Comparative Education

Globalising assessment: an ethnography of literacy assessment, camels and fast food in the Mongolian Gobi

Bryan Maddox[a]

[a] School of International Development, University of East Anglia,
Norwich, UK
Published online: 01 Apr 2014.

PLEASE SCROLL DOWN FOR ARTICLE

Globalising assessment: an ethnography of literacy assessment, camels and fast food in the Mongolian Gobi

Bryan Maddox*

School of International Development, University of East Anglia, Norwich, UK

What happens when standardised literacy assessments travel globally? The paper presents an ethnographic account of adult literacy assessment events in rural Mongolia. It examines the dynamics of literacy assessment in terms of the movement and re-contextualisation of test items as they travel globally and are received locally by Mongolian respondents. The analysis of literacy assessment events is informed by Goodwin's 'participation framework' on language as embodied and situated interactive phenomena and by Actor Network Theory. Actor Network Theory (ANT) is applied to examine literacy assessment events as processes of translation shaped by an 'assemblage' of human and non-human actors (including the assessment texts).

Introduction

What happens when standardised literacy assessments travel globally? This paper examines what the ethnography of literacy assessment events can tell us about the process. As the work of Steiner-Khamsi (2012) suggests, ethnographic research at the local–global interface provides insights into the character and dynamics of globalised literacy assessment. This case study of literacy assessment in Mongolia is presented against a backdrop of increasingly ambitious, globalised projects of educational assessment. National assessments are integrated into larger international programmes, locating them within a 'global field of comparison' (Rizvi and Lingard 2010). That process involves distinct challenges of cross-cultural comparison and validity (Hambleton 2005).

Our setting off point is Hamilton's (2001) critique of psychometric literacy assessment. Hamilton applied Actor Network Theory (ANT) to describe how literacy assessment regimes privilege and frame knowledge about literacy (also Hamilton 2011, 2012). She proposed a new direction of research, studying literacy assessment regimes from the inside, including ethnographic investigation into the institutional processes of knowledge production, 'procedures of evidence gathering' and how assessment is 'enacted at the local level' (2001, 193). As Hamilton argued, the complexity of psychometric literacy data can easily become obscured and 'technical decisions and controversies' rendered invisible in statistical results (192).

This paper builds on those insights with a detailed ethnographic account of a national level adult literacy assessment that took place in Mongolia in 2010 as part

---

*Email: b.maddox@uea.ac.uk

of the UNESCO Literacy Assessment and Monitoring Programme (LAMP). The paper is in two parts. It begins with a detailed ethnographic description of unruly behaviour during a literacy assessment event in the Mongolian Gobi. Ethnographic transcripts are used to explore what people's behaviour during assessment can tell us about its cross-cultural demands. The second part describes the tensions that can exist between textual and contextual resources as they are employed by respondents in literacy assessment tasks. This is particularly important for the analysis of the cross-cultural validity of 'realistic' test item content. The paper explores questions such as what familiarity with the Mongolian camel can tell us about test item relevance, and what *not* being able to fill in a simple form for employment in a fast food restaurant can reveal about the cultural, contextual and ideological agendas of standardised assessments. The ethnographic analysis of assessment events provides an opportunity to examine people's behaviour and their emotional and verbal interactions – the questions people ask and the responses they receive. This can inform understanding of how cognitive, affective, moral, contextual and cultural dimensions of assessment inform statistical outcomes. It goes without saying that these kinds of data are rarely included in large-scale statistical analysis.

This paper is informed theoretically by Goodwin's 'participation framework', and by Actor Network Theory (ANT). Charles Goodwin's work views the production of context as an interactive social phenomenon. This includes the role of material objects such as assessment texts in communicative acts (see Duranti and Goodwin 1992; Goodwin 2000, 2007). Goodwin's work emphasises the situated and embodied character of human interaction and its spatial, affective and temporal dimensions. The ethnographic transcripts below provide a rich source of qualitative data about how people respond to test items and assessment tasks.

The inclusion of Actor Network Theory (ANT) provides a complementary perspective, explaining social action in terms of an 'assemblage' of human and non-human actors (Callon 1986; Latour 2005; Waltz 2006). The study of literacy assessment events illustrates what ANT describes as the unpredictable and unstable processes of 'translation' as various actors and objects come together to form a network that constitutes particular settings and events (see Callon 1986; Fenwick and Edwards 2010; Hamilton 2011). In the case of literacy assessment, qualitative insights into the dynamics of the assemblage are illuminating. They illustrate, among other things, the negotiation of power and authority, sources of motivation, the implications of cultural norms and expectations, and the resources that test respondents bring to the assessment event.

The arrival in the Mongolian Gobi of distinctly foreign and unusual assessment texts, supported and accompanied by teams of staff and associated technical manuals, presents a novel context for ethnographic enquiry. While psychometric literacy assessment strives methodologically for standardisation as good practice, the contextualised process of assemblage in assessment involves features of social interaction that are unplanned, chaotic and particular. This phenomena is described in Actor Network Theory with the concept of 'irreducibility' (Fenwick and Edwards 2010; Latour 2005), and similarly in the work of Bloome's (1995) micro-ethnographic approach as 'indeterminacy'.

