

Fall Institute 2011

IALS (INTERNATIONAL ADULT LITERACY SURVEY)

ITS MEANING AND IMPACT FOR POLICY AND PRACTICE

OCTOBER 23-25, 2011

BANFF, AB

Some Thoughts on
IALS Measurement Validity, Program Impact,
and Logic Models for Policy Development

Author:
Stephen Reder
Professor
Portland State University

Document category:
Think paper

The institute offers a forum for diverse opinions and perspectives. The views expressed in this document are the author's, not necessarily those of institute organizers and sponsors.

The author was able to make revisions after the institute and has agreed to post this version.

Some Thoughts on IALS Measurement Validity, Program Impact, and Logic Models for Policy Development

In this background “think piece” for the IALS Institute, I look briefly at a few key issues in using the three IALS proficiency measures (prose, document and quantitative literacy) in large-scale national and international assessments as well as in program evaluation and policy frameworks for adult basic skills education. I consider in broad brushstrokes some questions about validity, program impact and policy development related to the IALS proficiency measures. There is space here only to skim the surface of and reflect on these issues – more in-depth and technical treatments are available in the cited references.

Validity Issues in IALS Assessments

Policy-makers and educators who rely on the measurement of literacy in IALS are naturally concerned with the validity of those measures. In science and statistics, many definitions have been developed for validity. In general, validity refers to the extent to which measurement of a concept such as literacy is well-founded and accurately corresponds to the real world. Fairly standard techniques were used to develop the proficiency measures used in IALS and related large-scale national and international assessments (Murray, Kirsch & Jenkins, 1998). There have been a number of critiques of the validity of these measures and how they have been used for programmatic and policy purposes. Some critiques have been fairly broad, aimed at what is viewed as inherent problems in using standardized tests for educational assessment and policy purposes. With respect to IALS, one such criticism is that this type of test assesses decontextualized skills and test-taking abilities rather than what individuals are able to do in contextualized everyday activities that involve reading, writing and math skills, whereas the results are interpreted in terms of everyday literacy performance (Hamilton, 2001; Hamilton & Barton, 2000; Street, 1997). Others have pointed out the strong cultural assumptions, hence bias, underlying the assessments (Hamilton, 2001; Street, 1997). Other critiques have focused more on various technical aspects of the measures, such as whether the three literacy scales actually measure different constructs; whether some critical dimensions of literacy (e.g., writing) are measured at all; where cut scores on the scales should be set for defining performance “levels” (or whether, indeed, there actually are qualitatively distinct performance levels at all); and how the pattern of correct and incorrect responses to test items is converted into proficiency scores (Blum, Goldstein & Guérin, 2001; Reder, 1998; Sticht, 2001).

Despite the often heated debates about these and other validity concerns, the IALS proficiency measures have managed to stand the test of time, buoyed largely I believe by a validity many attribute to them because of their strong correlations with measures of educational attainment and economic status in all participating IALS countries. These same correlations have placed the IALS assessment results in policy

frameworks in many countries (Coulombe, Tremblay & Marchand, 2004; Kirsch, Braun, Yamamoto & Sum, 2007; Organization of Economic Cooperation and Development, 2000).

Given the perceived relevance of the IALS proficiency measures to economic development, some of the expressed concerns about their use could be addressed by including other types of literacy measures in the large-scale assessments. By adding other kinds of measures to the IALS proficiencies, it would be possible to maintain historical trend data for proficiency while providing a more varied set of measures to inform future adult literacy programs and policy. In addition to proficiency, I have suggested (Reder, 2009b) using measures of literacy and numeracy *practices* (i.e., whether and how people use basic skills in everyday activities). Other possibilities include information processing sub-skills (e.g., decoding, vocabulary) and self-reported measures about the adequacy of individuals' skills for their work and other everyday activities.

Would we see fundamentally different pictures of how adult literacy is distributed and develops in society through the multiple lenses of these varied measures? A study that I was involved with provides an answer. The *Longitudinal Study of Adult Learning* (LSAL) followed a random sample of about 1,000 low education adults in the US over nearly a decade. It repeatedly administered IALS-like proficiency measures over time as well as closely tracking their educational activities, employment and earnings over the same time period. Six waves of interviews were completed, spanning about eight years of each individual's life. LSAL periodically repeated the measurement of literacy proficiency, literacy practices, self-reported skill improvement and reading sub-skills. Although all of these measures are positively correlated (e.g., those with higher proficiency also exhibit higher levels of engagement in everyday literacy practices, higher vocabulary scores, etc.), they show very different dynamics of change when measured repeatedly over time.

