UPPINGHAM SEMINARS
REPORT ON UPPINGHAM SEMINAR MARCH 2005:

MEETINGS THROUGH CONTEXTS: Ethnographic Approaches to Measuring Literacy

Tracey Costley

The Uppingham Seminars began as part of a series of seminars and conferences that were organised to promote the International Literacy Year of 1990. These particular seminars were originally run by Education for Development and were subsequently carried on by Alan Rogers in the form of the Uppingham Seminar Series. The seminars were conceptualised to provide three days of *'space'* for busy practitioners from a variety of professional and academic quarters to get together to discuss issues pertinent to the field of Literacy and Development. The Uppingham Seminars have always sought to provide a forum for open discussion and have never rigidly set goals for the seminar to arrive at pre-established conclusions. Participants invited to the seminars represent a range of diverse fields, and the seminars are the site of often conflicting opinions, approaches and experiences. What unites the participants at the Uppingham Seminars is their commitment to working through differences in the form of co-constructed dialogue and discussions that seek to move debates, issues and ideas forward through co-operative participation in discussion rather than allow differences to result in polarisations which hinder positive and fruitful discussion and progress.

This seminar, held at the Department for Education and Professional Studies (DEPS) at King's College London, was a one day follow-up seminar to the November 2003 *Measuring Literacy: ethnographic approaches to the collection of comparable data on literacy* seminar that was held over three days at The Falcon Hotel in Uppingham, Rutland. The central aims of UppSem 2003 were to begin to establish a *"common ground between ethnographic methods and more conventional measurement methods for assessing literacy and numeracy; and to seek out and document ways that ethnographic methods can contribute to measuring literacy"* (Petersen 2003: 2). The seminar further sought to *"examine the consequences of a 'multiple literacies' view, including multi-lingual literacies, for the measurement and assessment of literacy"* (ibid).

The seminar, through the presentation of a range of papers, group discussions and plenary sessions, outlined the above generated questions, focussing on the What, How and Whys of measuring literacy. Their findings are briefly summarised in the following examples:

- in measuring literacy and numeracy, *what* is it that we actually measure and *how* is this most 'accurately' achieved?
- what is *the purpose* of measuring literacy and *for whom* are we measuring it?

These questions naturally draw attention to the issue of scale and pose further questions such as:

- are we working at the macro-scale for international comparison and accountability or at a micro-'ground' level that seeks to inform practice and implementation?

A full and rich account of the discussions and conclusions of this seminar is publicly available via the following web address; www.uppinghamseminar.org/report-2003.htm in the form of a detailed report *Measuring Literacy: meeting in collision* written by Carolyn Petersen. A brief summary of the content of these discussions is that they enabled participants to forge a better understanding of the relative strengths and weaknesses of approaching the measurement of literacy from both qualitative and quantitative data perspectives and attempted to identify the role that ethnography can play at both the macro and micro level. The scope and nature of the questions raised at the 2003 seminar are such that they demanded further discussion and exploration; this is the grounding and context that shaped the follow-up seminar in March 2005, on which the rest of this report shall focus.

Ethnographic Approaches to Measuring Literacy

Four short papers, by Carolyn Peterson, Constant Leung, Clinton Robinson and Catherine Kell were circulated in advance of 3 March 2005 and provided the framing for the follow-up seminar (see Appendix 1). These papers engage with a number of issues from the *Meeting in Collision* report and therefore encouraged the continued progression of the discussion surrounding 'ways' of producing comparative statistical measurements of literacy (and numeracy) that seek to ensure contextual relevance and accuracy. The papers were drawn together by Alan Rogers (see Appendix 1) who through his assimilation of the papers provided five questions for a provisional agenda by which to guide the day's discussions. In the style that is true to the Uppingham Seminars, the papers and summary provided an initial platform from which the day's programme and structure were negotiated and decided by the participants.

The seminar developed over three interrelated stages. Whilst the structure of the discussion changed, the core content and focus of topics discussed remained consistent throughout. The opening stage was set in motion by Alan Rogers' posing of the five questions encapsulating the essence of the central issues raised in the four background papers. These questions brought forth a fruitful discussion which highlighted key themes, questions and issues which the group wished to continue discussion of throughout the seminar.

The second stage of the seminar was marked by participants narrowing down key topics for discussion and deciding upon the preferred format for the discussion. It was decided that, in order to capitalise upon the range of expertise and interests in the room, the group should remain as a whole rather than separating into smaller focus groups, which is often the format for the Uppingham Seminars. The agreed topics for discussion fell loosely into two areas: the *Contexts and Purposes* in and for which measuring could/should take place, and the potential *Units of Analysis and Methods* of measuring literacies. In order that each area received the maximum allocation of time, two informal chairs were chosen to facilitate the discussion; the first part was chaired by Mary Hamilton and the second by Constant Leung.

The third stage of the seminar was a group plenary with the day brought to a close with a discussion of possibilities for future seminar topics conducted by Alan Rogers and Brian Street. The remainder of this report will focus on presenting the details of the discussions from each stage of the seminar, in which the key points and underlying questions and ideas will be highlighted. In order that the account of the day's discussions is as coherent as possible, I have predominantly organised the report by key themes or points which emerged in each stage of the seminar rather than as they occurred chronologically in the unfolding discussion.

Stage One

As previously mentioned, the following questions raised by Alan Rogers provide a link between the 2003 seminar (as seen in the report *Meeting in Collision)* and the 2005 *Ethnographic Approaches to Measuring Literacy* seminar. The questions get to the heart of some of the key issues related to both identifying and understanding the relationship between ethnography and the measurement of literacy. It is fair to suggest that these questions represent the meta-questions of the seminar and were central to the discussions at all stages:

1) *How do we place ethnographic approaches to measuring literacy? Is a dichotomy of quantitative and qualitative approaches appropriate to describe the situation?*
2) *How should ethnographic 'assessments' be expressed?*
3) *What should the relations between ethnographic approaches and contemporary quantitative approaches be?*

4) *Do the two approaches provide any platform for comparability*?
5) *Is this discussion a valid and profitable one to be having?*

The macro level of the questions is such that they prompted a series of more micro or lower order questions that were necessary to begin to unpack and identify key issues. Unpacking these questions was a vital step in the process of clarification and conceptualisation. Below is a list of some of the principal questions that were raised and an account of the discussions that were generated surrounding each question:

- how can we reconcile the notion of measurement with ethnographic and qualitative data?
- what do we mean by literacy and how can we measure it appropriately?
- what are the uses and purposes of measuring literacy?
- what do the different approaches, i.e. quantitative, qualitative and ethnographic, each do well and what do they neglect, and what manipulations of the data are possible?

Reconciling ethnography and measurement

One of the dominant concerns expressed throughout the discussions was the need for a clearer understanding of what we mean by the term 'measurement' and more specifically, how; and if this is compatible with the notion of ethnography. The term 'ethnography' as suggested in the discussion by Nigel Norris *"doesn't automatically chime with"* the idea of measurement. These sentiments can be seen as expressing one of the fundamental tensions felt throughout the group which results from the idea of measurement being used in conjunction with the notion of ethnography. Inherent in the notion of measurement is the idea that one has access to something quantifiable and tangible that can naturally be rendered measurable. One of the goals of producing measurements is arguably to compare and contrast them with similar 'other' quantitatively measurable 'things'. One of the inherent contradictions between the concept of measurement and ethnography is that qualitative/ethnographic data is not often viewed as being something that is essentially measurable, nor is it often gathered for the purposes of measurement. Ethnographic/qualitative data generally takes the form of text, using description and illustration to enrich our understanding of aspects of social worlds and practices rather than the production of a set of measurements which seek to represent reality through numbers, charts and/or tables.