We can apply the principle of irreducibility to analyse literacy assessment events where globally mobile assessment texts are received in local contexts. Rizvi and Lingard (2010) argue that while globalised educational policy and practice reveals features of ideological and methodological alignment, there is 'nothing inevitable' about outcomes of globalisation that are 'mediated at the national and local levels by particular historical, political and cultural dynamics' (3). As Rizvi and Lingard argue, we

should consider how texts are received in sites that are far from those of their original production, into what Bourdieu describes as new fields of reception (Bourdieu 1999).

The analysis of assessment events draws our attention to their spatial, temporal and transcontextual character. These aspects of assessment involve processes of *movement and re-contextualisation* (see Bernstein 1996; Blommaert, 2008, 2010; Brandt and Clinton 2002; Kell 2000, 2006, 2011; Reder and Davila 2005). Three aspects of re-contextualisation are considered: (i) in test item production, content from 'everyday' life is re-contextualised in Bernstein's (1996) sense, as carefully crafted educational artefacts to be used in highly regulated domains of educational assessment; (ii) the 'traffic of texts' (Kell 2011, 613) along globalised networks of educational assessment. The circulation of assessment texts as valued exchange commodities or cultural goods; and (iii) questions of 'semiotic mobility' (Blommaert 2010, 3), or what Kell (2011) describes as 'recontextualisations of meaning' (609) as testees and other actors attempt to make sense of texts and images during assessment events by drawing on their local knowledge and cultural resources (see Bourdieu 1999).

As Steiner-Khamsi (2012) argues, 'the focus on reception and translation lends explanatory power to local policy contexts' (2012, 3). The local ethnographic analysis invites us to question assumptions about the hegemony of national and global institutions, and to explore new sources of institutional obligation, authority and affinity. As Blommaert argues:

> The 'margin', so to speak, is not necessarily a space in which people fail to meet norms, but it can as well be seen as a space in which different but related norms are produced, responding – 'ecologically', so to speak – to local possibilities and limitations. (Blommaert, 2010, 80–81)

An anthropologist among the psychometricians

In 2010 the author was invited to ethnographically observe the national adult literacy assessment conducted in Mongolia as part of the UNESCO Literacy Assessment and Monitoring Programme (LAMP). The aim of LAMP is to support national governments to produce robust statistics on adult literacy to inform educational policy and planning (Guadalupe and Cardoso 2011; UIS 2009).

LAMP is an example of the ambitious, technically demanding and increasingly globalised projects of large-scale literacy assessment (i.e. literacy as numbers), that are promoted as good practice by statistical organisations and trans-national institutions such as the OECD, UNESCO and the World Bank. LAMP shares many methodological features and some content with other assessment projects (e.g. IALS, PIAAC, ALL) including the use of Item Response Theory (IRT), test design frameworks and some test items.

LAMP adopts a psychometrically derived notion of a continuum of literacy and numeracy abilities presented as three 'levels' of performance across 'domains' of prose, document and numeracy (i.e. three common scales). The Mongolian assessment involved a randomised sample of 4000 respondents. The first full round of assessment took place in Mongolia, Jordan, Palestine and Paraguay. These countries are linked methodologically through processes of test item development, statistical procedures such as calibration and weighting, and ultimately through common reporting and comparison. They are also linked in networks and alliances, involving flows of materials (including test items), expertise and resources. These networked actors and institutions

can be viewed not only as technical experts, but also what Ball (2010) calls 'policy entrepreneurs'. Such networked expertise is vital in processes of methodological innovation, learning and problem solving – to deal with the challenges associated with the movement, translation and comparison of assessment items across culturally and linguistically diverse contexts.

I observed the national literacy assessment in Mongolia, and conducted informal interviews with respondents and testers, and detailed ethnographic observations of assessment events. I travelled and lived with assessment teams – as an anthropologist among the psychometricians. The assessment events described in the paper took place in the Gobi desert in the province of Dornogovi. The desert is sparsely populated and home to nomadic pastoralists who herd camels, horses, goats, sheep and cows. The nomads are resilient and capable, making their living in a harsh desert environment. There are few roads outside the provincial centre, only tracks in the sand, and the landscape is marked with rolling dunes and the occasional mining complex.

The assessments took place in people's homes (whether conventional houses in urban centres or in *gers* – the traditional, round, felt covered nomadic homes). Urban and nomadic homes are often one room constructions. That imposes certain dynamics on the testing situation. The Gobi desert is bitterly cold in November, so the testing team, the respondent and their family, including babies and young children, the local guide and driver, interpreter and the author were frequently present during assessment events. The adults would usually be offered tea and milk-based snacks while the assessment took place. It proved to be a very useful context to study standardised literacy assessment.

Mongolian people normally welcomed the LAMP assessment team, and were interested in the assessment tasks. The respondents were generally confident to ask questions and to make evaluative comments about the assessment. They sometimes provided running commentaries as uninvited 'think aloud' protocols. It was instructive to observe how respondents attempted to involve other people – the tester and family members to help them to complete the assessment task – asking for guidance, checking out the ground rules of the assessment. That interactive aspect was to some extent a feature of all the assessment events. However, in some cases attempts to involve others breached the norms of 'good practice' that are intended to isolate the respondent from others who might provide assistance (i.e. it is supposed to be an individual assessment). The following example of 'unruly' behaviour is an extreme case of a wider phenomenon – the significance of social interaction in assessment events.

