persistence depends on other problems in learners' lives. Educational outcomes depend on initial reading level, whether a learner has learning difficulties, how long they spend in class, how much homework they are prepared to do. Bottom-line payoff of workplace literacy depends on the organization of the workplace, skill demands of the job, cultural and social context within and outside the firm. Programs do not have the data to spell out the effects of such variables.

Performance accountability commonly uses indicators as a measurement tool: these must be both relevant and important. An indicator that

measures something unrelated to the problem – the percentage of brown-eyed adult learners, for example – is irrelevant. An indicator that measures something relevant – the skills and knowledge which a learner brings when they enter a program, for example – but in a very inadequate way is dangerous. And an inability to measure something important – affective changes in learners, for example – can be disastrous.

Indicators, usually numeric, serve as a useful way of capturing and communicating complex trends. They are a tool for policy-making, but also for all who want to use information in a systematic way to create change. Mutual accountability relationships shift the ways we think about and use measurement tools. Knowing who will use the information and for what purposes is at the core of the technical work of designing valid and reliable measures.

Indicators are an approximation of reality, not reality itself; tools, not solutions. They can be good, bad, or indifferent. A number of checklists assess the value of indicators such as SMART – specific, measurable, action-oriented, realistic, time-framed (see Abbot & Guijt, in press). In the context of adult education, this acronym suggests that:

- **Specific** indicators must be directly linked to the purposes and goals for learning, and reflect the different contexts and goals of learners, programs and communities. Example: if a group of adults enroll in a basic skills program because they want to pass the commercial drivers license test, then an indicator would be how many passed it, in what time frame. This would not preclude these individuals developing other learning goals along the way.
- **Measurable** means what can be documented in a systematic way, although not necessarily quantitative, as qualitative indicators also exist. Not everything that is important and valid is easy to measure, and vice versa: what is easy to measure is not always important. Performance accountability requires us to strive toward a synthesis of value and documentation. Example: if lifelong learning is important, then an indicator could be how many learners report that they have engaged in formal and informal learning opportunities in a series of follow-up surveys.
- **Action-oriented** suggests that indicators should focus on improvement of services to learners. To do this, they have to reflect something that can be controlled – if practitioners have no means of controlling something, it cannot be a useful indicator for action. Example: if learner satisfaction is important in keeping learners enrolled, then an indicator could be a customer satisfaction survey conducted at regular intervals – asking how learners define learning

objectives and evaluate their progress, as well as specific feedback about teaching, materials, facilities and so on – on the basis of which program changes might be made.

- **Realistic** means manageable – the indicators must be understandable, accessible, and above all do-able within the resources available. Example: if vocational or job training is important, then an indicator could be how many learners enroll in such programs within a specified timeframe, using the databases of other institutions cross-referenced with the ABE database.
- **Timely** means that the data are available at the time that action is needed. Elaborate research projects whose findings appear after four or five years are no good to program managers looking to act on program improvement, although they may help longer-term visioning and evaluation at the system level. Example: if it is important that learners stay in the program long enough to learn, program managers could use attendance records each month to plot retention, looking for patterns in particular classes or groups of students. Accumulated data would show "normal" attrition rates and unexpected variations.

These examples draw on a variety of types of data collection methods, including customer satisfaction surveys, follow-up surveys after leaving the program, data from other databases, documentation of credentials, and internal program monitoring.

A complete performance measurement system would include different approaches (called here technologies) to data collection and analysis that meet varying accountability purposes and distinct notions of performance. Research, monitoring, and evaluation can be seen as different accountability technologies that gather different kinds of data to answer particular questions over specific timeframes.

Research tries to answer important questions about associations, correlations and meanings. It can help with some of the kinds of "big" questions policymakers often have. Does adult education impact people's lives? What are the benefits to individuals and society? What policy initiatives are needed, and what levels of resources should be provided? What long-term or short-term outcomes are associated with particular program designs? What kinds of resources are needed to support specific program designs? While we often think of research as being conducted by professional researchers – in a university or consulting institution – teachers and program managers may conduct their own research to answer questions about teaching approaches, learner perspectives, community needs. Increasingly, action research is

finding a place in classrooms and programs as a tool for local investigation (see, for example, Quigley & Kuhne, 1997).

Evaluation can answer questions about "how are we doing?" State program administrators often have such questions about local programs. Are programs meeting their objectives? Do they meet program quality standards? How are learners being served? Are students making appropriate progress? Evaluation can answer questions like: what happened, how did people in different relationships perceive things, what were the problems and barriers as well as success and achievements? Evaluation can be a useful tool for program improvement at all levels – local, state and national. It can document program services and other system components, including professional development, technical support, monitoring, planning. Evaluation may be ongoing, but it is not routine: it seeks answers to particular questions at particular points of time. Again, while often conducted by outside evaluators, there is a substantial body of experience in other fields, and some in adult education, of practitioner involvement in evaluation (Feuerstein, 1986; Greene, 1994; Rugh, 1986).