A question at the heart of this is: *what is the role of ethnography in measuring literacy?* The goal of large scale survey projects such as the International Adult Literacy Survey (IALS) and UNESCO's Literacy Assessment and Monitoring Programme (LAMP) is to produce data on national levels of literacy (UNESCO). Samples are used to make projections about a country's literacy levels with supporting data generated through assessment methodologies which are survey based and predominantly quantitative in nature. The claims to validity, comparability and accuracy of the data of such projects have been widely criticised from the standpoint of ethnographers and qualitative data scientists on the grounds that the complexities of the real world cannot be realistically and accurately encapsulated or expressed through statistics alone. The objectives of projects like LAMP are essentially very different from the objectives of ethnographers and qualitative social scientists. A central theme of the discussions therefore focussed on the ways in which ethnographic approaches could be used in the production of this type of data. The differences between the two are ideological and epistemological in nature and the question of the possible uses of ethnography was further extended by many participants asking if ethnography should be used at all to support the work of such projects.

Defining literacy and measurement

One of the principal points to be raised in relation to this heading was the need to voice what we mean and understand by the term 'literacy'. The fundamental conceptualisation of literacy at the seminar was one that draws from the work of the New Literacy Studies (NLS) group and takes the view that literacy is a social practice. Literacy is therefore not a single monolithic practice that remains consistent over time, space and context; rather there are multiple literacies and literacy practices in which and through which individuals engage and navigate the social worlds they encounter. In her paper, Catherine Kell highlights the widely felt sentiment that *"it doesn't make sense ... to measure literacy"*. As Kell suggests, measuring literacy doesn't make sense when the term 'literacy' is reified and abstracted from the practices to which it is inextricably linked. A key part of the discussions was therefore to tease out a collective understanding of actions and practices that are recognised as being instantiations of literacy and to begin to identify the 'sites' in which these practices take place.

The conceptualisation of literacy as representing a plurality of practices occurring over a variety of contexts and settings provides a further challenge to the notion of measurement. Measuring, for example, the number of hours an adult may spend attending a literacy learning programme requires wholly different measurement techniques and methods from those needed to measure an adult teaching themselves literacy outside of a 'formal' educational context. The discussions highlighted the dangers of reification in using the singular form of the term 'measurement' and emphasised that an ethnographic approach to measurement would by nature draw on and require data generated from a range of "*measurement practices"* (Paul Black) rather than a singular or *'autonomous model of measurement'* (Brian Street). As an eloquent note on one of the flip chart papers highlighted, our tools, methods and practices are determined by the contexts and settings in which we work. What an ethnographic approach would provide could then be described as 'measurements of literacies' or a measurement or account of the *'states of literacy'* (Aaron Benavot). What was also discussed is that any ethnographic or qualitative account of the 'states of literacy' will always generate data that is contextually determined, data which is therefore inextricably linked to and inseparable from the specific people, places, times and spaces of its enactment. The role that context must play in any understanding of or attempt to measure literacy was importantly highlighted by the recognition that the term 'literacy' itself may not even be a term, label or concept that individuals being surveyed may identify with or recognise, for example; in many contexts, 'numeracy' is a more widely accessible and invoked concept than literacy (David Baker).

IALS and LAMP are confident that checks and measures are in place to ensure that the data they are producing on literacy levels is accurate and comparable across countries. What such intentions and objectives assume is that what is being compared is essentially the same. By default, there is an implicit understanding or belief that 'literacy' is the same across countries and contexts. The discussions highlighted that an ethnographic approach to measuring literacy presents a fundamental challenge to this belief through assuming that literacy practices are different across countries and contexts. An integral belief within IALS and LAMP is that the surveys and samples are empirically comparing like with like; however the overriding view of the participants in the seminar was that such a belief is dangerously ill-founded and detached from the reality of 'real world' literacy practices. A valuable contribution to the strengthening of this sentiment came from Amina Osman who, in her work, has found that many of the IALS survey participants found the questions they were being asked at odds with or abstracted from the on-the-ground reality of their everyday lives and practices. Such findings challenge the notion that it is possible to develop a *one size fits all* model or approach to measuring social practices (Paul Black).

The uses and purposes of measuring literacy

A major part of the discomfort felt by participants at the seminar in relation to the uses and role of ethnography in the measurement of literacy stems from concerns and ideological clashes with the ultimate purposes and uses of the measurements themselves. The group by and large voiced widespread concern about the relationship between projected measurements of literacy and the movement, distribution and allocation of aid monies and funding. Statistical measurements are widely used by such agencies as UNDP and national governments as a development measure for donor funding (Alan Rogers). The Global Monitoring Report (GMR) highlights the pressure on such programmes as LAMP, IALS and national governments to produce statistical measures of literacy that can be used for the purposes of international comparisons. A major concern for many of the participants was the idea that the act of developing ethnographic measurements of literacy could inevitably become something of a statement of support for the macro-level, international drive for statistically driven league tables and their related policies (Kathy Maclachlan).

The group felt that in order to begin to help allay some of these fears, it would be necessary to develop a "sociology of the people" (Harvey Goldstein) involved in the design and implementation of the actual measuring of literacy practices and a more comprehensive understanding of the ways in which the data is ultimately used. Framing these discussions was an awareness of the structures and relations of power inherent in the practice of measuring and representing literacy levels. Current distributions of power favour and ultimately support the intentions and objectives of macro-level policies and programmes, and there was a sense within the group that statistical representations of literacy are produced predominantly with the goal of securing money rather than as a means of ensuring and promoting equity and social justice (David Baker). Whilst many in the group felt that entering into any kind of work on the production of statistical representations of literacy was not the business of ethnography and qualitative data scientists, there was an overwhelming sense that it is ultimately necessary to engage with the discourses in order to create a possible space or opportunity from which to influence and shape future policy and practice.

This led the group to return to the idea of the role of ethnographic measurements of literacies and to consider the ways in which data and understandings yielded from such approaches might seek to provide information that could offer an alternative or 'challenges' to the current overwhelming privilege given to traditional statistics (Kathy Maclachlan). A defining feature in this discussion was the idea that an understanding of the relative strengths and weaknesses of any model/s of measurement used for the purposes of understanding literacy levels must be developed in order that we are able to recognise what particular models do well and what they omit from and reject in their findings (David Baker).

Approaches and manipulations

One of the key areas of discussion not only in the context of this seminar but also in the wider context of research itself, centres on the ways in which different approaches i.e. quantitative, qualitative and ethnographic both collect and present data, and how these approaches actually impact, shape and define the data which is collected. One of the major criticisms of such projects as LAMP is that their fundamental claims to produce data that guarantees global comparability is not supported by data that is able to withstand or conform to the rigours of commonly accepted testing standards of science (Harvey Goldstein). What lay at the heart of this, in the context of the seminar, was not only a discomfort with the ultimate goal of comparability but also a recognition of the limitations inherent in the apparently unquestioned use of *"long-standing, hegemonic,*[and] *dominant models of measuring".* Such models bias and privilege traditional statistical measurements which, the group felt, are unable to recognise, document and account for *"the transitions and dynamisms of the field"* (Aaron Benavot) and which therefore fail to capture fully the realities of the social world. A principal danger in this is that these measurements are used in the long term by a host of agencies, organisations and individuals to make what are believed to be informed decisions about the distribution of funds, the development of aid programmes and interventions, and educational and pedagogic practices.