The ethnography of assessment events

This ethnographic account of an assessment event with a middle aged, educated man was conducted in his home. A background questionnaire was completed, and the choice of respondent took place by randomised selection. The assessment took place in an open plan room consisting of a sitting area at one end (with settee, chairs and television) and an attached kitchen.

There were four of the testing team present – the tester (T), myself, interpreter and a visitor from UNESCO. Of the respondent's (R) family, his wife (W) and teenage daughter (D) were present, and later in the assessment his son (S) arrived. The tester sat on a stool opposite the respondent at a low table. R's wife and daughter moved during the interview between the kitchen area and the sitting area. I sat on the settee observing the event with the interpreter and the visitor from UNESCO.

This 'assemblage' of actors contrasts with the individualistic, cognitive construct of literacy that informs psychometric assessment design. The assessment events did not involve isolated, disembodied individuals. They were embodied social actors with status, emotional lives and agency. For ease of analysis the transcript is presented in sections. The transcript was produced through simultaneous translation and note taking. In that sense the transcript is incomplete as it does not capture details familiar to sociolinguistic transcription – such as tone, the nuances of the Mongolian language or overlapping speech. We join the event where the respondent is presented with the first of the testing booklets:

R: Whoah! They're asking me to do this stuff!
W: Don't make the booklet dirty.
R: Well then, ask my daughter to do it.
T: She can't do it, she wasn't selected.

At this point R's wife (W) and daughter (D) come from the kitchen area to help and to lend moral support. The interviewer says that they can't help.

R: They can't help? Oh damn it!

The respondent's wife and daughter nevertheless read the document instructions with the respondent, and begin to read the questions out loud for the first assessment item.

T: The important thing is that you can't help him. Don't be stressed, it's for the Mongolian government, for the national literacy assessment, it will be confidential. Nobody will have access to the information.
R: What is this? How am I supposed to do this kind of stuff?

This opening section of the assessment event contains many of the characteristics of other assessments observed – involving the interactive co-construction of the event and negotiation of the 'rules of engagement'. In this case there is a rapid emotional response to the assessment tasks and a process of negotiation where the tester establishes the individual orientation of the assessment tasks. The respondent is clearly not keen to undertake the assessment. The tester asserts the rationale of randomised selection: 'She can't do it, she was not selected'. Then, as the respondent's wife and daughter come to offer help and emotional support, the tester again attempts to establish and enforce the rules of the assessment by use of a 'scale jump' (Blommaert 2008) that locates the event within a wider national context and, by association, obligations of citizenship: 'Don't be stressed, it's for the Mongolian government, for the national literacy assessment, it will be confidential'.

This is an ideal moment to reflect on the way that social context shapes the assessment event. The history of state power in Mongolia, particularly in the former communist period, and the uncertainties of life in the contemporary 'market' era mean that the respondents have reasons to be cautious in their interactions with the State. The long and often repressive regimes of the communist era (1921–1990) imposed powerful obligations of nationality and citizenship where non-compliance with requests from government had severe implications. Now, in what is referred to as the 'age of the market' (Sneath 2002), people have less reason to be wary of state power, but remain reliant on the good will of social networks including those of the district ('*sum*') administration for access to goods and services (Empson 2011; Sneath

2002). As Empson (2011) argues 'The Mongolian government still retains enormous decision-making powers over the district, its people and resources' (24). Furthermore, in the Mongolian Gobi, there are strong obligations toward hospitality, and any overt expressions of displeasure in encounters with visitors are thought to be inappropriate and risky (Humphrey 2012). The obligations to others (in this case to representatives of the State conducting a literacy assessment), may be viewed within the cultural notion of 'harnessing fortune' ('*hishig, hurtech*'), which Empson (2011) describes to mean 'grace, favour, benefit, fortune' (14). This helps to explain the outcome of the initial period of negotiation of the respondent's participation, why he is formally 'reassured' by the tester's comments. Returning to the assessment event, the interviewer has just explained about the procedures for answering the questions, and about the relationship between the comprehension texts and the associated questions. The respondent (R) looks through the booklet, and smiles nervously. Then he calls his wife:

R: Hey! I don't understand this question. What is this word – 'foundation'?
T: You can't help [speaking to his wife]. You [the respondent] have to look for the information in the text.
R: How am I supposed to know?
T: All the information related to the question is in the text. It is not testing your knowledge, just trying to look at your literacy.

The interaction between the respondent, his family and the tester is illuminating. Here, he raises one of the fundamentally important questions about the assessment process – that is, the relationship between textual and contextual knowledge. The tester attempts to establish the principle that the text is the primary source of information. As we will see later, the assessment process expects the respondent to privilege textual content over contextual knowledge. Secondly, the tester is establishing a position about the construct in the assessment: that 'literacy' can be reduced to an isolated (and individualised) text-based act, disentangled from wider processes of knowledge production and interpretation.