At each point in time, LSAL's data show strong positive correlations between literacy proficiency and employment and earnings, just as seen in IALS. Such correlations between proficiency and earnings do not, of course, necessarily imply causality, and care is appropriate when thinking about this relationship. But LSAL shows something else that's equally important to think about: *changes* in individuals' measured proficiency scores over time are correlated with *changes* in their employment and earnings over those same time periods. Understanding why some individuals gain proficiency in adult life while others lose proficiency is thus likely to be very important for adult literacy policy and program design (Willms & Murray, 2007). In LSAL, key life events and prevailing economic conditions seem to influence proficiency growth over the lifespan. For example, changes in an individual's employment and earnings are associated with observed changes in their proficiencies; individuals gaining employment or increased earnings tend to show increasing proficiencies, and vice-versa (Reder, 2010b). However, no relationship was observed in LSAL between proficiency change and participation in adult basic skills programs (Reder, 2009a).

Measuring Program Impact

This is a striking finding because current accountability regimes hold programs accountable for producing short-term “learning gains” on proficiency measures when, according to LSAL results, programs do not have short-term effects on this type of literacy measure. This lack of impact may seem at odds with the small positive proficiency learning gains that programs typically report in comparing their participants’ pre- and post-test scores on proficiency tests; however, analyses of program data typically do not include gains measured for comparable groups of adults who do not participate in programs. LSAL found overall small gains in proficiency among program participants, as program evaluations typically report, but found equivalent gains in proficiency over time in comparable program non-participants.

In sharp contrast, the LSAL data do exhibit a strong positive relationship between program participation and changes in literacy and numeracy *practices* measures (i.e., how often individuals do things like read a newspaper, use math at home, etc). With many statistical controls in place, LSAL found strong relationships between participation in adult education programs and increased engagement with literacy (e.g., reading books) and numeracy (e.g., using math at home) practices (Reder, 2009a). The sequence of the observed changes makes it clear that program participation influences practices rather than vice-versa. This finding is consistent with the relationship that Purcell-Gates and colleagues (2004, 2000) have reported between types of adult education programs and changes in practices observed among program participants. Purcell-Gates *et al* (2000) found that adult education students from programs that focus instruction around authentic literacy materials and practices report greater changes in their literacy practices than students from programs not centered around authentic literacy practices. LSAL contrasts the development of literacy and numeracy practices between program participants and non-participants whereas Purcell-Gates *et al* contrast the development of literacy practices among participants in different types of programs.

Sheehan-Holt and Smith (2000) analyzed the U.S. National Adult Literacy Survey (NALS) data, looking at cross-sectional differences between recent program participants and non-participants. With many background characteristics statistically controlled, they found no significant differences between participants and non-participants in the IALS-like proficiency measures. But they did find significant differences between the same groups in measures of reading practices. Their cross-sectional findings are thus consistent with the longitudinal findings of LSAL. Three reviews of research on the impact of program participation on literacy proficiency have all concluded that there is no evidence of systematic impact when the analyses involve comparison groups and statistical controls (Beder, 1999; Brooks *et al*, 2001; Smith, 2009).

LSAL respondents were also asked regularly to report on changes in their reading, writing and math since their previous interview. A set of structured questions was asked on each wave of interviews about self-perceived changes in respondents’ reading, writing and math since the previous wave. These questions were asked at Wave 2 about changes since Wave 1, at Wave 3 about changes since Wave 2, and so forth. For each of reading, writing and math, questions were asked about changes in skill level, changes in the

frequency of using skills in everyday activities, and changes in the types of tasks and materials used. These measures show clear effects of recent program participation. We found clear relationships between individuals' participation between two successive time points and their reported changes in basic skills between those same time points compared to changes reported after periods in which they did not participate in a program. These effects remain strong even after many other factors are statistically controlled (Reder, 2010a).

While key life history events also appear to be related to self-perceived changes in basic skills, program participation has stronger influences on basic skill changes than life history events do (Reder, 2010a). Examples of life history events in LSAL that affect literacy development include starting a job after a period of non-employment, which has a significantly positive relationship with the likelihood of reading better. The recent addition of a child to the household has a positive impact on the frequency of reading, perhaps because it engages parents in reading about child-raising or perhaps in reading *with* children. The presence of a new partner in the household significantly enhances the likelihood of doing math more often, whereas a recent increase in household income is associated with increased likelihood of doing math with new kinds of materials. Financial record-keeping and budgeting become more frequent as the number of adults in a household increases, which occurs with the addition of a new partner. This may lead to more use of math in everyday activities. As total household incomes rise, new kinds of tax and employment-related forms that involve math, for example, may need to be completed. Nevertheless, program participation has stronger influences on basic skill changes than these key events.