Monitoring can answer ongoing questions about day-to-day operations. Teachers and local program directors may pose questions like: Are we recruiting the kinds of students we intended to? What do these students say they want from their learning experience? Do students stay long enough to experience learning? How satisfied are they with the program and classes? State and national administrators may use monitoring to look at enrollment targets, program quality indicators, learning gains, customer satisfaction. Monitoring data are usually guided by pre-determined indicators. Some monitoring data may be aggregated and passed to the state and national levels, some would only be used at the program level. Monitoring data may also be collected at the state level – on technical assistance and professional development, for example.

Each of these accountability technologies shines a flashlight from a different angle to illuminate different aspects of reality. They provide triangulation, enabling us to cross-check findings from one set of data with another. From a system management perspective, wise use of resources

requires that the different strengths of research, evaluation and monitoring be used selectively. Routine monitoring at the program level can be extremely expensive and resource intensive: with literacy programs already resource-poor, it makes sense to select carefully and economically the data that they must collect. Not all data need to be collected on every student, and some can be collected on sample

populations instead. Properly selected, such samples can provide valid and reliable data. California, with a very large ABE/ESL population already reports sample data rather than data on the entire universe.[19]

Some data are not easily collected at the local level. Follow-up data – for example on learner outcomes – may require tracking of individuals over periods of time after they have left the program. When important measures cannot be tracked through monitoring, research is needed. Meta-data, that is, data about data, can be used to identify valuable follow-up data about the sample population that are already collected and housed in other databases, like employment records, drivers license records, further education institution records. Once identified, some of these data sources can become part of the ongoing monitoring system, or may remain part of research.

In separating the different domains of research, evaluation and monitoring, it is not intended to privilege any one – all are needed. Nor should research be seen as the domain of "researchers" alone. Research for action, conducted by practitioners, is a valuable tool in the learning organization. Professional researchers and practitioners often have different interests in measurement – the one to examine problems, the other to construct solutions. Data from outside research may not readily translate into data for action at the program level. Researchers are often slow at publishing reports, limiting the potential for use of research data for program improvement, which depends on timely response. For these reasons, there is an important role for practitioner action research, as well as monitoring, evaluation, and system level research.

Chapter 3 outlines some of the issues of concern about current measurement tools, especially of learning. It is clear that new approaches are needed. Performance assessment approaches, used in K-12 education and in other countries like Britain and Australia, hold some promise. These should be research-based and validated, and to be effective should be linked with a set of external standards or criteria against which learning can be assessed.

The NIFL Equipped for the Future project is in the process of developing standards for adult education. Content standards set out what learners should know and be able to do. Performance standards enable judgments to be made on whether an individual meets the standard. In K-12

[19] Interview with Pat Rickard, CASAS, April 1997.

standards-setting work, performance standards commonly have three parts: a succinct description of what the individual should know and be able to do (the content standard), samples of student work to create an image of the kind of work that meets the standard, and commentaries on the work to explain what features raise them to the standards (Tucker and Codding, 1998). Only when these standards are agreed can the work of creating new approaches to assessment of learning be completed.

Next steps: Ways forward

Performance accountability should enable the adult education system to be flexible and responsive to changing social and individual needs, with a clear, common purpose. Designing it should engage all stakeholders, because only with that participation are all parties in the mutual relationships informed, engaged, and equipped for their roles. To get from here to there will not happen quickly or simply, but it is vital to begin. The process of designing performance accountability for adult basic education can use three approaches to bring about change – public discussion, action, and research.

Public discussion is at the heart of what Harry Boyte, long-time community activist and political scientist, calls "citizenship as public work" – "citizenship as effective, skilled, public-spirited work in solving our common problems." (1995, p. 1). The creation of public values comes about through ongoing efforts of people with different interests and views to address common concerns. This does not gloss over questions of unequal power or conflicting values. The common concerns of the many and varied stakeholders in adult education need to be arrived at through joint discussion and effort.

Action is required to try out ideas, learn from them, and amend them for future trial. It is a crucial step in recent thinking about education reform. Michael Fullan, for example, in his work on reforming public education, suggests the formula: "ready, fire, aim" (rather than "ready, aim, fire") (Fullan, 1993). In education circles, it is common for so much time to be

taken up in discussion about how to aim, that action (firing) never happens. It is important not just to talk, but to try things out, and in the trying, to learn how to aim better next time.

Research is needed both to learn systematically from experience and also to collect new data, analyze, validate and evaluate it. The research needed is of different kinds, including practitioner action research as well

as more traditional approaches. It can help us identify what we know, and what we don't know, so that we can move forward.