The participants were conscious of not entering into a debate or discussion which sought only to decry projects such as LAMP and their methodologies and simply argue for replacing one approach with another, but to work towards a more fruitful understanding of what *particular* approaches could bring to and add to the process of understanding literacy levels in a global context (Alan Rogers). As was highlighted earlier in this report, there was an understanding at the seminar that quantitative methods and data appear to enjoy a more privileged position than data and methods of a more ethnographic or qualitative nature. The relative position of ethnographic/qualitative data in this hierarchy is illustrated by the ways in which data generated from these methods and approaches is often referred to as 'anecdotal' and viewed as simple 'stories' of the world. Claims to 'truth', 'accuracy', 'fact' and 'science' appear only to be viable through statistically based projections and representations. The distance between different approaches to understanding aspects of the material world is too vast and longstanding for the seminar and its participants to have had a realistic hope of reconciling the differences. However, what the group was able to do was to move beyond personal preference and opinion to begin to look at the pragmatics and practicalities of the potentialities of a *"mixed methods approach"* (Aaron Benavot) and collectively stressed the importance of ethnographers and quantitative and qualitative data scientists working successfully together (Jay Derrick).

Stage Two

The discussions outlined above generated a range of topics and issues, all of which were and are still deserving of further time, thought and discussion. To maximise the use of time at the seminar, participants decided to focus the time that remained of the day on key areas under the headings of:

- the Contexts and Purposes of measuring literacies
- and the Units and Methods by which and through which literacy could potentially be measured.

The discussions overlapped on a number of occasions which is a result of the interrelated nature of the distinctions. The key goals of the discussions were to further our understanding of the ways in which ethnographic methods for measuring literacies could be used in co-operation with more traditional quantitative methods in ways that minimise ideological contradictions or contestations, and to develop a clearer understanding of what units and practices constitute literacy and would be accessible to research and necessary for understanding.

Contexts and Purposes of measuring literacies

The notions of context and purpose were felt to be central to the idea of measuring literacy in that the two are inseparably and inextricably linked. The ultimate motivations and purposes of collecting data determine and define what data is needed, where it will be sought and how it will be viewed and interpreted (Paul Black). Some of the contexts in which ethnographic data and approaches may be located as suggested in the discussions are outlined below along with an overview of the discussions that accompanied them.

Policies

Policies at the moment they are set in action upon the world conceal the processes, struggles, nuances and agents of their development. The untangling of these processes and the identification of involved parties was a suggested area in which ethnographers and social scientists might work (Constant Leung). A clearer understanding of the 'messiness' of the development process could provide us with a clearer understanding of the implicit designs and objectives of policies and their relationship to subsequent interpretations and uses of data. Working at the level of policy development is something that traditionally ethnographic researchers have been reluctant to do or uncomfortable in undertaking (Brian Street); however, engaging with policy at such a level may be a route by which a better understanding of the distributions of power may be gained. The design, development and implementation of policies and associated practices necessitate the involvement of a number of actors and agents. There are significant issues of power involved specifically in regards to the notion of choice. It is often the case that governments, institutions and individuals do not always occupy the position of decision-maker in being able to choose their own policies and practices (David Baker). Ethnography could potentially offer a better account of the landscapes over which policies and practices travel as well as an understanding of where and in what forms they are practised.

Practices

Another route to understanding literacy which is open to contributions from ethnography is in providing a more detailed understanding of both new and existing measurement practices (Alan Rogers). An ethnographic approach would seek to understand which literacies are actually being measured, for example; who are the people that surveys have previously targeted and encountered and who are the people that have previously been missed out? A natural follow on from this is in identifying what particular literacies and literacy practices are targeted in the surveys and which ones are overlooked. Understanding the nature of the data that is collected is an essential way of understanding the concept of literacy that is in use in such projects as LAMP and IALS. A richer sense of understanding the conceptualisation of literacy in such projects is an essential channel through which a better understanding of the measuring process and the collected data could be fed back to key agents within the project. Such a process is essential in understanding how, for example, the choice of language used for interviews, the interviewer themselves and the questions asked shape and define the ultimate picture or understanding of literacy that is being gathered. An ethnographic approach would bring to light the dynamics of interpersonal relationships and power structures inherent in the processes of measurement and how they impact upon the 'sense' of literacy that is developed.

Literacy

As was highlighted in **Stage One,** the participants felt that a more comprehensive understanding of what literacy actually is would be necessary for any attempt of measurement to be made. The contextually defined nature of literacy practices calls for a survey of what literacy is and means in the contexts of its enactment (Harvey Goldstein). As suggested by Clinton Robinson, it is impossible to separate the contexts from the purposes of literacy. Any assessment of literacy requires the isolation of key contextual factors in the *"literate environment"* (Clinton Robinson). An ethnographic approach would be needed to access such data but would be insufficient in terms of comprehensiveness and would therefore require quantitative data and skills to gain a fuller picture (Harvey Goldstein).

Pedagogies

An essential area from which a better understanding of the contexts of literacy or the 'literate environment' could be understood is through looking more closely at the different ways in which literacy programmes are being undertaken in different contexts (Brian Street and Alan Rogers). What are the dominant pedagogies of particular programmes and how did they come to be so? Such questions would provide an understanding of literacy at a 'grassroots' level and would offer an insight into the ways in which the purposes of literacy are conceptualised and shape pedagogic practice. Again ethnographic methods and qualitative data would be insufficient to produce a full picture and would require input from quantitative research.

Contexts

Nigel Norris provided the group with a valuable summary of the contexts in which ethnographic and qualitative methods might be used in order to understand and potentially inform some of the key considerations and objectives in the context of policy.

1) ***Targeting of resources***: a valuable pathway to understanding the purposes and goals of policies is through understanding the routes and methods by which they access funding and monies. Policies need to secure capital in order that their development and enactment can take place. Key stakeholders and their impact upon the nature of policies would be necessary to understand fully the nature of a particular policy. Understanding the formula by which policies access funding is a potentially rich route by which to understand the ways in which the policies and their agencies subsequently distribute funds.
2) ***Informing strategy***: working at the level of policy development, 'feeding-up' contextually based data could be an area in which ethnographic data and methods can be utilised in informing and shaping programme design. One of the key concerns at the seminar focused on establishing ways in which power to inform and shape policy could be gained or more equally shared by ethnographers and qualitative data scientists. This would involve ethnographic/qualitative practitioners and researchers working at a policy level, and a significant aspect to this relationship would be the fact that they would be working with 'their own' data as well as that collected via alternative but complementary approaches.
3) ***Measuring effectiveness***: policies represent the objectives, wishes and capital of a variety of sources and are therefore tied to their donors, investors and stakeholders in relationships likely to be guided by issues of performativity and accountability. Ethnography and qualitative data can play a vital role in accessing and providing information about such relations. A by-product of this would be a better understanding of the ways in which top-down pressures interact with policy and on-the-ground practices.
4) ***Justifying the level and distribution of expenditure***: The role of alternative approaches to quantitative data and research in this context is similar to the one outlined above. An additional dimension to this however, is an insight into the ways in which funding decisions are made and carried out. Within this space, there is scope for the data and insights provided by ethnography and qualitative data to direct funding and inform the criteria upon which it is allocated.
5) ***Political resources***: The ways in which policies are understood and intended to function at the sites of their conceptualisation are not always the ways in which they are understood and utilised at the sites of their practice. Governments and agencies are often unfamiliar with the ways in which policies and programme operate in their lived contexts, and this is another space in which ethnography and qualitative data could be valuably used.

Comparisons

The macro-level objectives of measuring literacy in the light of the aforementioned projects are to produce data for cross-contextual comparisons. It was suggested that the comparisons themselves may become fruitful sites of investigation, analysis and the positive employment of a mixed methods approach.

Some of the key contexts and purposes in which qualitatively based data may be validly used which were identified in these discussions can be summarised as follows,

Contexts and Purposes

- Producing detailed accounts of the peoples and practices of different literate environments.
- Enriching conceptualisations of literacy through examining the ways in which current survey practices, policy stakeholders and macro-level objectives influence and determine data.
- Developing a richer understanding of the ways in which policies actually operate at the level of contact.
- Using contextually derived data to mediate relations and decisions at all stages of policy development and planning.
- Using a richer understanding of literacy practices to encourage a more sensitive understanding of the dynamics of the literate environment, highlighting some of the limitations and implications of the act of comparison.