A Dilemma for IALS and Adult Basic Skills Education

These findings pose a critical dilemma for adult education programs and policy frameworks. On one hand, proficiency measures are strongly related to economic outcomes within IALS and other large-scale surveys, and these relationships are used to justify investments in programs designed to increase proficiency. On the other hand, research shows that programs have demonstrable impact on measures of literacy and numeracy practices but not on proficiency measures, at least not over the relatively short time intervals typical of program participation and of program accountability and improvement cycles.

We thus have a major misalignment between the effects programs are actually having on their students' literacy and numeracy development, on one hand, and the short-term proficiency gains for which programs are accountable under the dominant policy and funding regimes. As the stakes rise in these accountability schemes, such misalignments are likely to produce substantial distortions in educational practice (Reder, 2009b).

Furthermore, these results suggest that using short-term proficiency gains as the main criterion in program improvement processes may be very problematic. If short-term proficiency gains are not closely linked to program participation, then they may not differentially reflect variations in types and quality among programs, either. Because LSAL did not gather many details about the types of programs that respondents participated in, it is difficult to tell from its data. The findings of Purcell-Gates and colleagues, however,

demonstrating that practices-based measures are sensitive over relatively short intervals to differences among types of adult literacy programs suggest that practices-type measures may be effective indicators of program quality and thus suitable for program improvement processes.

Resolution: A Multiple-Measures Approach

To resolve this dilemma, we need to better understand the role of literacy and numeracy practices in producing long-term gains in proficiency. Practice engagement theory (Reder, 1994; Sheehan-Holt & Smith, 2000) hypothesizes that higher levels of engagement in literacy practices (which programs do produce) can lead to longer-term changes in proficiency. Analyses of the repeated measures in the LSAL data provide clear evidence of *practice engagement effects* on long-term proficiency development (Reder, 2009b). Adults at similar proficiency levels at one point in time wind up many years later at different proficiency levels depending in part on their earlier levels of engagement in literacy practices. Those with higher levels of engagement at an initial point in time have higher levels of proficiency at a later point in time even with initial levels of proficiency controlled. Purcell-Gates and colleagues (2000) provided evidence that programs – especially programs using authentic materials and practices in the classroom – foster higher levels of engagement in literacy practices in their students after they leave the programs.

There is thus a chain of evidence linking programs to increased engagement in practices and linking practice engagement over longer periods of time to increased proficiency levels. Without the use of literacy practices measures, a systematic connection between programs and proficiency is not evident in these data. We can make the case for using practices-based measures by arguing that increased levels of practice engagement – something that programs do produce – have long-term consequences valued by policy makers, that is, increased proficiency.

The research we have considered here suggests that measures of adults' engagement in literacy and numeracy practices should be used in conjunction with the well-entrenched proficiency measures such as those of IALS to provide a richer quantitative framework for understanding and assessing adult literacy and numeracy development. Longitudinal data about learners indicates that adult education programs are more closely aligned in the short-term with practice engagement measures than with proficiency measures. Program participation leads to increased practice engagement which over time leads to the very gains in proficiency valued by policy makers.

Although proficiency measures will likely retain a place in any policy framework because of their close empirical relationship to schooling and economic status, additional measures are needed that focus on literacy and numeracy practices and are more closely aligned with the impact that programs actually have on adult literacy and numeracy development. Adding appropriate practices-based measures to policy and programmatic frameworks would broaden and enrich policy-makers' perspectives on adult literacy and

numeracy development and, when used in program improvement efforts, would lead to more effective programs.

Other types of measures are not as well-suited for this purpose. We have already seen that proficiency measures are not well-suited to be indicators of short-term program impact. How about measures of the sub-skills involved in reading, things like vocabulary knowledge, decoding and fluency measures, and so forth? A recently completed program of experimental research studies indicates that pre-post changes on a wide variety of commonly used sub-skill measures do not systematically vary across different types of adult basic reading instruction. These studies were randomized control trials (RCTs) in which adult students were randomly assigned to various instructional conditions, with pre- and post-measurement of numerous sub-skills. All studies found essentially the same thing: although students are learning, there are no systematic, significant effects of instructional conditions on a wide range of commonly used sub-skills measures (Miller, Esposito & McCardle, 2011). This striking result is seen again in another randomized control trial study with non-native English speakers in adult literacy classes: the absence of systematic, significant effects of instructional conditions on measured changes in reading (Condelli, Wrigley & Yoon, 2009).