T: Nobody other than you can help to answer the question. If you feel that you can't do it, you can skip it. The information is on the left and the question on the right. You should look carefully and find the answer in the text.
R: Can I put any number?
W: Just put 40! [she calls from the kitchen area]

The respondent's daughter comes in from the kitchen area and explains the task to her father. The respondent asks:

R: Is it real? Is the foundation really in Mongolia?
T: No – it's not real. It's just for the assessment. You have to look to find the answer.

What we see in the section above is a clear process of negotiation about rules and procedures. But there is something else. The respondent is exploring an ontological question about the status of the text – i.e. 'is it real?'. In this case, the tester responds that it is 'not real' (implying that the text is a genre of realistic fiction). However, the content of the text is serious and realistic in nature. The ability to distinguish between 'real life' information and inferences drawn from real life about hypothetical content is one of the more nuanced aspects of the assessment tasks. We will return to the question of what's real later in the paper.

'Sister, you can't help him'

The 'unruly' behaviour of the respondent and his family are a challenge to the integrity of the assessment event and the procedural requirements of good practice. The respondent's family want to help and indeed, the respondent expects them to help. So rather than a 'focal event' (Goodwin and Duranti 1992, 11) involving the respondent and the assessment text, what we have is a collective event, in which the respondent and the assessment text is the focus of attention. The tester has been co-opted as a resource for the respondent – where questions can be asked about rules and boundaries, questions of clarification and strategy.

The respondent's daughter is now leaning over her father's shoulder, reading the tasks with him. The respondent's daughter asks the tester for clarification:

> D: Does he count the number or the percentage?
> R: It's talking about honey or something.
> T: Well, the information is in the text. It's easy to find the answer.
> D: I don't think it is honey, I think it is oil. [she whispers help]

The assessment continues in that style with the daughter providing help and asking the tester questions about the test items.

> D: Can he write it down here? [she asks the tester]
> T: Sister – you can't help him! Why not go and watch TV.
> R: Daughter! Come here.
> R: What is this?

As the assessment continues the respondent's daughter and wife are increasingly involved.

> D: You should calculate 10%, and half of that, then add to get 15%. You should get 10%, and then 5%, and then deduct it from the price.
> T: Hey! You should not help him.
> R: This is my first day of leave, and I have to do this difficult task!
> R: Is this correct?
> D: No!
> R: Why?
> D: Because you are using the wrong numbers.
> R: What the hell is this? Daughter – come here!
> T: You can't call her – she is not supposed to help.

The daughter does come to help – and the man's wife also joins them. It is a touching example of support for a family member. The daughter is openly going against the guidance provided by the tester.

> R: This is such a stupid question.
> T: Sister, I told you. You can't help.
> R: Why is this survey coming to *my* house?
> T: It's randomly selected.

As this conversation takes place, the respondent's wife joins them and leans over to support him, whispering the answer to a question. The man is clearly stressed by the task. He uses phrases including: 'These questions are really confusing'; 'boy, this is such a pain!'; and 'I'm being killed, and I have no help'. But he does have help.

The man's wife and daughter stay with him (either leaning on him or standing close by), and provide help and emotional support. In many of the assessment events family members took up the same physical stance, leaning over the respondent to read the assessment booklet to provide moral support. They did not always provide the answers. In this case, despite the tester's best efforts, there does not seem to be a way to exclude the family members from the process.

The man's son returns home and asks what's going on.

R: Your sister is not helping me. [to his son, who joins him and leans over to read the question]
S: The answer is right there. [he points to the text]
Then later in the assessment:
D: You see, I am helping you. You said that I was not helping!

The social in context

What can be learned from this episode? It's clear from the transcript that the tester tried but failed to prevent the family members from collaborating in the assessment event. Should that simply be attributed to 'bad practice' in testing, or the behaviour of unruly people? To answer that question the following exchange after the assessment between the respondent and tester is quite revealing:

R: That was a heavy exam!
R: Couldn't I get help? Sometimes I do.
T: No, you're not supposed to.

This provides an insight into how the respondent and tester viewed the assessment event and the collective nature of everyday literacy practices. That is (a) the respondent acknowledges that he gets help with literacy, i.e. 'sometimes I do', and (b) that the tester refers to an external source of authority or standards in the assessment 'you're not supposed to'. The first statement relates to the realm of everyday life, while the second refers to the attempt by the assessment regime to construct a competing context for the assessment.

The transcript illustrates the dissonance produced as the participants contest competing constructs of literacy. The testing situation is based on a construct of literacy abilities that are cognitive attributes of an isolated individual. Therefore, the respondent's acknowledgement that he sometimes receives help with literacy appears to show that his functional literacy abilities are lacking.

Another way to look at the event is that the respondent *has* been able to complete the assessment tasks with the help of others. That explanation seems entirely inappropriate from the perspective of an assessment regime and the educational construct of individual competence. However, it is one way to understand the behaviour of the respondent and his family. Anthropological studies of literacy describe its inherently social character. They suggest an alternative construct based on a situated model of literacy and cognition. That perspective is supported by Goodwin's 'participation framework':

A primordial site for the analysis of human language, cognition, and action consists of a situation in which multiple participants are attempting to carry out courses of action in concert with each other through talk while attending to both the larger activities in which their current actions are embedded and relevant phenomena in the world around them. (Goodwin 2002, 22)

Goodwin's framework seems apt. The literacy assessment event is fundamentally inter-
active and the participants are attentive to multiple sources of information that are
present in the event and outside the event, some of which are textually based, and
others that relate to the 'wider world around them'.