Units and Methods of analysis

Identifying the units of analysis and the methods by which measurement and subsequent understandings of these might be achieved was felt to be a key process in identifying the role that ethnographic and qualitative methods might play in the measurement of literacy. One of the key objectives in these discussions was to identify spaces in which ethnographers and qualitative data scientists could work to both understand and improve the effectiveness of current measuring techniques, and to consider ways in which different approaches to and paradigms of measurement could work together in the processes of seeking data that is able to be used in making comparisons.

Units

One of the principal motivations for identifying units of analysis for ethnography to focus upon comes from the collective drive to understand literacy in its widest possible contexts and to offer these understandings and conceptualisations as enrichments to current definitions and understandings. Ethnography would seek to identify the various levels or strata that constitute the 'literate landscape', highlighting the practices and individuals involved in community literacies, household literacies and family literacies as well as educational and professional literacies. Any measurement of literacy has to include more than data on formal, national and dominant literacies and must include those practices outlined above, practices which the group felt have often been overlooked or absent from previous measures and perceptions of literacy. An account that offers data on only the numbers and percentages of people involved in literacy practices would fail to recognise the ways in which the domains of literacy interact with other socio-cultural practices and phenomena, and the interpersonal and social relations that guide practices could be overlooked or omitted. An ethnographic approach would seek to understand these interrelations and show how they are essential to an understanding of how policies are played out in reality.

Indexes were also suggested as representing units of analysis that would be a highly profitable area to focus upon (Mary Hamilton). Multiple indexes are commonly used in quantitative approaches to gather and interpret data, whereas qualitative approaches tend to use description and classification in order to delimit units of study. In the context of understanding and informing the measurement of literacy, an understanding of the ways in which practices are indexed and classified could provide an insight into the differing views and conceptualisations of literacy (Harvey Goldstein). Understanding the indexes of various approaches could provide opportunities to better understand the ways literacy is understood by the different approaches and in the development of mixed-method indexes (Mary Hamilton).

Methods

It is possible to suggest that it was in the drive to deconstruct the divide between quantitative and qualitative methods of research and analysis that the seminar and participants made the most fruitful and dynamic progress. One of the predominant criticisms or suspicions that ethnographic or qualitative approaches level at the door of quantitative approaches is that the complexities of the social world are subject to simplification when interpreted and presented mathematically. 'Noise-free' statistics (Harvey Goldstein) become powerful and convincing forms of representation often used to support and direct a variety of policies, interventions and pedagogies. At the same time it was recognised that ethnographic and qualitative data is open to criticism on a similar level in that they could produce accounts of things such as literacy that appear as or become reified 'ideal types' (Brian Street).

Any ethnographic or qualitative analysis of events and practices involves implicit acts of classification and assessment, and it is within such processes that data is assigned or given its particular category of meaning. The acts of measuring and classifying are in themselves crude measurements (Harvey Goldstein) and analytic frameworks that are open to criticisms of subjectivity and simplification. Failure to document or make explicit the grounds and reasoning for particular classifications arguably make any ethnographic representations as 'noiseless' as their quantitative counter-parts (David Baker). All approaches are therefore charged with the same duties of making explicit any implicit 'observations' and 'calculations' that are made. The ways in which data is manipulated in the process of analysis and presentation could provide a profitable insight into the points at which the boundaries between various methods of data collection could be blurred (Jane Freeland), and a method by which mutual understanding and co-operation within and across approaches could be forged.

A significant area in which the discussions focused was in relation to the concept of comparison and the methods by which contextually relevant measurements of literacy practices might be obtained. It was suggested that the levels or orders of comparability upon which measurement practices operate will significantly influence and determine any subsequently obtained measurements (Harvey Goldstein). It is widely assumed that comparisons are generally made between average values of particular measured outcomes. In relation to literacy, these comparisons may be in the form of test scores, participation figures, expenditure etc… However these 'types' of comparison, as previously outlined, do not necessarily result in the comparison of substantively similar things. As suggested by Harvey Goldstein, these particular types of 'first order' comparisons are only achievable through the use of statistics and are almost impossible to support with empirically verifiable data.

'Second order' comparisons were suggested as a more appropriate and operational level at which to work. 'Second order' comparisons involve a study of the ways in which units of measurement interact and relate to other aspects of culture and society, e.g., literacy practices and education. If literacy practices are studied in relation to other things, there is scope and potential for comparability if these same related units are studied in alternative contexts. A fuller discussion of these orders of comparability and ethnographic and statistical models for social realities has been provided by Harvey Goldstein and is attached to this report in Appendix Two.

Some of the key units for analysis and methods with which qualitatively based approaches may profitably work that were identified in these discussions can be summarised as follows;

- Units of analysis need to be contextually derived and cannot be solely generated out of context if any comprehensive understanding of the reality of literate environments is to be had.

- All aspects of the literate environment including literacy development programmes, policies, policy practices and stakeholders are considered to be important units of analysis.
- The goal of comparability could potentially be realised through using lower-order levels of comparison which would require a mixed-method approach to ensure that the units of comparison are consistent over a range of contexts.
- Differing approaches to measuring and identifying literacy practices can be reconciled by a richer understanding of their inherent strengths, weaknesses and biases.
- Any form of classification or measurement is informed by conceptual and analytic constructs which need to be explicit at all times. The act of making these decisions explicit is a source from which a better understanding of the ways in which literacy is conceptualised by different agents in the data collection processes and analysis can be achieved.
- The development of a more sophisticated method of measurement, informed by quantitative, qualitative and ethnographic approaches could result in the development of better policies.

Stage Three

The final, and shortest, stage of the seminar was a case study. Looking at various statements about literacy statistics including those in the Global Monitoring Report, it is often stated that *"The large bulk of illiterates are women"*. This example was taken to look at and discuss *"How ethnographic approaches to measurement might deal with issues of gender in literacy?"* (Alan Rogers). Below is a list of points that individual participants made in relation to this statement.

- Some immediate questions that an ethnographic approach would seek to answer and clarify through empirically derived data, as posed by Paul Black were, *"Are the measurements wrong or biased? What is it about the tools used which created, generated and/or influenced such a gender spread? What is the range due to?"*
- Jane Freeland suggested that such a measure of variation is likely to be determined by sex and not gender. Ethnography could therefore be utilised to understand gender relations in terms of access to social goods and capital. A key area of understanding to be developed in relation to such a statement would be to understand social theories behind gender.
- Harvey Goldstein used examples such as the 11+ and IQ test developments to illustrate how test design and outcomes can be gender biased and how relatively simple changes can alter and shift results in radical ways and suggested that validation of tests can only really occur through reference back to the populations tested.
- Brian Street offered the Indian literacy programme 'Nirantar' as an example of the ways in which classification procedures lead to serious distortions of data and figures relating to women's literacy.
- Kathy Machlaclan suggested that claims to validity of such statements as "gender is or is not an important literacies' issue" need to be supported by ethnographically derived data. It is through researchers' observations and understandings of the lived experiences of women and men that more 'accurate' pictures of literacy and literacy practices can be developed. The value of an ethnographic approach is that gender can be used as a lens through which one can view the social landscape rather than as a

category or index for assessment. Traditional quantitative approaches to gender risk reducing it to a series of binary oppositions that can be 'easily' assessed and measured rather than viewing gender as representing a complex array of multiple perspectives which are necessary to understand how social practices and events are constituted and viewed by on-the-ground participants.