Changing Our Logic Model

Program designers and policy-makers expect to see positive relationships between students' hours of attendance in adult literacy programs and their proficiency gains. This expectation is based on the assumption that the longer students sit in class, the more they should learn and progress on these measures. This is the traditional *logic model* at the core of program design and policy frameworks in adult basic skills education. Administrative data gathered in federally funded adult education programs in the U.S. are often used to display the relationship between hours spent in programs and gains on proficiency test scores. Researchers and evaluators have reported that increased test score gains are associated with increased hours of participation in programs (Rose, 2009). Such relationships, of course, prompt policy-makers and practitioners to redouble efforts to increase student persistence in these programs, reasoning that increased hours in programs will lead to increased proficiency gains (Comings, 2009).

Reasonable as this sounds, it may not be quite so straightforward. When I have re-analyzed such data from individual states in the U.S., I have seen how program testing policy inadvertently creates an important and misleading effect. Many programs mandate repeated administration of proficiency tests after so many hours of attendance or so many weeks or months of enrollment. This has the unintended consequence of confounding the students' hours in class with the number of times they have taken the proficiency test. There is good reason to believe that students get better at taking such tests the more often they take them (even when alternative forms of the test are used). When I have analyzed such test-retest data in ways that separate the effects of times the test has been taken from hours of instruction, there are no net effects of instructional hours on reading and math scores for native speakers and considerably diminished effects for non-native speakers.

This does *not* mean that programs are ineffective. It could be that this is just a further reflection of programs not having short-term impact on proficiency measures to begin with, as seen above with the LSAL data. Perhaps we would see such relationships if practices-based measures were used in the pre-post data. After all, Purcell-Gates and colleagues found differences in literacy practices growth in students participating in different *types* of programs. In LSAL, we did not find clear effects of the amount of participation on literacy and numeracy practices measures, just overall differences according to whether people participated or not. The experimental randomized control trial classroom studies of reading instruction described above also examined the effects of hours of attendance on gains in various reading sub-skills and, generally speaking, found no systematic effects of classroom participation time on a wide range of outcome measures (Condelli et al, 2009; Miller et al, 2011).

So perhaps we need to rethink the underlying logic model that we have been using for so long to link participation in the adult education program to proficiency growth. Leander (2009) has called the prevailing logic model of adult education a “parking lot” model, in which what matters for learning is how long students are “parked” in the program. Instead, he suggests, we should think of adult education as a “busy intersection”, in which what matters for learning is not how long they spend waiting in the intersection but the direction they take when they leave the intersection. In this conception, the adult education program tries to help students, who come in to the program from many different directions and depart towards many different destinations, choose the best path forward as they leave the program and provide them with the resources and supports they will need to become persistent lifelong learners and reach their destinations (Lesgold & Welch-Ross, 2012; Miller et al, 2011). In this logic model, the program’s impact on learning may be best seen in different ways at different points along the adult’s trajectory forward. According to the research we have considered here, the initial impact of adult basic skills programs is best measured in terms of changing literacy and numeracy practices. Over time, these changes in practice will lead to increased proficiency levels and with them the enhanced economic development that societies and public agencies are seeking. To support this growth, we need to broaden the IALS portfolio of assessment tools and revise the logic model we use to connect adult education programs with long-term learning outcomes.

References

- Beder, H. (1999). *The outcomes and impacts of adult literacy education in the United States*. (Report No. 6). Cambridge, MA: National Center for the Study of Adult Learning and Literacy. Retrieved from <http://ncsall.gse.harvard.edu/research/report6.pdf>
- Blum, A., Goldstein, H., & Guérin-Pace, F. (2001). International Adult Literacy Survey (IALS): An analysis of international comparisons of adult literacy. *Assessment in Education*, 8(2), 225-246.
- Brooks, G., Davies, R., Ducke, L., Hutchison, D., Kendall, S., & Wilkin, A. (2001). *Progress in adult literacy: Do learners learn?* London: The Basic Skills Agency.
- Comings, J. (2009). Student persistence in adult literacy and numeracy programs. In S. Reder & J. Bynner, (Eds.), *Tracking adult literacy and numeracy: Findings from longitudinal research* (pp. 160-176). New York and London: Routledge.
- Condelli, L., Wrigley, H. S., & Yoon, K. S. (2009). "What works" for adult literacy students of English as a second language? In S. Reder & J. Bynner, (Eds.), *Tracking adult literacy and numeracy: Findings from longitudinal research* (pp. 132-159). New York and London: Routledge.
- Coulombe, S., Tremblay, J.-F., & Marchand, S. (2004). *Literacy scores, human capital and growth across fourteen OECD countries*. Ottawa: Statistics Canada.
- Hamilton, M. (2001). Privileged literacies: Policy, institutional process and the life of the IALS. *Language and Education*, 15(2-3), 178-196.
- Hamilton, M., & Barton, D. (2000). The International Adult Literacy Survey: What does it really measure?, *International Journal of Education*, vol 46, no 5, pp 377-389.
- Kirsch, I. S., Braun, H., Yamamoto, K., & Sum, A. (2007). *America's perfect storm: Three forces shaping our nation's future*. Princeton, NJ: Educational Testing Service.
- Leander, K. (2009). Studying and designing for adult literacy learners on the move: What's missing from the 'network' metaphor? Position paper prepared for the *Summit on the Future of Adult Education in the New Digital World*, Virginia Commonwealth University, Richmond VA, November 5-6, 2009.
- Lesgold, A., & Welch-Ross, M. (Eds.) (2012). *Improving adult literacy instruction: Options for practice and research*. Washington DC: National Academies Press. Prepublication copy retrieved from http://www.nap.edu/catalog.php?record_id=13242
- Miller, B., Esposito, L. & McCardle, P. (2011). A public health approach to improving the lives of adult learners: An introduction to the *Special Issue* on adult literacy interventions. *Journal of Research on Educational Effectiveness*, 4, 87-100.
- Murray, T. S., Kirsch, I. S., & Jenkins, L. B. (Ed.) (1998). *Technical report on the first International Adult Literacy Survey*. Washington DC: National Center for Education Statistics.

