Developing languages of description to research pedagogy

Paula Ensor and Ursula Hoadley

Introduction

How do we make trustworthy claims about pedagogy? How do we, in both small and large scale studies of classrooms, gather and analyse data in such a way as to make confident claims about teaching and learning? This is an issue of ongoing concern for educational researchers, and perhaps more urgently now in the current context in which interventions are proposed to bring about, and measure, school improvement. 'Looking into classrooms' has become the preoccupation of those who want to measure pedagogic variation over time, and/or establish the link between pedagogic practice and learner performance. It remains the focus of those with an ongoing theoretical interest in pedagogy and symbolic control. Whatever the interest, the ways in which we generate and analyze classroom data has implications for the kinds of claims we can make about pedagogy.

The purpose of this paper is twofold. Firstly, it highlights some of the complex issues involved in researching pedagogy and the sense we make of how teachers and learners go about the business of negotiating school knowledge in classrooms. Secondly, it demonstrates how some of the difficulties identified might be addressed through developing languages of description (Bernstein, 2000).
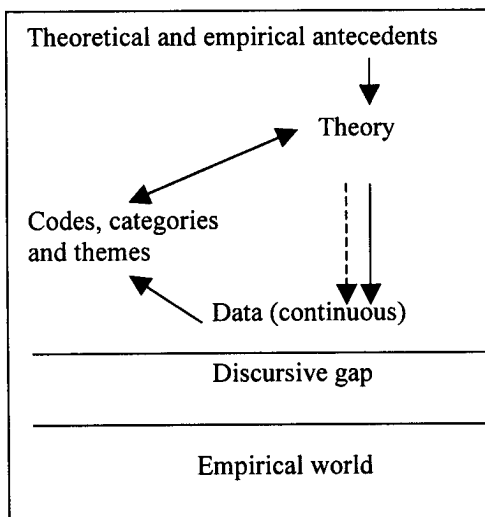
Two approaches to observing classrooms

Two broad approaches to observing classrooms have emerged in the research literature. *Inductive approaches*, often described as classroom ethnography (Delamont and Hamilton, 1993; Galton and Delamont, 1985; Hammersley, 1993), and often but not always associated with grounded theory, call for the generation of the fullest possible records of classroom life from which theoretical frameworks can be inductively derived. Inductive approaches are usually but not always associated with exploratory, small-scale studies

involved in theory construction. A notable exception is the TIMSS video study (Stigler, 1997; NCES, 1999) which adopted an inductive, theory building approach but which was relatively large in scale (and hence very costly). *Deductive approaches*, in the past often referred to as systematic observation (Croll, 1986), operate deductively from theory to the development of categories and subcategories which are used to sample aspects of classroom life. Deductive approaches are more commonly used in large-scale studies and tend to be more concerned with theory testing than theory development.
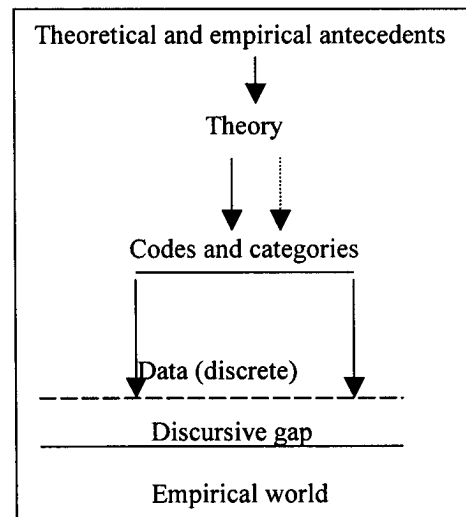
We can represent these two approaches as two ideal-types, bearing in mind that very often classroom research incorporates both of these approaches.

Figure 1:   Ideal-typical approaches to classroom observation

1a. Inductive approach   1b. Deductive approach

In the inductive approach depicted in Figure 1a, data are collected as a continuous narrative, using *open instruments* such as field notes, video recordings, a combination of field notes and audio-recording and so forth. By continuous here we are referring to the attempt of researchers to capture as complete a record as possible of classroom life over time. We are not asserting that this aim for completeness can ever be fulfilled, as invariably continuous data must constitute a selection from classroom life. Field notes, for example, cannot capture everything that is said and done, and video cameras inevitably capture some details and not others (focusing only on the teacher, for example, rather than students).