- Aaron Benavot gave the example that it is often argued that men have a greater access to languages of the wider community and higher status codes than their female counterparts. Women tend to be associated with local, mother-tongue languages of limited circulation and status. Ethnographic approaches could offer an understanding of the mechanisms in place which promote such spreads and distributions and identify opportunities for wider access for women.
- David Baker suggested that concepts such as 'women' obscure who the populations actually are. Within any population there are multiple bodies and an ethnographic approach would seek to identify who is contained within the population. Who are the various women which the statement reifies into one homogeneous unit? What are their lived experiences and differences in relation to literacy?
- Priti Chopra suggested that such statements lead to conceptualisations of women that fail to recognise the actualities of everyday lives, experiences and practices. Ethnography would seek to recognise that their lives are linked to factors beyond labels.
- Harvey Goldstein suggested that such statements demand a need for information about the allocation of resources and how these are distributed within populations.
- Alan Rogers suggested that it is necessary to recognise that learning of literacy skills very often takes place outside of school contexts and there is therefore a need to examine the range of literacy practices that are available and develop an understanding of how the range of literacy programmes, teaching and acquisitions may be gendered.
- Paul Black suggested that an understanding of how gender impacts upon or affects things would require multivariable analysis with the main priority being the avoidance of any type of correlation and cause assumptions.

Summing up, Alan Rogers suggested there were two possible approaches which ethnography could provide to such a simplistic statistical statement – it could challenge the validity of such generalisations, and it could elaborate such statements to provide more detailed explanations rather than simplistic causal relationships.

What might ethnographic approaches offer to the measuring of literacy?

The discussions that took place over the course of the day's seminar provided the opportunity for an exploration of a wide range of key issues, concerns and hopes for the ways in which ethnographic approaches may contribute to the measuring of literacy practices. Key ideological and epistemological concerns and conflicts were given a space to be voiced, and an opportunity for considering ways in which these may be overcome was provided by the contributions of the group.

For my part, I feel that what came through in the course of the seminar was the idea that an ethnographic perspective can play a crucial role in keeping at the forefront the whos, whats, wheres, whens and whys that are often obscured, marginalised or rendered neutral in common concepts, practices, labels and assumptions. An ethnographic approach to measuring literacy would seek to avoid operating at the level of abstraction and reification through anchoring both *qualitative and quantitative data to empirically supported contextualisations. Such an approach would raise awareness to the idea that literacy practices are not consistent over time and space and that any form of comparison risks the possibility of comparing like with unlike. The dangers of such misrepresentative comparisons have a variety of serious implications for individuals on-the-ground as well as for the development of policies and their subsequent enactments.*

One of the most encouraging aspects of the seminar was that a precedent has been set in which, to use Aaron Benavot's words in a slightly different context, quantitative and qualitative approaches, whilst remaining strange bed-fellows, can at least be seen as living in the same house. What such a view enables is the idea that any approach has its own inherent strengths, weaknesses and biases. It is necessary to recognise that a preference towards one particular approach will result in data and conceptualisations that are shaped and defined by these very characteristics. The goal of understanding and representing the states and variations of literacies across and within different contexts and boundaries and correlating figures of measurement can only realistically be achieved if approaches are drawn together and co-participate at all levels of operation.

In writing this report, I have tried to represent the nature and content of the discussions as accurately as possible. However, it is necessary to recognise the role that my own interpretations of the discussions will have played in the ways in which it has shaped this report. I would like to thank all of the participants for their contributions to the seminar and to this report and hope that the Uppingham Seminar Series can continue to draw such a wide range of participants to continue such positive and productive discussions in the future.

Appendix One

UPPINGHAM SEMINARS
FOLLOWUP SEMINAR ON ETHNOGRAPHIC APPROACHES TO MEASURING LITERACY, 3 MARCH 2005

BACKGROUND PAPER

Following the Uppingham Seminar of November 2003 on Ethnographic Approaches to Measuring Literacy, the report of which is available on our website (www.uppinghamseminars.org[1]), a one-day follow up seminar has been arranged.

The following comments have been culled from comments provided by the following:

- Carolyn Peterson from Edinburgh University who wrote the report on the November 2003 seminar (currently in South Africa)
- Cathy Kell from the University of Aukland
- Clinton Robinson, freelance consultant
- Constant Leung from Kings College London

all of whom were at the earlier Seminar.

Their comments have been brought together by Alan Rogers into a draft agenda – this is open to negotiation at the opening session of the Seminar.

ISSUES ARISING FROM THE NOVEMBER 2003 SEMINAR AND FURTHER CONSIDERATION.

1. CONCEPTUALISING THE FIELD

Is it adequate/appropriate to express the distinction in terms of a dichotomy of quantitative versus qualitative/narrativising approaches? Among the distinctions so far identified are the following:

individual action - collective/joint activity
individual language - international discourse
skill in language - language-in-action
product - process
mono-construct - multifaceted phenomena
formal - realist
conventional - critical
efficient (low unit cost) - inefficient (high unit cost)
economic rationalism - human dignity
stability - contingency

Another might be the needs of practitioners and those of the managers.

(For the sake of convenience, I will use the terms quantitative and ethnographic for the distinction under examination, although these terms may be rejected by the seminar participants).

[1] Copies can be made available if requested.

2. HOW SHOULD ETHNOGRAPHIC ASSESSMENTS (?is this the right term?) BE EXPRESSED?

e.g. by single figures? how do we express multiple literacies?

Can numeracy be included or do we need a separate assessment for numeracy?

Would the examination of the 'literate environment' as an expanded concept be a better means of understanding contextual factors as parameters of measurement? If so, how can such an examination be expressed?

3. RELATIONS BETWEEN ETHNOGRAPHIC APPROACHES AND CONTEMPORARY QUANTITATIVE APPROACHES:

Do the two approaches operate at different levels (macro-micro) and different scales (e.g. local and national/regional)?

Should ethnographic approaches be included in current quantitative literacy assessments such as LAMP or be pursued separately?

Are there different purposes being pursued? –
e.g. current large-scale quantitative assessments to measure school educational attainment and ethnographic approaches to measure development outcomes?
or ethnographic approaches to identify 'models of' literacy/ies and quantitative to develop 'models for'?
or ethnographic approaches to enhance the practice and quantitative to develop policy?

What **challenges** do ethnographic approaches make to current large-scale quantitative literacy assessments - e.g. what can they contribute to the Global Monitoring Report?

Equally, what challenges do quantitative approaches make to ethnographic approaches?

4. COMPARABILITY

Do we need to settle for comparing like with unlike when measuring literacy across contexts?

5. SHOULD WE BE ENGAGING IN THIS DEBATE AT ALL?

It can be argued that the current emphasis on quantitative assessment is part of the contemporary 'audit culture'; should practitioners and researchers tacitly accept, tolerate, challenge or ignore such a culture? Do we **need** to try to influence providers of literacy learning programmes? and if so, at local (i.e. practice) level or at policy level (e.g. campaigns) also?

I hope I have represented the comments made by our commentators fairly. To ensure that all of you get the opportunity to read their comments in full, they are attached.

Alan Rogers
11 February 2005

CAROLYN PETERSON

Suggestions for discussion in the follow-up:

1. Contribution of ethnographic methods to large-scale literacy assessments; and to policy on large-scale assessment:

LAMP – comments on process and outcomes of consultation / participation of the UppSem group in LAMP; plans for further action if required
Any current or forthcoming large-scale assessments and opportunities for comment on contribution of ethnographic methods (e.g. EFA Global Monitoring Report 2006 – how is literacy/-ies and numeracy being assessed? Is literacy assessment being used as an indicator for school educational attainment or for development outcomes? How will it be presented e.g. in a single figure per country? How far is the multiple view of literacies being used / incorporated?)