These approaches to change can be applied to the four major task areas that have been reviewed in this chapter: agreeing on common definitions of performances, negotiating relationships of mutual accountability, building capacity for accountability, and developing and using accountability technologies. These "steps" or areas of work need to be overlapping and interconnected. Each step will require all three change approaches, although the weight given to each might vary.

Step 1. Agree on performances: The key first step is the development of agreement about vision and purpose for adult education, in which the multiple roles and expectations of all key players are consulted and acknowledged. Accountability systems that do not clearly define "performances" cannot be effective. Many different players must be involved in defining vision, purposes, and goals. It is a question of values not of technical solutions. The diversity of voices involved in the goal-setting exercise will be a measure of how far the goals can speak to the whole field.

Broad consensus is needed on a new framework that allows for diversity within a clear and agreed common vision. Lessons from experience in education and other fields suggest:

- Asking the question of what performance means. This cannot be skipped over or rushed. Without knowing what is important, measurement becomes an exercise in "gaming the numbers" to satisfy external demands, often with perverse results.
- Involve stakeholders and seek consensus. Without broad public debate it is difficult to frame performance goals that reflect the "big tent."
- Reflect newer understandings of literacy and connect performance with real life. This is an opportunity for literacy research to connect with and support practice.
- Acknowledge multiple performances. Too narrow or tight a definition of goals will exclude learners and programs outside or force them to falsify their data.

Next steps could include:

⇒ **Public discussion:** A variety of specific activities could be undertaken that involve the full range of stakeholders in discussion on

a broad vision and definition of performances. These might include consultation (like the initial Equipped for the Future consultation with adult learners), working groups at program, state and national levels, conferences and meetings. The activities would highlight disagreements, lack of a common vision, as a first step toward consensus-building and common ground.

⇒ **Action:** Agreed performances and their associated indicators and measures need to be developed and piloted at local and state levels. Again, the involvement of and validation by diverse stakeholders is required for these to be accepted and acted-upon. The Equipped for the Future project is already developing a common framework of purpose, roles and standards, which could become the framework for the field as a whole.

⇒ **Research:** The performances identified through public discussion have to be validated with all stakeholders – students, practitioners, community leaders, policymakers, business leaders, and other agencies. Other research contributions might include qualitative studies documenting the working assumptions about performances currently held (what goals do learners bring? practitioners hold? what submerged assumptions are guiding teaching?). New Literacy Studies researchers could develop practice-linked research to bring new concepts of literacies into operation at the level of teaching, management, and system development.

Step 2. Develop mutual accountability relationships: Mutual accountability depends on the various partners knowing, understanding and accepting their roles and obligations to one another and to the system. These relationships have to be negotiated, and require discussion and consent. Formal contracts might be appropriate in some cases. Information needs to flow freely in all directions. Inequalities of power inherent in any system need to be explicitly addressed. Lessons from experience in education and other fields suggest:

- Bring the full range of stakeholder groups into the process, including teachers and learners who often have not been 'at the table.'
- Provide support for stakeholders who have least access to information and power, like adult learner organizations at national and state levels.
- Construct information channels among and between all stakeholders.
- Develop learning organizations at the program and state levels which would emphasize learning and continuous improvement, shared responsibility, and engagement in monitoring results.

Activities that could support next steps on mutual accountability include:

⇒ **Public discussion:** Discussions among different players in the field could focus on the concept of mutual accountability, what the relationships are or should be, who should be held accountable for what and by whom, what kinds of formal or informal contractual relationships may be appropriate.

⇒ **Action:** A pilot state or program, perhaps one that already has performance accountability work underway, could take on the task of developing written contracts and trying them out as part of programs. Action projects are needed to identify and test the information flows that are crucial to mutual accountability, highlighting specific information needed by each player in the field.

⇒ **Research:** Research should document the pilot action projects. Communication audits would document existing information flows and use, as a basis for change. Case studies could examine more closely the questions of power which affect players' ability to engage in mutual accountability. Action research would clarify the costs and resources associated with developing mutual accountability.

Step 3. Build capacity for learning and system improvement:
Building the capacity of players at all levels of the system to collect and evaluate data and use them to plan work requires serious attention and targeted resources. We are not starting from scratch: there is already considerable experience with program improvement efforts in a number of states – through program quality indicators, teacher inquiry, action research at the program level, and staff development (see for example the work of Lytle and others in Smith & Lytle, 1993; the special issues of *Focus on Basics*: Garner, 1997; and Quigley & Kuhne, 1997). Increasing experience with participatory approaches to the development of indicators and performance measurement is accumulating at the international level (see, for example, Guijt & Sidersky, 1996; World Bank, 1997).