Context and affect

The ethnography of literacy assessment events reveals content about social interactions
and emotional affect that may influence the outcomes of assessment, and that are not
usually captured in statistical assessment data. That includes levels of stress, pleasure,
fatigue and motivation. Some respondents were distracted by other things going on in
their lives such as prior appointments or the arrival of visitors. In those cases they some-
times skipped questions and rushed the assessment tasks. The verbal responses of
respondents capture this kind of detail. For example, the respondent below was due
to take her baby for a health appointment later that day:

> I want to do this fast. I'm busy in the afternoon. I have to fetch water. I don't have much
> time.

Most of the respondents showed a high level of concentration and commitment to the
assessment. They put other activities on hold and asked others to prepare food, to look
after babies and children, and said that other demands such as looking after animals
would have to wait. The following excerpt from an assessment with a nomadic
herder illustrates this point:

> R: These graphics are confusing.

Outside some visitors are arriving to look at horses that the respondent intended to sell.

> Visitor: It's getting dark.
> T: number the answers to make them clear.
>   [The respondent smiles, and concentrates on the task. There is movement of
>   animals outside. A visitor comes in, pours himself some tea and lights a pipe.]
> T: If you think it is difficult you can skip the question.
>   [He does not skip the question and is really concentrating on the task. Three
>   herders are now sitting watching the assessment event. A fourth herder enters]
> Visitor: The horses are coming in.
> R: Okay. You go for the cows.

In that example the respondent is attentive to the assessment task while life goes on
around him.

These examples provide insights into how the assessment event is co-constructed by
human and non-human actors, and how the respondent deals with their competing
demands and expectations. The second part of the paper develops this further to
examine how the respondents tackle test items.

This is not a camel

The international character of large-scale assessment programmes highlights the chal-
lenge of maintaining cross-cultural validity, and associated questions of test item adap-
tation, relevance and comparability. In the second part of this paper two contrasting test
items are examined to see how they were received by respondents in the Mongolian

Gobi. The test item content was informed by a methodological framework that promoted the use of 'realistic' materials, while maintaining cross-cultural validity (Kirsch 2001).

The Mongolian camel test item was initially developed by the Mongolian LAMP team. The test item provides an image of a Bactrian camel followed by some descriptive text about camels in Mongolia. It has three questions and the answers can ideally all be obtained from the descriptive text without the need for local knowledge. The Gobi is the home of the Mongolian camel, and the people there are experts on camels, so it was instructive to see how they responded to 'realistic' test item content.

The test item proved to be very popular with the Mongolian respondents. They frequently identified the test item as one of the most enjoyable of the assessment. These questions clearly produced a positive emotional affect.

One might anticipate that local knowledge would enable people who own Mongolian camels to perform well on this test item (perhaps better than others taking the assessment outside Mongolia). However, this was not always the case. Indeed, in statistical terms, Mongolia's respondents were no more likely to obtain the correct answer than respondents from other countries. To explain this, we need to consider Bernstein's (1996) concepts of educational framing and re-contextualisation. We can safely assume that nomads in the Mongolian Gobi are experts on Mongolian camels. However, to paraphrase the artist Magritte, 'this is not a camel', rather it is a representation of a camel, positioned within the particular requirements of a performance-oriented test item. As non-experts in Mongolian camels we have limited knowledge resources to help us to answer the question. In contrast, the Mongolian herders have extensive and nuanced knowledge and experience of camels. Within the assessment there is a clearly demarcated sense of legitimate knowledge and correct answers. Implicit in the design of the test item is that the respondent should privilege the knowledge that is available in the text and, by implication, to relegate their intimate knowledge of camels to secondary importance. The question clearly points the reader in that direction:

> *Use the article* about Mongolia's Bactrian Camel on the opposite page to answer questions 3 through to 5. (Emphasis mine)

The refrain of testers that 'the answer is in the text' alerts us to the dominant construct of literacy that informs the assessment design. The manual for testers emphasises the importance of precision to ensure that the wording of the answer matches the content of the assessment text. The respondent is presented with a dilemma. They are 'invited' by the familiar topic to use local knowledge to help them to complete the task. However, what seems local is actually only 'rendered local' (Latour 1996). As an assessment item it is a *global* camel – one that has travelled from Mongolia, attended global meetings on literacy assessment, and then, having been framed within the assessment text, is re-introduced into a Mongolian setting. The analogy with Mongolia's political history is not without a sense of irony. Knowledge of camels arguably operates as a 'nuisance dimension'. This can be illustrated with a second transcript below.

The following account of assessment was conducted with a woman nomadic herder (R) in the Gobi. She had been reluctant to take the assessment and openly questioned its value. She had previously been visited by a team conducting a survey about child labour and asked if this assessment was part of that project. The tester (T) had persuaded her to take part, saying that we had travelled a long way across the desert to be there. It was cold, late in the day, and the sun was going down. Unlike most of

the respondents, the women complained about the assessment and some of the questions in the background questionnaire. When asked about her household possessions she replied:

> R: What are you talking about? Are you crazy? Do you see any stuff like that here? I do everything with my bare hands!