2. Multi-lingual literacies – particularly equity and comparability issues in literacy assessment:
Comments on any progress of the group so far plus future conferences and strategies
Multi-lingual literacies in large-scale literacy assessments (see above)

3. Literacies AND numeracy – how could these be better integrated in literacy assessment? How could ethnographic methods contribute?

CONSTANT LEUNG

I have re-read the extended notes produced by Carolyn and am beginning to feel that one key question that may follow quite naturally from the Nov 2003 meeting is: what is/may be the division of labour in conceptualizing measuring literacy from quantitative and qualitative/ narrativizing perspectives?

And to begin that process of discussion, I thought some bipolar pairs of descriptors may help oil the wheel of thinking, for instance:

- individual action - collective/joint activity
- individual language - international discourse
- skill in language - language-in-action
- product - process
- mono-construct - multifaceted phenomena
- formal - realist
- conventional - critical
- efficient (low unit cost) - inefficient (high unit cost)
- economic rationalism - human dignity
- stability - contingency

There are undoubtedly many more such pairings. Given that we will be talking to colleagues from a range of disciplines, I think any attempt to give an account of literacy (measurement, description etc) would probably benefit from all participants sharing a sense of understanding in terms of their own position/s.

CLINTON ROBINSON

Two issues to consider further

The two issues mentioned in the announcement of the follow-up seminar – ethnographic approaches to assessment and comparability of assessment – continue the debate of the 2003 seminar, where the role of ethnography and the need for comparability found vigorous protagonists on both sides. These issues are indeed at the heart of the dilemmas which attempts to measure literacy lead us into, and I would like to raise two issues which impinge on that debate:

- first, with regard to comparability, is it inevitable that we must agree to compare like with unlike when measuring literacy across contexts?
- second, with regard to ethnographic approaches, could the examination of the literate environment, as an expanded concept, be a means of better understanding contextual factors as parameters of measurement? This question is closely linked to the first one.

Comparing like and unlike

The 2003 seminar agreed that the key questions in the need for comparison of literacy levels were the 'who' and the 'why'. Who is calling for a comparison across different contexts and what is the purpose of the comparison? There was some agreement that the needs and purposes of international agencies were different from those of programme managers and practitioners and that measurement undertaken for the former would probably not serve the purposes of the latter, and vice versa, whatever the methods used. The strong calls for consideration of context also underlined the difficulties of any meaningful comparison.

Putting aside for a moment the fundamental question of whether there really is any need for international comparison, or indeed for any comparison across different contexts within one country, is it in fact inevitable that any attempt at comparison must turn a blind eye to context and therefore compare like with unlike?

With the increasing understanding that literacy must be assessed in context, the literacy performance standards set by countries use national criteria, defining certain levels and targets. Across countries, therefore, what is being compared are national understandings of literacy levels, not any absolute standard set internationally. LAMP seeks to get around this by setting some absolute levels, but taking account at the same time of contextual factors, such as the language of literacy. However, as soon as any level is set for assessment across contexts, crucial contextual factors are lost, such as the purposes of written communication and the opportunities which certain literacies offer. In this perspective, to seek to compare like with like ends up decontextualising literacy performance and ignoring the unlike. What are in fact the like and unlike in literacy? Can literacy be sufficiently abstracted from context to say in any meaningful way that we are comparing like with like? Can we rather, or should we, compare contexts, if we want to undertake comparison at all? This would mean identifying key parameters in each context which shape literacies, their acquisition, use and consequences. This leads to the second issue.

Examining the literate environment

The concept of the literate environment may serve as a means of structuring our understanding of how literacies are acquired and used in context, and has the potential to be expanded to include factors such as institutional aspects, production of written communication, purposes of literacy, and language issues. Could therefore a more systematic description and analysis of the literate environment facilitate some kind of comparison across contexts? Could parameters of sufficient similarity be isolated so that comparison would be warranted? Although developed primarily for the interplay between language and literacy, could the dimensions of Hornberger's biliteracy framework be useful in this regard?

Could literate environments be characterised in some way? It might be valuable to look at certain types of environment (necessarily overlapping), defined by languages, availability of (types of) literature, individual and institutional use of text, modes and opportunities to acquire literacy, and so on. It is because context matters so much for assessment that the concept of the literate environment deserves further development and closer analysis.

In terms of comparability, an analysis of the literate environment can get at precisely some of those factors which make for differences in the use of literacies. To take the example of language, the literate environment is very different where a small minority language is concerned, indeed totally different from those who operate only in a majority or mainstream language. This distinction must <u>always</u> be signalled in any assessment. Similarly, with regard to the acquisition of literacy, situations differ in major ways according to whether schools (which ones?) produce a large majority of learners with the capacity to use literacy easily, critically and creatively, and those that do not. It is for example meaningless to compare literacy levels of the following two groups, without a full understanding of context:

- the young people of urbanised elites in India, who have attended English-medium private schools, whose dominant literacy is in English, with adequate competence in the language of the state of residence and whose home environment includes both the state language and English;
- the young people of an Indian tribal group who have attended a local school, with teaching of low quality, and where the state language is the medium of instruction. They have had little or no chance to develop literacy through/in their own tribal language, the one used predominantly in home and community.

In such situations, the modes of acquisition, languages of literacy, and opportunities through literacy are totally different. Comparison of literacy levels must necessarily take account of these differences of literate environment, with the added complexity of the organic and changing links that exist between the two contexts. We should note also that the approaches and methods of literacy assessment are themselves pertinent aspects of the literate environment, as part of the institutional framework and the modes of acquisition.

CATHERINE KELL

In my view, three issues emerged from Uppingham which may provide useful points of reference in the current seminar. I will refer to the results of the Uppingham Seminar outlined in the report as "Uppingham". I will state the issues and then touch on each of these in order to explain why I think they are important. I hope I am not straying beyond my brief, and hope that you will accept these comments in the style of a polemic and in an attempt to stimulate debate!

First, Uppingham claimed that there are three main ways in which ethnographic and qualitative approaches can contribute to assessment of literacy: through influencing policy; through large scale assessment and at the micro-level. We did not really discuss the first (influencing policy) in any depth and I would like to suggest that this needs to be opened up to scrutiny. I will refer to this as the "levels" and "scales" issue.

Secondly, I suggest that the notion of "models of" and "models for" may open up a productive area of inquiry, and draws on some of the bi-polar descriptors that Constant has suggested. This also intersects with the levels and scales issue.

Thirdly, the broader historical and global canvas on which this debate is being played out. Here, I would like to refer particularly to the role of academics and of research in an era of "audit cultures" generally. Uppingham seemed to tacitly accept or tolerate the measurement paradigm and seek to find ways of engaging with it. This is very important but as academics, perhaps our key responsibility lies in researching further, the questioning of and articulating of the problems with the paradigm.

<u>Levels and scales</u>

The fact that Uppingham did not address the level of influencing policy is regretful. (I take as a given the importance of working to influence the micro-level [which is relatively straight-forward], and the large-scale which is hugely difficult). But each of the levels feeds into the other in both directions, and if large scale assessment or micro-level work is not engaging with policy, then it is difficult to understand what the point of it is. Surely large-scale assessments are aimed at influencing policy in themselves, particularly financial policy. But the really hard question is how does the work of ethnographers (let's call them the "in-dwellers") influence the work of the measurers (the "bean counters") and therefore the legislators (to use Zygmunt Bauman's broad term).

Of course, this comes back to the question of levels and scales and sequencing. If the in-dwellers have influence at the micro level and the large scale level "we" may be able to have influence at the broadest policy level – but what if the paradigm at either of the lower levels is wrong in the first place? We need to engage with and question at each of those levels all the time.