Data are unlikely to be complete, and it is unlikely that data can ever be collected independently of theoretical orientation. Theory inevitably shapes the collection of continuous data, guiding what the researcher foregrounds and backgrounds. The theoretical framework used may be well-developed in advance of the study (reflected by the solid arrow) or more tenuous (reflected by the broken arrow), but in both cases there is a relative openness in the way in which theory will be developed to read the data. Data analysis is an iterative process that brings theory and data into dialogue with each other in order to generate categories and claims.

The deductive approach depicted in Figure 1b uses theory to generate a network of categories prior to the process of data collection. Again, this theory may be strong and well-developed (depicted by the solid arrow) or less so (depicted by the broken arrow). Categories are used to develop classroom observation instruments in order to select and record aspects of classroom life. Such instruments may be interested only in one aspect of classroom life, such as teachers' questioning techniques, and therefore focus only on instances in the classroom when questioning is used. Because classroom life is sampled in this way the diagram depicts the data set as discrete rather than continuous. Sampling classroom behaviour is usually undertaken using category systems which incorporate categories, signs, checklists and rating scales (Evertson and Green, 1986). Closed systems may or may not explicitly aim to record events in time, for example recording the occurrence of particular events, of particular behaviours at pre-determined times (noting, say, every 30 seconds what the teacher is doing/saying), or particular behaviours over pre-determined intervals.

Careful attention to sampling from classroom life is a concern *a priori* in the use of closed schedules. Researchers are required to decide in advance of data collection what aspects of classroom life they will record, about and from whom, and how often. In the case of open schedules, the issue of sampling emerges in data collection in that the researcher needs to decide what to focus on. Sampling emerges as an issue also in the process of analysis, as the researcher attempts to address issues of trustworthiness in terms of how exhaustively he/she treats the collected data texts. Whether one uses codes, categories, themes or critical incidents, one is expected to demonstrate the extent and range of their presences and absences in the data in order to make robust claims about pedagogy. Analysis and findings need to be presented in such a way that the reader gains access to the method of analysis as well as a sufficiency of data to satisfy the requirements of validity and reliability. Silverman (1993) warns us against the "anecdotal" incorporation of data and

the need to provide what he terms "a sense of the flavour of the data as a whole" (op. cit., p.163). As Bryman argues:

> There is a tendency towards an anecdotal approach to the use of 'data' in relation to conclusions or explanations in qualitative research. Brief conversations, snippets from unstructured interviews, or examples of a particular activity are used to provide evidence for a particular contention. There are grounds for disquiet in that the representativeness or generality of these fragments is rarely addressed (Bryman, 1988, p.77).

'Exhausting the data text' is a challenge for researchers working with either closed or open instruments. For those using a deductive approach with closed instruments, 'exhausting the text' means providing an appropriate sampling frame to select data which are in some way representative of the slice of classroom life defined. For researchers working with open instruments, collecting continuous data in the form of field notes or video recordings, exhausting the text arises at the stage of data analysis rather than of data collection. Whether one develops coding systems, extracts themes, or focuses on critical incidents or cases, these need to be positioned against the data set as a whole in order to specify to what extent these foregrounded elements make sense of the section of classroom life selected for analysis.

It might well be argued that the approach we sketch out above is but one in a range of research possibilities, which vary according to the theoretical commitment of the researcher. We want to suggest, though, that these imperatives for data collection and analysis apply regardless of the epistemological position of the researcher, which we illustrate towards the end of the paper through a discussion of the discursive gap. Before turning to this, however, we wish to discuss some of the key issues which arise in classroom-based research.

Problems in classroom observation

Classroom observation requires selection at a number of levels:

- a research question
- the setting which we wish to observe (e.g. which classrooms, how many)
- the aspect of classroom life which is to become the focus of enquiry (teachers' questioning techniques, forms of classroom interaction)
- tools to record and store this data for study and analysis (observation schedules, video recording etc.)
- procedures for observing (where to sit/stand, when to observe)
- the subjects or events to be observed (individual, group, behaviour type, strategy)
- the analysis procedures appropriate for the question and data collected
- the method of reporting the data collected

All of these aspects are important, but in this paper we focus primarily on the last five points which we discuss in relation to the study of a sample of observation schedules which have been used both inside and outside South Africa. In South Africa, we studied 18 schedules developed for the President's Educational Initiative (PEI) Project.[1] The 18 projects from which these schedules were drawn were for the most part relatively small scale qualitative studies. In addition to these schedules, we also considered instruments which we gathered via an internet search and those that we had assembled over time from different studies. Of the total of 30 schedules that we studied, 24 were closed schedules, three were open, and three were mixed.