The conditions for the assessment were not good. There was no table, and there was no fire. The assessment was completed by the respondent while she stood under a weak lamp. The respondent complained about her eyesight, and held the assessment booklet close to her face:

> R: I can't read, I'm blind.

The respondent complained that the assessment contained 'a load of shit questions'. It appeared that she did not appreciate the value of the national literacy assessment. However, she continued (skipping the questions that she could not answer), and when she got to the question about the camel she clearly showed an increased interest and motivation.

The respondent read the questions slowly and out loud and joked with the tester.

> Does a camel drink that much? I'm kidding!

The respondent answered two of the questions correctly, but provided the answers before she had read the descriptive text by using her local knowledge. In the first two cases her answers corresponded with the information provided in the text. However, for the third question she provided a very detailed answer based on her expert knowledge of camels. She described in some detail about how camels are patient about the cold and the heat, how they don't have to drink every day and how they can cope with hunger. Her answer did not match the text and was marked as incorrect.

Expertise about camels helped the respondent to complete the task (and increased her enjoyment and motivation), but that did not help her to answer questions based on obtaining information from the assessment text. Local knowledge about camels in the Gobi seems to have undermined the rigour of the assessment item. This may be a case of 'too much context'.

The research literature on 'realistic' content in assessment reveals how it can unexpectedly become the source of test item bias. Cooper and Dunne (1998) for example, describe how girls from working class backgrounds are disadvantaged because they can't anticipate the educational codes and norms associated with realistic content in mathematics test items, and the 'educational ground rules' that enable them to produce the 'correct' answer. They argue that too much realism from the respondent will produce multiple answers, many of which might sound plausible, but will not produce the 'correct' answer: 'The child is required to introduce only a small dose of realism – "just about enough"' (120).

The question 'is this real?' asked by several of the Mongolian respondents during the LAMP assessment events is telling. The realistic content of the assessment can cause confusion about the level of realism involved. There were numerous occasions in the assessment events observed where the participants appeared to struggle over

realistic content in the text. In some cases respondents declined to answer questions because they said that they were not familiar with the realistic content.

A simple application form for employment

This second and contrasting example involves a test item requiring applicants to complete a realistic application form for employment in a fast food restaurant. The questions ask the respondent to complete part of a form and to indicate their availability for work. The test item is used in LAMP, but was originally developed for the International Adult Literacy Survey (IALS) for use in OECD countries (OECD 2000) and some modifications took place for use in LAMP.[1] This is an example of a high stakes assessment since 'not being able to fill out a simple application form' is frequently cited in the media as evidence of poor levels of adult literacy. That illustrates how a sense of deficit and crisis can be developed from literacy assessment data to legitimise neoliberal policy interventions – framing respondents as potential wage labourers in need of training.

The test item application form is realistic, but the question of 'how real' is part of the puzzle. The form appears to be real and has the design and hallmarks of authenticity associated with textually mediated institutional encounters. Furthermore, the instructions address the respondent directly as 'you'. The respondent is therefore invited to 'play' in a fictional scenario. The respondent must understand that it is not really them, but a make believe 'you'. The assessment text is not supposed to be about their actual life – i.e. it has nothing to do with their actual availability for work. The content is realistic, but definitely *not* real (as the tester noted in the earlier transcript). As Bernstein's (1996) concept of a pedagogic device suggests, the real life content is re-contextualised as an assessment task.

In the Mongolian context, the test item clearly illustrates processes of geographical movement and re-contextualisation. The test item is based on particular Western experiences of markets for employment and culinary traditions.

How did the Mongolian respondents manage the application form? In contrast with the camel question, many of the respondents commented that they found the employment form difficult, and said that they did not understand the task. However, the statistical DIF (Differential Item Functioning) analysis did not reveal any problems with the performance of the item. Difficult items are a necessary feature of literacy assessment that uses Item Response Theory (IRT) to reveal the distribution of abilities within the population. However, the sources of difficulty matter. In this case it appears that some respondents found it difficult to make sense of the test item, and in doing so they had little experience to draw from:

R: What does it mean? I'm not going to answer this question.

And another respondent said:

R: Should I fill it in with my information?

The ethnographic transcripts reveal a semiotic dissonance associated with the ambiguously realistic content of the item. They show that respondents were unclear about the nature of the assessment task and as a result they sometimes skipped the test item (which would be marked as 'incorrect').

R: How do you know what the answer should be?
T: The answer is given in the information.
R: What does this sentence mean?
T: This is about the distance from the house to the work. Not your actual work, it's the information in the text. Not real life.