Reframing the entire adult education system as a learning system can build on existing efforts to support innovation. Experiences suggest:

- Key elements of building the capacity to perform include
    - ∗ increased resources,
    - ∗ focusing on quality,
    - ∗ staff development and training,

  - technical support,
  - use of performance data for continuous improvement.
- Key elements of building the capacity to be accountable include
  - accountability demands match resources,
  - user involvement in developing better measurement tools,
  - staff training and support,
  - timely information loops,
  - rewards for improved performance.

Next steps might include:

⇒ **Public discussion:** Issues for discussion would include the matching of resources with accountability, concept of learning organizations, how to use information for continuous improvement. Arenas should cover the full range of stakeholders and levels, from learners and practitioners at the local level to state and national policymakers and other stakeholders outside the system.

⇒ **Action:** States or programs might pilot approaches to collecting, analyzing, and using information for improvement and develop staff training, technical assistance and support. Pilot learning organization projects could be established at local and even state levels.

⇒ **Research:** Studies could identify where capacity needs to be built and what specific kinds of information (and in what form) seem to help program improvement the most. All current data collection and reporting must be reviewed with a critical eye to assess whether it is useful and useable for the various purposes of different stakeholders. Research is needed to design timely and effective information feedback loops so that it can be used for program improvement.

Step 4. Design and develop accountability technologies: Once performances have been negotiated and agreed, there will be pressing need for new ways of measuring them. Two connected areas of work here are to design new and useful indicators and measures and to develop appropriate methods to collect and analyze data. Many existing performance measures, especially for learning gains, are disliked by practitioners and researchers alike. There is a pressing need for new ways to assess learners' performance in terms of literacy practices rather than the indirect approaches of standardized tests which 'stand for' real-life practices, usually inadequately. Existing methods of data collection have also proved themselves inadequate, and new approaches need to

be developed that take into account resources and capacity. Experience suggests:

- External standards need to be developed, as criteria against which individual student learning can be measured, and through which program performance can be assessed.
- Performance assessment tools for measuring learning need to be designed. Initiatives in performance assessment in countries such as Britain and Australia may provide useful models for measuring and assessing learning.
- The full range of potential of research, evaluation, and monitoring technologies needs to be utilized to meet the needs of different stakeholders. These approaches to gathering, analyzing, and using information are based on different kinds of data and meet different purposes. Using them in appropriate ways, adult education can develop a dynamic system of information, analysis, and reporting.

Next steps could include:

⇒ **Public discussion:** Stakeholders should be involved in evaluating current measurement methods. Public discussion can begin to identify information needed. External standards would need to be developed and validated. On this basis, performance indicators and new measures can be developed.

⇒ **Action:** Collaborative approaches to indicator development could be explored in pilot sites. States could experiment with different accountability technologies – e.g. using sample data in monitoring, research surveys to track outcomes over time, identifying other databases to use for outcomes documentation. Practitioners and learners should be involved in designing the measures to be used in monitoring, evaluation, and research at the local level to make sure that they are useful and useable. The process of designing measures will be an iterative one, going back and forth between local, state, and national levels, and involving a variety of stakeholders as well as researchers.

⇒ **Research:** New methodologies and measures for performance need to be designed and validated, along with testing/validating what existing tests actually measure. Research funds could be allocated to identifying and validating assessment tools, particularly looking for alternative approaches to assessing performance. The different contents and approaches of the three broad accountability

technologies (research, evaluation, and monitoring) should be defined and further elaborated.

BOX 4.4 Summary of next steps

	Public Discussion	Action	Research
1. Agree on performances	Identify different purposes, visions, definitions of performance	Develop an agreed framework for performances, on which standards and indicators can be based	Validate performances with all stakeholders; link New Literacy Studies research with practice
2. Negotiate mutual accountability relationships	Identify contractual relationships, what each should be held accountable for, and by whom	Pilot development of contracts, information flows needed for mutual accountability; provide support for learner organizations	Communication audits of information flows; identify power differentials
3. Build capacity to learn and improve	Discuss how resources and accountability should be matched, and how to use information for improvement	Pilot continuous improvement projects; design training and support; construct information loops	Identify capacity issues, information needed; evaluate all data collection for usability and usefulness
4. Develop and use accountability technologies	Identify indicators and measures, link with appropriate technologies	Collaborative pilots for indicator development, test technologies	Design new methods and measures, validate assessment tools

Conclusions

New research and theory about literacies rooted in social and cultural contexts suggests the concept of performances as a useful one on which to build accountability systems. There is little agreement on what constitutes performance in adult basic education, because there are many "literacies," in different contexts, and diverse purposes for literacy learning. But some common ground and agreement must be negotiated if adult education is to be held accountable for results.