Mignonne Breier, one of the co-ordinators of the Social Uses of Literacy project in South Africa, claimed that ethnography at this stage cannot influence policy directly, but it can influence policy discourse. I think this is a helpful distinction. Not being able to influence policy directly is related to a number of issues. Worth mentioning are

- first, the problems of capturing, interpreting and generalising in-depth interview and observation material, and
- secondly, the issue of engaging effectively with policy-making in an attempt to make it more beneficial while at the same time attempting to maintain a critical distance from its 'disciplinary' effects (see Kell 2001 and Spiegel et al 1999).

This is a huge question (surely there are numerous studies looking at how quantitative and qualitative paradigms intersect); and reviewing these and exploring the specifics of this debate for adult literacy work may well be the central focus of the current seminar and further work for Uppingham Seminars and others.

When engaging with policy discourse, two provisos are necessary. First, it seems to me that the bean counters across a sweep of fields (not just adult literacy) focus only on outcomes. (This is related to the audit cultures issue and is an example of engaging at a level where the overall paradigm is wrong). In a field as complex as adult literacy, this seems to me to be completely the wrong emphasis. (One) As we all know, the outcomes can be unrelated to what they purport to be indicators of (see Strathern for an eloquent description of this). (Two) I believe strongly that in the adult literacy field the emphasis should be on inputs (see Archer for some sensible ideas about inputs). I believe we should advocate for a moratorium on measuring outputs in adult literacy until we have some agreement on what inputs are necessary for outcomes to be reached and some indication that those inputs are in place!

Secondly, another area where engagement at the middle level (large scale assessment) is problematic is the coherence of the concept of the "measurement of literacy" itself. It does not make sense to me to measure literacy. It does make sense (if inputs are in place) to measure the number of hours spent in a literacy programme. Then perhaps useful decisions can be made by the legislators. This is consistent with my argument made in various places that we need to separate out literacy and an adult education system, and we need to think about these both together and separately.

As a final point on engaging with policy discourse, I think that some of "our" efforts should go into "in-dwelling" in the corridors of power in order to understand more about how the legislators draw on the work of the bean counters and where the problems are with this. Again this applies across a sweep of fields, not just adult literacy. Basically, drawing on Latour and other's work, I am interested in how what is called "facticity" is established. And above all I am interested in its costs and in its effects. I would like to know all about how much it cost to bring out the EFA Global Monitoring reports and what their effects are. Some of you may know that what I am referring to can be called "ethnography-up".

"Models of/ models for"

When I raised this at Uppingham, I referred to Goffman's use of this term. I have searched for the reference but cannot find it. These are my thoughts:

Ethnographers of literacy have provided "models of" adult literacy. They have and are making serious attempts to use these to provide coherent and consistent "models for" (ie the implications of ethnographic findings for policy and provision, see for example the work of the Lancaster group of NRDC, my own small studies in South Africa, Robinson Pant's work, Street's work over the past twenty years). The **relation** between models of and models for is of course hugely interesting and could be studied as a part of the levels and scales above.

The legislators and bean counters are providing powerful and persuasive "models for" (sometimes by default – eg the effects of IALS), but they have not provided coherent and consistent "models of", despite token nods to the idea of "literacies" etc. In fact, the understandings and definitions of adult literacy that they are working with are generally incoherent. This state of affairs is entirely consistent with what is happening in schooling and in many other fields, like building construction, social services and health – tick off the outcomes on the checklist but don't check the inputs. In New Zealand in the building construction industry, the result of this approach has been rotting buildings!

If we are to grapple seriously with influencing large scale assessment and influencing policy (taking as a given the necessity for diversity and context-specificity in adult literacy work), and before we can make normative claims about what can be done, we need to have coherence and consistency in both our "models of" and "models for". I believe that the bean counters don't know whether or not they are using apples to measure pears, but they are making claims about fruit.

I take the micro level area of work as fairly unproblematic on this count.

The role of academics

Finally, in a time of shifts from what have been called Mode1 to Mode 2 knowledge production, I believe that decisions about forms of engagement need to be made very carefully. Related to Uppingham's concerns, there is a spectrum of forms of engagement.

Disengagement sits at one end of this spectrum. In this position, the in-dwellers would only work at the micro-level mentioned above, believing that their work can only contribute to the development of accountability and appropriacy at detailed programme and case study level and that any forms of generalisation are too difficult at this stage. The feedback loop is a long one, and it travels the usual academic routes of research studies, peer-reviewed publications, finally entering the public domain. Mode 1 knowledge production.

Jumping to the middle of the spectrum, there is the position that many of us are in at the moment. A kind of rapprochement. Engaging cautiously, feeding into LAMP and other initiatives, testing out the waters of perhaps transforming measurement instruments and processes, but the promise of having effects. This is a shorter feedback loop but it may be a compromising position.

At the opposite far end of this spectrum, there may be two positions. One is classic Mode 2, involving going into policy and assessment forums and organisations and using the language and results of ethnographic research to mitigate against the problems generated by large scale aggregate and individual forms of measurement. This is a short feedback loop.

A different position would also be a strong engagement but to use the ethnographic approach to "study up", perhaps work with those in the corridors of power to do ethnographic evaluations of their own work processes, uncovering problems in the establishment of "facticity" and so on.

Appendix Two

Ethnographic and statistical models for social reality

These notes by Harvey Goldstein present some thoughts on the nature of measurement in both ethnographic and survey contexts, the issue of how social complexity is described, and some thoughts on the issue of 'comparability' across cultures and societies.
(*following the Uppingham/ King's Seminar on Literacy and Measurement, held at DEPS, King's College London on 3.3.05*)

Kinds of measurement

Measurement is often construed as the kind that is involved in most mental testing, essentially providing a number located on a (semi) continuous scale that is used to characterize an individual, group or institution. More generally, however, statisticians recognize different kinds of measurement ranging from those like length that have a well defined zero and are truly continuous, through to those that are simply classifications of individuals or objects into different categories.

I would suggest that measurement in ethnography can be thought of most usefully as the latter type whereby the descriptions applied to individuals, groups etc. (let me refer to them as 'units' for simplicity) can formally be thought of as classifications of those units. The labels or descriptions attached to the categories that form a classification represent attempts to capture some (hopefully) theoretical construct that underlies the process of description, for example of types or 'levels' of workplace literacy. As with all kinds of measurement, there will be uncertainty surrounding the allocation to a category, and it may even be possible to express this numerically (e.g. the probability that unit x belongs to category A rather than category B, or individuals may 'belong to' more than one category). Such 'nominal' scales of course are a logical abstraction and represent a deliberate choice about what information to ignore, again as with any kind of measurement – but this is implicit in ethnographic recording and description.

The point is that we can manipulate the information in the classification to obtain further useful descriptions. This may simply be a matter of counting how many are in each category and perhaps making comparisons of percentages across groups or across time. It could also lead to a more elaborate 'modelling' of relationships and there is a great deal of literature with examples of this. Naturally, any inferences that are drawn as a result of such 'manipulations' are only as good as the information they are based on, including all the assumptions about how to classify etc.

The paradigmatic mental testing measurement tends to use particular ways of obtaining its information and then manipulating it and in the course of this making assumptions about the ways in which reality is captured by those measurements and their manipulations that can be debated. The major difficulty is that the techniques used have evolved to such a level of computational/mathematical complexity that it is very difficult for most practitioners to understand the essence of what is being done, and in many quarters, such as those where international comparative studies of literacy are carried out, it is in the professional and indeed commercial interest of those concerned to maintain impenetrability. (see http://www.aea-europe.net/page-164.html and click on 'Harvey Goldstein' for an elaboration of these issues).