The fast food test item appears to promote a neoliberal social imaginary, framing nomadic herders as potential restaurant employees.[2] The irony in this case is that for many respondents in rural areas, the 'realistic' content is unfamiliar. Fast food restaurants are not a feature of the Mongolian Gobi and the depiction of the formal labour market was alien to many respondents. Nomads in the Gobi make a living from herding animals in a harsh environment. Where opportunities exist some nomadic families do take paid employment (Humphrey and Sneath 1999; Sneath 2002). In the Gobi area those opportunities include work in the rapidly expanding mining industry. A few of the younger respondents commented that they had recently completed application forms for work in the mining sector. However, others commented on the difficulty of the test item, saying that they were not familiar with application forms and had never applied for employment. This example is a case of 'not enough context'. For this test item the ethnographic transcripts suggest that test item difficulty is associated with test item movement and re-contextualisation – how respondents received and made sense of its unfamiliar context.

Conclusion

The global scale of educational assessment programmes is said to offer powerful comparative insights into educational achievements. However, the forces that drive globalisation in assessment (i.e. the desire for international comparison) are not necessarily sensitive to the difficulties that are created when test items travel. The movement and re-contextualisation of test items across radically different cultural and linguistic contexts presents formidable challenges to the validity of international assessments.

This paper has sought to investigate those dynamics with an ethnographic account of standardised adult literacy assessment Mongolia. It introduced the 'assessment event' as a focal feature of ethnographic analysis, and its scope to inform the analysis of test item performance. The paper used the concepts of 'irreducibility' and 'assemblage' from Actor Network Theory (ANT) to examine how apparently standardised assessment events display unruly and idiosyncratic characteristics. These are the result of multiple factors (often neglected in statistical analysis) – such as the agency and behaviour of actors, emotional affect, sources of motivation and obligation, and the relationship between textual and contextual knowledge. The use of ANT in this paper has illustrated how non-human actors (including technical framework documents and test item booklets) shape assessment encounters and help to transmit global and neoliberal agendas. Nevertheless, social reproduction cannot simply be read off from the content of test items. The transcripts of assessment events demonstrate the ANT principle of 'irreducibility', as the assemblage of human and non-human actors produces its own dynamics in local–global encounters.

The opening ethnographic account of an 'unruly' literacy assessment drew on the work of Goodwin to show how assessment events are co-constructed through embodied social interaction, and its spatial, temporal and affective dimensions. The context of the

event is extended to include interactions between the tester, the respondent, the assessment text and others present – such as family members. The inclusion of small-scale ethnographic analysis provides insights into how respondents make sense of the assessment, and helps to identify unexpected sources of difficulty that were not intended in the assessment design. In the opening example, that includes difficulty associated with the 'realistic' content of the test items, and the associated anxiety with the ontological question 'is it real?'.

The theme of realistic test item content was explored further in the second part of the paper. 'Mongolia's Bactrian Camel' and the 'Employment Application' highlight some of the challenges associated with 'realistic' test item content, cross-cultural validity and adaptation. Those test items represent contrasting aspects of contemporary Mongolian society: traditional herding and the rapid changes to labour markets associated with resource extraction.

The test item on the Mongolian camel described the problems that nomadic herders faced in dealing with the dissonance between realistic textual knowledge and their own expertise in camel herding. It revealed how the assessment design privileges written information in assessment texts over local knowledge. The local content increased people's motivation and interest in the questions. However, one could argue that the test item was 'too real', since local knowledge acted as an unexpected distractor. Thus people who are less familiar with camels or the geography of Mongolia might have been more likely to obtain their answers from the textual information.

The fast food employment application form highlights a contrasting aspect of Mongolian society linked with globalisation. In this case, the realistic content was 'not real enough'. Fast food restaurants are not currently a feature of life in the Mongolian Gobi. Nor are people experienced in dealing with textually mediated access to labour markets, though this may change rapidly with the expanding economy. The test item is about how people engage with the political economy and the changing political ecology of literacy.

Notes
1. Lennon, ETS personal communication, May 2012.
2. On neoliberal social imaginary and globalisation in education see Rizvi and Lingard (2010) and Ball (2012).

Notes on contributor
Dr Bryan Maddox is a senior lecturer in education and development at the School of International Development, University of East Anglia, UK. He has conducted ethnographic research on literacy practices and assessment in Nepal, Bangladesh and Mongolia, and has written widely on the field of literacy studies.

References
Ball, S. 2010. *Global Education Inc. New Policy Networks and the Neo-liberal Imaginary*. Routledge: Abingdon and New York.
Bernstein, B. 1996. *Pedagogy, Symbolic Control and Identity: Theory, Research, Critique*. London: Taylor and Francis.
Blommaert, J. 2008. *Grassroots Literacies: Writing, Identity and Voice in Central Africa*. London and New York: Routledge.
Blommaert, J. 2010. *The Sociolinguistics of Globalisation*. Cambridge: Cambridge University Press.

Bloome, D. 1995. "Necessary Indeterminacy and the Microethnographic Study of Reading as a Social Process." *Journal of Research in Reading* 16 (2): 98–111.

Bourdieu, P. 1999. "The Social Conditions of the International Circulation of Ideas." In *Bourdieu: A Critical Reader*, edited by R. Shushterman. Oxford: Blackwell.

Brandt, D., and K. Clinton. 2002. "Limits of the Local: Expanding Perspectives on Literacy as Social Practice." *Journal of Literacy Research* 34 (3): 333–356.