Performance accountability should be viewed as having multiple dimensions – many crosscutting lines of accountability, and different views of performances. In order to be accountable, adult education needs to recognize and become clear about its multifaceted goals, develop mutual accountability relationships between the various constituencies who have a stake in adult learning, and develop its capacity to learn from experience, to improve its practices, and to be accountable.

The issue of capacity is a thorny one that must be uppermost in the minds of policymakers as they advance accountability initiatives. The need to build the capacity of the field both to get results and to be accountable for performance is widely recognized. Adult basic education has done a lot with a little. Its capacity for performance accountability needs to be deliberately strengthened, and may require new resources, as well as coordinating and harnessing existing resources (especially technical assistance, research, and professional development) to a common agenda.

There are no quick answers. The principles for action which have been outlined here require consultation with the field and with stakeholders. To move forward will need meetings and taskforces, and it will take time. It requires the field to learn lessons from elsewhere when appropriate, build on current initiatives when they are underway, and create new tools when none exist. Policymakers have the capacity to set the stage, harness resources, and create a common agenda. Commitment to high performance requires the contributions of many players. The goal is nothing less than developing a learning system and culture, so that high performance is agreed, expected, valued, and achieved.

REFERENCES

Abbot, J., & Guijt, I. (in press). *Changing views on change: A working paper on participatory monitoring of the environment.* London: International Institute for Environment and Development.

Alamprese, J. A. (1993). *Systematizing adult education: Final evaluation report of the Connecticut Adult Performance program (CAPP)*. Washington, DC: COSMOS Corporation.

Auerbach, E. R. (1989, May). Toward a social-contextual approach to family literacy. *Harvard Educational Review*, 59 (2), 165-181.

Auerbach, E., & McGrail, L. (1991). Rosa's challenge: Connecting classroom and community contexts. In S. Benesch (Ed.), *ESL in America: Myths & possibilities* (pp. 96-110). Portsmouth, NH: Boynton/Cook/Heinemann.

Barton, D. (1994a). *Literacy: An introduction to the ecology of written language.* Oxford: Blackwell.

Barton, D. (1994b). Preface: Literacy events and literacy practices. In M. Hamilton, D. Barton, & R. Ivanic, *Worlds of literacy* (pp. vii-x). Clevedon, Avon: Multilingual Matters, and Toronto: Ontario Institute for Studies in Education.

Barton, D., & Hamilton, M. (1998). *Local literacies: A study of reading & writing in one community.* London: Routledge.

Beder, H. (1991). *Adult literacy: Issues for policy & practice.* Malabar, FL: Krieger Publishing Company.

Beder, H. (in press). *The outcomes and impacts of adult literacy education in the United States*. Boston, MA: NCSALL.

Behn, R. D. (1993). The growing use of performance indicators in state government. In R. D. Behn et al., *Performance indicators for state government: A report to the nation's governors and their chiefs of staff* (pp. 1-22). Durham, NC: The Governors Center at Duke University.

Boyte, H. (1995). *Beyond deliberation: Citizenship as public work*. Paper presented at the Conference on Citizen Competence and the Design of Democratic Institutions, Committee on the Political Economy of the Good Society, February 10-11.

Bradshaw, D. (1996). Putting on a performance. *Open Letter: Australian Journal for Adult Literacy Research and Practice*, *6* (2), 44-58.

Brizius, J. A., & Campbell, M. D. (1991). *Getting results: A guide for government accountability.* Washington, DC: Council of Governors' Policy Advisors.

Burt, M., & Keenan, F. (1995, September). *Adult ESL learner assessment: Purposes and tools*. ERIC Digest (EDO-LE-95-08).

Bush, B. (1989, October). Literacy: Our shared goal. *The Reading Teacher*, 10-12.

Caulkin, S. (1997, December 1). Stampede to replace the principle of profit. *Observer* (UK).

Center for Applied Linguistics (1983). *From the classroom to the workplace: Teaching ESL to adults*. Washington, DC: Author.

Center for Applied Linguistics (1989). *Basic English skills test (BEST).* Washington, DC: Author.

Chisman, F. P. (1989). *Jump start: The federal role in adult literacy.* Final report of the project on adult literacy. Southport, CT: Southport Institute for Policy Analysis.

Condelli, L. (1994). *Implementing the enhanced evaluation model: Lessons learned from the pilot test.* Prepared for the U.S. Department of Education, Division of Adult Education and Literacy. Washington, DC: Pelavin Research Institute.

Condelli, L. (1996). *Evaluation systems in the adult education program: The role of quality indicators.* Prepared for the U.S. Department of Education, Division of Adult Education and Literacy. Washington, DC: Pelavin Research Institute.