- Organization for Economic Cooperation and Development (2000). *Literacy in the information age: Final report of the International Adult Literacy Survey*. Paris: Author.
- Purcell-Gates, V., Jacobson, E., & Degener, S. (2004). *Print literacy: Uniting cognitive and social practice theories*. Cambridge, MA: Harvard University Press.
- Purcell-Gates, V., Degener, S., Jacobson, E., & Soler, M (2000). *Affecting change in literacy practices of adult learners: Impact of two dimensions of instruction*. Cambridge MA: National Center for the Study of Adult Learning and Literacy, Harvard graduate School of Education.
- Reder, S. (2010a). Participation, life events and the perception of basic skills improvement. In J. Derrick, J. Field, P. Lavender, S. Meyer, U. Howard, & T. Schuller (Eds.), *Remaking Adult Learning: Essays on Adult Education in Honour of Alan Tuckett*, London: Institute of Education.
- Reder, S. (2010b). *Adult literacy development and economic growth*. Washington DC: National Institute for Literacy. <http://lincs.ed.gov/publications/pdf/AdultLiteracyDevEcoGrowth.pdf>
- Reder, S. (2009a). The development of literacy and numeracy in adult life. In S. Reder & J. Bynner, (Eds.), *Tracking adult literacy and numeracy: Findings from longitudinal research* (pp. 59-84). New York and London: Routledge.
- Reder, S. (2009b). Scaling up and moving in: Connecting social practices views to policies and programs in adult education. *Literacy and Numeracy Studies*, 16.2/17.1 (1), 35-50.
- Reder, S. (1994). Practice engagement theory: A sociocultural approach to literacy across languages and cultures. In B. Ferdman, R.M. Weber & A. Ramirez (Eds.), *Literacy across languages and cultures*. Albany NY: State University of New York Press.
- Reder, S. (1998). Issues of dimensionality and construct validity of the NALS assessment. In M. C. Smith (Ed.). *Literacy for the 21st century: Research, policy, practices, and the National Adult Literacy Survey*. (pp. 37-57). Westport, CN: Greenwood Publishing (Praeger).
- Rose, S. (2009). Using a longitudinal approach with state administrative records to evaluate adult education programs in the United States. In S. Reder & J. Bynner, *Tracking adult literacy and numeracy: Findings from longitudinal research* (pp. 278–292). New York and London: Routledge.
- Sheehan-Holt, J., & Smith, C. (2000). Does basic skills education affect adults' literacy proficiencies and reading practices?, *Reading Research Quarterly*, 35(2) 226-243.
- Smith, M. C. (2009) 'Literacy in adulthood'. In M. C. Smith (Ed.), *Handbook of Research on Adult Learning and Development*. New York and London: Routledge.
- Sticht, T. (2001) 'The International Adult Literacy Survey: How well does it represent the literacy abilities of adults?' *The Canadian Journal for the Study of Adult Education*, 15, 19-36.
- Street, B. (1997) *Literacy, economy and society: A review*. *Working Papers on Literacy*. Montreal: The Centre for Literacy.
- Willms, D., & Murray, T. S. (2007). *Gaining and losing literacy skills over the lifecourse*. Ottawa: Statistics Canada.