Callon, M. 1986. "Some Elements of a Sociology of Translation: The Domestication of Scallops and the Fishermen of St. Brieuc Bay." In *Power, Action and Belief: A New Sociology of Knowledge*, edited by J. Law, 196–233. London: Routledge and Kegan Paul.

Cooper, B., and M. Dunne. 1998. "Anyone for Tennis? Social Class Differences in Children's Responses to National Curriculum Mathematics Testing." *Sociological Review* 46 (1): 115–148.

Duranti, A., and Goodwin, C. (Eds). 1992. *Rethinking Context: Language as an interactive phenomenon*. Cambridge: Cambridge University Press.

Empson, R. 2011. *Harnessing Fortune: Personhood, Memory and Place in Mongolia*. London: Oxford University Press and British Academy.

Fenwick, T., and R. Edwards. 2010. *Actor-Network Theory in Education*. London and New York: Routledge.

Goodwin, C. 2000. "Action and Embodiment within Situated Human Interaction." *Journal of Pragmatics* 32: 1489–1522.

Goodwin, C. 2002. "Time in Action." *Current Anthropology*. 43 (Supplement): 19–35.

Goodwin, C. 2007. "Participation, Stance and Affect in the Organization of Activities." *Discourse and Society* 18 (1): 53–73.

Goodwin, C., and A. Duranti. 1992. "Rethinking Context: An Introduction." In *Rethinking Context: Language as an Interactive Phenomenon*, edited by A. Duranti and C. Goodwin, 1–42. Cambridge: Cambridge University Press.

Goodwin, C., and M. Goodwin. 1992. "Assessments and the Construction of Context." In *Rethinking Context: Language as an Interactive Phenomenon*, edited by A. Duranti and C. Goodwin, 147–185. Cambridge: Cambridge University Press.

Guadalupe, C., and M. Cardoso. 2011. "Measuring the Continuum of Literacy Skills Among Adults: Educational Testing and the LAMP Experience." *International Review of Education* 57 (1–2): 199–217.

Hambleton, R. 2005. "Issues, Designs, and Technical Guidelines for Adapting Tests into Multiple Languages and Cultures." In *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*, edited by R. Hambleton, P. Mereda, and C. Spielbeger, 3–38. Mahwah, NJ: Erlbaum.

Hamilton, M. 2001. "Privileged Literacies: Policy, Institutional Process and the Life of the IALS." *Language in Education* 15 (2–3): 178–196.

Hamilton, M. 2011. "Unruly Practices: What a Sociology of Translations can Offer to Educational Policy Analysis." *Educational Philosophy and Theory* 43 (S1): 55–75.

Hamilton, M. 2012. *Literacy and the Politics of Representation*. London: Routledge. July 2012.

Humphrey, C. 2012. "Hospitality and Tone: Holding patterns for strangeness in rural Mongolia." *Journal of the Royal Anthropological Association* 18 (S1): 63–75.

Humphrey, C., and D. Sneath. 1999. *The end of nomadism? Society, State and the Environment in Inner Asia*. Duham, NC: Duke University Press.

Kell, C. 2000. "Teaching Letters: The Recontextualisation of Letter-Writing Practices in Adult Literacy Classes in South Africa." In *Letter Writing as a Social Practice*, edited by D. Barton and N. Hall, 209–232. Amsterdam: John Benjamins.

Kell, C. 2006. "Crossing the Margins: Literacy, Semiotics, and the Recontextualisation of Meanings." In *Travel Notes from the New Literacy Studies*, edited by K. Pahl and J. Rowsell, 147–169. Clevedon/Buffalo/Toronto: Multilingual Matters.

Kell, C. 2011. "Inequalities and Crossings: Literacy and the Spaces-in-Between." *International Journal of Educational Development* 31 (6): 606–613.

Kirsch, I. 2001. *The International Adult Literacy Survey (IALS): Understanding What Was Measured*. Princeton, NJ: Educational Testing Service.

Latour, B. 1996. "On Interobjectivity." *Mind, Culture and Activity* 3 (4): 228–245.

Latour, B. 2005. *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford: Oxford University Press.

OECD. 2000. *Literacy in the Information Age*. Paris: OECD.

Reder, S., and E. Davila. 2005. "Context and Literacy Practices." *Annual Review of Applied Linguistics* 25: 170–187.

Rizvi, F., and B. Lingard. 2010. *Globalizing Education Policy*. London: Routledge.

Sneath, D. 2002. "Pastoral Adaptation and Subsistence in Mongolia's 'Age of the Market." In *Rethinking Development in East Asia: from Illusory Miracle to Economic Crisis*, edited by P. Masina, 161–184. Curzon Press.

Steiner-Khamsi, G. 2012. "The Global/Local Nexus in Comparative Policy Studies: Analysing the Triple Bonus System in Mongolia Over Time." *Comparative Education* 48 (4): 455–471.

UNESCO Institute of Statistics. (2009). *The Next Generation of Literacy Statistics: Implementing the Literacy Assessment and Monitoring Programme*. UIS: Montreal.

Waltz, S. B. 2006. "Nonhumans Unbound: Actor-Network Theory and the Reconsideration of 'Things' in Educational Foundations." *Journal of Educational Foundations* 20 (3–4): 51–68.