Condelli, L., & Kutner, M. (1992). *Quality indicators for adult education programs: Lessons learned from other programs,* Prepared for Office of Vocational and Adult Education, U.S. Department of Education. Washington, DC: Pelavin Associates.

Condelli, L., & Kutner, M. (1997). *Developing a national outcome reporting system for the adult education program*. Report for the U.S. Department of Education, Office of Vocational and Adult Education, Division of Adult Education and Literacy. Washington, DC: Pelavin Research Institute.

Cook, W. D. (1977). *Adult literacy education in the United States.* Newark, DE: International Reading Association.

de Castell, S., Luke, A., & Egan, K. (1986). *Literacy, society, and schooling: A reader.* Cambridge: Cambridge University Press.

Feuerstein, M.-T. (1986). *Partners in evaluation: Evaluating development and community programmes with participants.* London: Macmillan Educator Ltd.

Fingeret, H. (1984). *Adult literacy education: Current & future directions.* Columbus, OH: ERIC Clearinghouse on Adult, Career, and Vocational Education. Information Series No. 284.

Fingeret, H. (1992). *Adult literacy education: Current and future directions. An update.* ERIC Clearinghouse on Adult, Career, & Vocational Education. Information Series No. 355.

Fingeret, H. (1993). *It belongs to me: A guide to portfolio assessment in adult education programs.* Durham, NC: Literacy South.

Fingeret, H., & Jurmo, Paul (1989). (Eds.). Participatory literacy education. *New Directions in Adult Education: No. 42*. San Francisco: Jossey-Bass.

Fitzgerald, N., & Young, M. (1997). The influence of persistence on literacy learning in adult education. *Adult Education Quarterly*, *47* (2), 78-91.

Freire, P., & Macedo, D. (1987). *Literacy: Reading the word and the world.* South Hadley, MA: Bergin and Garvey.

Fullan, M. (1993). *Change forces.* Bristol, PA: Falmer Press.

Futrell, M. H. (1989, September). Mission not accomplished: Educational reform in retrospect. *Phi Delta Kappan*, 9-14.

Garner, B. (Ed.) (1997, February). Connecting research to practice. *Focus on Basics, 1*(A).

Gee, J. P. (1990). *Social linguistics and literacies: Ideology in discourses.* Bristol, PA: The Falmer Press.

Gee, J. P., Hull, G., & Lankshear, C. (1996). *The new work order: Behind the language of the new capitalism.* Boulder, CO: Westview Press.

General Accounting Office (1995, September). *Adult education: Measuring program results has been challenging*, Report to Congressional Requesters. Washington, DC: U.S. General Accounting Office (GAO/HEHS-95-153).

Gillespie, M. K. (1990). *Many literacies: Modules for training adult beginning readers and tutors.* Amherst, MA: Center for International Education, University of Massachusetts.

Gillespie, M. K. (1996). *Learning to work in a new land: A review and sourcebook for vocational and workplace ESL.* Washington, DC: Center for Applied Linguistics.

Greene, J. (1994). Qualitative program evaluation: Practice and promise. In N. Denzin, & Y. S. Lincoln (Ed.), *Handbook of qualitative research* (pp. 530-544). Thousand Oaks, CA: SAGE Publications Ltd.

Griffith, W.S., & Cervero, R.M. (1977). The Adult Performance Level Program: A serious & deliberate examination. *Adult Education*, 21, 209-224.

Grummon, P.T.H. (1997, February). Assessing students for workplace readiness. *Centerfocus*, *15*. Berkeley, CA: National Center for Research in Vocational Education.

Guijt, I., & Sidersky, P. (1996). Agreeing on indicators. *ILEIA Newsletter*.

Hamilton, M. , Barton, D.,& Ivanic, R. (1994). *Worlds of literacy*, Clevedon, Avon: Multilingual Matters, and Toronto: Ontario Institute for Studies in Education.

Harman, D. C. (1985). *Turning illiteracy around: An agenda for national action* (Working Paper No. 2). The Business Council for Effective Literacy.

Heath, S. B. (1983). *Ways with words: Language, life and work in communities and classrooms.* Cambridge: Cambridge University Press.

Hull, G. (Ed.). (1997). *Changing work, changing workers: Critical perspectives on language, literacy, and skills.* Albany, NY: State University of New York.

Hunter, C. St. J., & Harman, D. C. (1979). *Adult illiteracy in the United States: A report to the Ford Foundation.* New York: McGraw-Hill Book Company.

Johnston, W.B., & Packer, A.E. (1987). *Workforce 2000: Work and workers for the twenty-first century.* Indianapolis, IN: Hudson Institute.

Kutner, M., Webb, L., & Herman, R. (1993). *Management information systems in adult education: Perspectives from the states and from local programs.* (NCAL Technical Report TR93-4). Philadelphia, PA: National Center on Adult Literacy.