Accounting for real life complexity

One of the difficulties with traditional statistical 'models' applied in the social sciences, has been that heir relative simplicity cannot match the real world complexity. It is just that complexity that ethnographic approaches seek to capture. While statistics performs well in distinguishing the performance of agricultural crops or even drugs, the assumptions that have typically gone into modeling (e.g. using multiple regression) educational systems have been highly simplified. Thus, for example, the standard item response models (so-called IRT) used by psychometricians to scale and provide cross-country comparisons in IALS, PISA and the like make certain assumptions (e.g. uni-dimensionality) that are not only a priori unreasonable but can be shown empirically to be false. What emerges, therefore, are comparisons that are difficult to interpret because they are not anchored in any substantively interesting theory or view of the world. I also believe that the political requirements to provide 'objective' comparisons of literacy lead inevitably to these kinds of over-simplified statistical procedures, because of the need to avoid taking account of diversity and complexity.

More recently, many quantitative social scientists have been developing statistical modeling techniques (multilevel models and to some extent 'structural equation models') whose aim is to provide a framework that is elaborate enough to capture previously unexplored complexities. This has been aided by modern powerful computing, and while still at an early stage of development has shown considerable promise and is widely accepted as an appropriate approach. As it develops my view is that it will bring about a real possibility of dealing with levels of complexity commensurate with those studied by ethnographers. (see http://www.mlwin.com/hgpersonal/Professorial%20address%20Bristol%20Oct%202002.pdf)

Comparability

The common assumption when this term is used is that any comparison, for example between countries or groups within a given country) is between average values on achievement or performance scales. I would call these 'first order' comparisons. Not only are these very difficult to interpret (especially since almost all the studies use cross-sectional rather than longitudinal data so that any observed differences cannot simply be ascribed to schooling variations per se), but they are rather uninteresting.

Of more interest are what I call 'second order' comparisons. For example, the strength of the relationship between 'literacy' and formal education level may vary between countries and to study such an issue one does *not* need to use the same instrument to measure literacy in each country, although one would need to be satisfied that in each country the instrument used (which could of course be 'ethnographic') was measuring the appropriate literacy, but does allow local definitions of literacy to be used.

Another example would be a study of the extent to which achievement varied across schools, which might be interpreted as a measure of educational homogeneity. The requirement here to move away from the imposition of a single measuring instrument to allow locally specified procedures is attractive. It avoids the reliance on over-simple psychometric assumptions and gives more power to local expertise.

One can anticipate many who will view this with apprehension, since it would seem to remove the ability to define definitions of literacy etc., from those who currently control comparative studies – so the scientific debate inevitably will be paralleled by a political one.

Harvey Goldstein
4.3.5

Elaboration in correspondence between Brian Street and Harvey Goldstein, March 2005:

Brian Street to Harvey Goldstein:
Harvey,

Many thanks for these very helpful notes – and for your creative and considerable contribution to the seminar. I look forward to further discussions – and indeed to that Cross London Seminar with Mary.

You mention in your notes that: “One of the difficulties with traditional statistical ‘models’ applied in the social sciences, has been that their relative simplicity cannot match the real world complexity. It is just that complexity that ethnographic approaches seek to capture”.

I would say that ethnographic approaches are also ‘simplifications’ – which is of course how models work. Whether ethnographic models are ‘nearer’ to the real world than statistical models is an open question and probably made more blurred by the newer statistical modelling you talk about. I am attaching (see below) a piece I use on ‘case studies’ that I think captures some of the principles of case study methodology that would seem to apply to both ethnography and to what you call ‘second order’ comparisons. I look forward to further discussions.

Best wishes,

Brian

Harvey Goldstein to Brian Street:

Thanks Brian,

I was interested in the piece on case studies. I am worried by the last sentence. If you get a case study where your 'theory' explains things that were not previously explicable, that seems to me to be only the start of any explanation. First you would want to know why your particular theory does 'better' than any rivals - assuming there are rivals. Secondly, I don’t see how you can be satisfied with just one case study. How do you know that there are not other situations where it would fail? The only way you can do that is by choosing other situations, i.e. carrying out some kind of sampling so that you explore as many instances as possible. Thus in a sense you are surely just as interested in making inferences to a 'population' as the statistical sampler? In other words it seem to me that finding a 'good' case study promotes a hypothesis that needs further exploration and testing.

Now, in reality, we often have information that is relevant to the 'generalisability' of a case study finding. This may result from other established results, or may depend on theoretical notions or assumptions. They may, for example in the physical sciences, rely on the idea that physical events are repeatable rather than unique as many social events are. The medical example here is useful - case studies have an important role in generating hypotheses, but medical researchers also recognise the need for replication and experimentation to justify generalisability or theory generation.

Best wishes

Harvey

Brian Street to Harvey Goldstein

Harvey,

Thanks for those comments. I agree with you when you say: a 'good' case study promotes a hypothesis that needs further exploration and testing'. But the case study need not itself be 'generalisable', only offering insights at a theoretical level that might then be taken further. That offers other researchers a starting point for further investigation, some of which may indeed be about sampling, generalisability and statistical procedures. Jennifer Platt argues that 'there are sometimes good reasons for being interested in cases in their own right', although she mainly recognises that, if we want to generalise, then case studies can 'be useful in the early stages of research'. In this sense, case studies lay the basis for more statistical and generalising methods so that the outcome would be greater insight and knowledge derived from a mix of methods. And Hammersley takes up Mitchell's argument and like Platt, suggests 'the issue of generalizability to a relatively large population *is* relevant to case study work. He also suggests the distinction between enumerative and analytic induction may not be so great and that much qualitative and ethnographic research does implicitly if not explicitly assume comparison: we can improve and assess the generalizability of case study findings, which I think is the position you are suggesting. From the other direction, I would argue (as in the Literacy and Measurement sessions) that too often enumerative induction proceeds without having clarified the items to be generalized from – as in the international literacy measurements that assume 'literacy skills' as given and then correlate, analyse statistically etc before they have really clarified what that literacy starting point was in the first place. That starting pint can be clarified, or at least hypothesized, through the case study in Mitchell's sense of the term. When students use case studies in their research I always ask them to explain 'what is it a case *of*?' At a general level this might be eg of literacy as social practice eg of the role of gender relations in education. Having generated hypotheses in this way and clarified the concepts and categories with which we are working, the field is indeed open as you say to 'recognise the need for replication and experimentation to justify generalisability or [further] theory generation'.

Best wishes,

Brian

From **Mitchell,J** 1984 "Typicality and the Case Study" pp. 238-241 in R.F. Ellen (ed.) Ethnographic Research: a guide to conduct Academic Press: New York

Problems have arisen from the inappropriate extension to case studies of the procedure of making a particular kind of inference from quantitative data. When quantitative data are used to support theoretical interpretations, two different and unconnected types of inference are involved. The first is that of inferring that relationships observed in a sample of instances available to the analyst exist in the wider population from which the sample has been drawn. Sampling theory in statistics devotes itself to providing numerical estimates of the likelihood that the population values lie within some defined range or that established from the sample - provided that the sample has been chosen in such a way as to meet the mathematical conditions to justify the computation of the probabilities concerned. The sophistication and elaboration for choosing a "representative" sample in this restricted sense has overshadowed the other kind of inference involved when analytical statements are made from associations uncovered in a statistical sample.

This is the inference that the *theoretical* relationship among conceptually defined elements in the sample will also apply in the parent population. The basis of an inference of this sort is the cogency of the theoretical argument linking the elements in an intelligible way rather than the statistical representativeness of the sample: such inference is based on *analytical induction* rather than *enumerative induction* (These arguments are more fully set out in Mitchell 1983). It should be obvious that the inference from case studies is based on analytical induction. What the anthropologist using a case study to support an argument does is to ask how general principles deriving from some theoretical orientation manifest themselves in some given set of particular circumstances. A good case study, therefore, enables the analyst to establish theoretically valid connections between events and phenomena which previously were ineluctable. From this point of view, the search for a "typical" case for analytical exposition is ‘likely to be less fruitful than a telling’ case.