Kutner, M., Webb, L., & Matheson, N. (1996). *A review of statewide learner competency and assessment systems.* Washington, DC: U.S. Department of Education, Division of Adult Education and Literacy.

Lam, T. C. (1995). *Fairness in performance assessment*. ERIC Digest (ED391982 95).

Lankshear, C. (with Gee, J. P., Knobel, M. & Searle, C.) (1997). *Changing literacies*. Buckingham: Open University Press.

Laubach Literacy Action (1996). *National quality standards for volunteer literacy programs.* Syracuse, NY: Author.

Literacy South (1997). *Phenomenal changes: Stories of participants in the portfolio project.* Durham, NC: Author.

Lytle, S. L., & Wolfe, M. (1989). *Adult literacy education: Program evaluation and learner assessment.* Columbus, OH: ERIC Clearinghouse on Adult, Career, & Vocational Education. (ERIC Information Series No. 338).

Macedo, D. (1994). *Literacies of power: What Americans are not allowed to know.* Boulder, CO: Westview Press.

McGrail, L. (1995). Memories of Mami in the family literacy class. In Weinstein-Shr, G. & Quintero, E. (Eds.). *Immigrant learners and their families: Literacy to connect the generations* (pp. 77-89). Washington, DC: Delta Systems, Center for Applied Linguistics.

Merrifield, J., Bingman, M. B., Hemphill, D., & deMarrais, K. (1997). *Life at the margins: Literacy, language & technology in everyday life*. New York, NY: Teachers College Press.

Miller, P. (1995). Adult education's mislaid mission. *Adult Education Quarterly*, *46*(1), 43-52.

Mislevy, R. J. (1995). *Linking adult literacy assessments.* Princeton, NJ: Educational Testing Service.

Moore, M. T. & Stavrianos, M. (1995). *Review of adult education programs & their effectiveness: A background paper for reauthorization of the Adult Education Act.* Submitted to U.S. Department of Education. Washington, DC: National Institute for Literacy.

Murnane, R. J. & Bickerton, B. (1997, February). A productive partnership. *Focus on Basics, 1*(A), 1, 3-5.

National Center for Education Statistics (1994). *Adult literacy in America*. Paper prepared by Educational Testing Service for the National Center for Education Statistics, Office of Educational Research and Improvement, U.S. Department of Education.

National Center on Education and the Economy (1990). *America's choice: High skills or low wages!* Rochester, NY: Author.

National Commission on Excellence in Education (1983). *A nation at risk: The imperative for educational reform.* Report to the nation and the Secretary of Education, U.S. Department of Education. Washington, DC: Author.

National Education Goals Panel (1993). *Goals 2000 educate America: The national education goals report*. Washington D.C.: U.S. Department of Education.

National Governors Association (n.d.). *Performance-based governance*. Strategy Group on Performance-Based Governance, NGA.

National Institute for Literacy (NIFL) (1995a). *A review of the Oregon and Texas experience in building performance measurement & reporting systems.* Prepared by M. Campbell, Brizius and Foster. Washington, DC: Author.

National Institute for Literacy (1995b). *Literacy works*. Washington, DC: Author.

New Economics Foundation (1997). *Quality of life briefing.* London: Author.

Osborne, D. & Gaebler, T. (1993). *Reinventing Government: How the entrepreneurial spirit is transforming the public sector.* New York: Plume. (Original work published 1992, Reading, MA: Addison-Wesley).

Pedler, M., Burgoyne, J. , & Boydell, T. (1991). *The learning company: A strategy for sustainable development*. London: McGraw-Hill Company.

Pennsylvania Bureau of Adult Basic and Literacy Education (n.d.). Informational material about project EQUAL. Lancaster, PA: Project EQUAL.

Pennsylvania Bureau of Adult Basic and Literacy Education (1997). *How adult education program improvement is working in Pennsylvania.* Lancaster, PA: Project AXIS.

Prinsloo, M. & Breier, M. (1997). *The social uses of literacy: Theory and practice in contemporary South Africa,* Cape Town: Sached Books.

Quigley, B. A. & Kuhne, G. W. (1997). *Creating practical knowledge through action research: Posing problems, solving problems, and improving daily practice.* New Directions for Adult & Continuing Education, No. 73. San Francisco, CA: Jossey-Bass.

Rosen, D. J. (1997). NLA Discussion: Accountable to whom for what? Posting to NLA list-serve, 13 May 1997. Quoted with permission.

Rudner, L. M. & Boston, C. (1994, Winter). Performance assessment. *ERIC Review*, *3*(1), 2 - 12.

Rugh, J. (1986). *Self-evaluation: Ideas for participatory evaluation of rural community development projects.* Oklahoma City, OK: World Neighbors Inc.

Secretary's Commission on Achieving Necessary Skills (SCANS) (1991). *What work requires of schools: A SCANS report for America 2000.* Washington, DC: U.S. Department of Labor.

Secretary's Commission on Achieving Necessary Skills (SCANS) (1992). *Learning a living: A blueprint for high performance.* A SCANS report for America 2000. Washington, DC: U.S. Department of Labor.

Shavelson, R. J., Baxter, G. P., & Pine, J. (1992, May). Performance assessments: Political rhetoric and measurement reality. *Educational Researcher*, 22-27,

Smith, M., & Lytle, S. (1993). *Inside/outside: Teacher research & knowledge,* New York: Teachers College Press.

Stagg, D. D. (1992). *Alternative approaches to outcomes assessment for postsecondary vocational education.* Berkeley, CA: National Center for Research in Vocational Education.

Stedman, L. C. & Kaestle, C. F. (1987, Winter). Literacy and reading performance in the United States, from 1880 to the present. *Reading Research Quarterly, XXII*(1), 8-46.

Stein, S. G. (1993). *Framework for assessing program quality*. Washington, DC: Association for Community Based Education (ACBE).

Stein, S. G. (1995). *Equipped for the future: A customer-driven vision for adult literacy and lifelong learning.* Washington, DC: National Institute for Literacy.

Stein, S. G. (1997). *Equipped for the future: A reform agenda for adult literacy and lifelong learning.* Washington, DC: National Institute for Literacy.

Sticht, T. G. (1975). *Reading for working: A functional literacy anthology.* Alexandria, VA: Human Resources Research Organization.

Sticht, T. G. (1983). *Literacy and human resources development at work: Investing in the education of adults to improve the educability of children.* Alexandria, VA: Human Resources Research Organization, Professional Paper 2-83.

Sticht, T. G. (1988). Adult literacy education. Chapter 2 in E. Rothkopf (Ed.) *Review of research in education, 15.* Washington, DC: American Educational Research Association.

Sticht, T. G., & Armstrong, W. (1994). *Adult literacy in the United States: A compendium of quantitative data and interpretive comments.* Washington, DC: National Institute for Education.

Stites, R. (1996, December 19). Retention of literacy skills. Posting to NLA list-serve. Quoted with permission.

Street, B. (1984). *Literacy in theory and practice*. Cambridge: Cambridge University Press.

Street, B. (1995). *Social literacies: Critical approaches to literacy in development, ethnography and education.* London and New York: Longman.

Swadley, R. & Ziolkowski (1996). *National Institute for Literacy PMRIS evaluation: Final report, cross-site evaluation and case studies.* Prepared for NIFL. Washington, DC: National Institute for Literacy.

Szwed, J. F. (1981). The ethnography of literacy. In M. F. Whiteman (Ed.) *Writing: The nature, development and teaching of written communication* (pp. 13-23). Hillsdale, NJ: Lawrence Erlbaum Associates.

Taylor, D. (1983). *Family literacy: Young children learning to read and write.* Exeter, NH: Heinemann.

Toms, K. (1995). *The adult education training and workforce preparation system in New York: System description*. (Critical Issues Discussion Paper # 1). Albany, NY: University at Albany, State University of New York, School of Education.

Tucker, M. & Codding, J. (1998) Raising our standards for the standards movement. *Education Week*, February 18, 56f.

United Nations (1997). *World investment report*. New York, NY: Author.

U.S. Department of Education (1992, July). *Model indicators of program quality for adult education programs*. Washington, DC: Office of Vocational and Adult Education.

Van Horn, B. L., Carman, P. S., Askov, E. N., & Jenkins, P. S. (1996). *Assessment and adult learners: Getting the most from standardized and informal assessment instruments.* Prepared for Pennsylvania Bureau of Adult, Basic and Literacy Education. University Park, PA: Institute for the Study of Adult Literacy, The Pennsylvania State University.

Venezky, R. L. (1992). *Matching literacy testing with social policy: What are the alternatives?* (NCAL Policy Brief, Doc. No. PB92-1). Philadelphia, PA: National Center on Adult Literacy, University of Pennsylvania.

Wallerstein, N. (1983). *Language and culture in conflict: Problem-posing in the ESL classroom.* Reading, MA: Addison-Wesley.

World Bank (1997). *Indicators and methods to measure participation, demand orientation and local organizational capacity in community driven projects,* (Report of a workshop sponsored by Social Policy and resettlement Division, Environment Department). Washington, DC: Author.

Young, M. B., Fleischman, H, Fitzgerald, N., & Morgan, M. A. (1995). *National evaluation of adult education programs: Executive summary.* Prepared for U.S. Department of Education. Arlington, VA: Office of the Under Secretary by Development Associates Inc.