Two key issues emerged from an analysis of these instruments which affect the kinds of claims we are able to make about what goes on in classrooms.

1.   Very few studies appear to be driven by a theory of pedagogy (or any other related theory). It is usually difficult to establish the main features of the conceptual framework from which the indicators set out in the schedule were derived.

In very many cases, classroom observation schedules, whether open or closed, were driven by uninterrogated views of what constitutes 'good teaching practice'. Relating this to the diagrams in Figure 1 above, we suggest that the

---

[1]   We are grateful to the Joint Education Trust for giving us access to these schedules.

theory of pedagogy in both inductive and deductive approaches was weak, and research was driven by common-sense notions of teaching, or ideologically driven commitments to 'good practice'. Group work is considered a good thing; teacher exposition is not. Many of the PEI observation schedules were preoccupied with the pacing of lessons, variety in the selection of teaching resources and language used, drawing on everyday knowledge, sequencing (linking previous with new knowledge) and aspects of the moral order of classrooms (empathy, respect for dignity of students, etc). Far less emphasis was placed on conceptual development, except that a few instruments asked the researcher to note whether the teacher demonstrated sound knowledge of subject content, communicated clearly, involved students in problem solving activities and used appropriate questioning skills. Most of the PEI schedules reflect a concern with the aims of Curriculum 2005. While the kind of information which these schedules attempted to collect is not unimportant, the schedules, whether open or closed, were largely normative in that researchers entered the field with strong views about what constituted good teaching practice. Because of this, these instruments would be unable to capture information about what teachers might have been doing that fell outside of these categories.

The apparent absence of an explicit theory of pedagogy, a theory which guides the exploration of classroom life, has in our view resulted in many schedules which are unarticulated assemblies of classroom features with little or no in-depth description of any particular aspect of classroom activities. As such they do not readily suggest a research problem, and the unit of analysis – whether this be the teacher, the learners, the materials, tasks, utterances, or subject knowledge – is not always clear. In studying these assemblies of classroom features, we found that in very many cases the instruments focused on what we, following Bernstein (1990) refer to as the regulative discourse, the social relations and moral order of the classroom. Focusing on the regulative discourse foregrounds features such as teacher-learner relations and the degree of intimacy and distance entailed in these. Few instruments concentrated on instructional discourse, the knowledge and skills transmitted to learners, and only two of the PEI instruments made reference to the actual subject area under investigation. For example, in a project that aimed to explore best practices in mathematics and science, no reference was made to either subject area and the schedule called rather for observations about "Learners use of highly interactive materials" and "The creation of a conducive learning environment". By focusing on these two aspects, the instrument foregrounded regulative features, and, we fear, may have resulted in attempts to "read" the instructional through the regulative discourses. In other words, data collected

by means of a schedule asking for evidence of how learners are seated in the classroom (groups or rows), or about the availability and variation of 'interactive' learning materials and activities might be used to draw conclusions about 'learner-centred' or 'innovative' classrooms, without saying much at all about the quality of the pedagogic discourse which learners are offered.

The absence of a theory of pedagogy also means that criteria for what is to be grasped by the observer are not made available, and reliance on commonsense understandings and the judgment of the observer is increased. For example, one schedule exhorts the observer to "write down your own comments on the use of materials in this lesson. Be very honest", and another asks observers to "Please comment on values and attitudes displayed by the teacher in the lesson". We return to these issues below.

2.   *In many instances there appear to be threats to both reliability and validity.*

Closed and open classroom instruments are associated with different kinds of threats to reliability and validity. In the case of open instruments, issues of reliability (the soundness of the data collection process over time) are addressed by specifying precisely the ways in which data are to be collected. The TIMMS video study (Stigler, 1997) for example provides careful detail of decisions taken about when video recording took place, by whom, and what aspects of classroom life were focused upon. Validity, in the case of open instruments, arises largely at the stage of analysis when relationships are set up between the theory, the categories and themes developed inductively, and the data.

In the case of closed instruments, issues of validity and reliability arise most significantly at the time of instrument design and data collection. Closed schedules, which approach classroom life with a set of pre-conceptualised categories, can be either high or low inference measures (Evertson and Green, 1986), both of which have implications for validity and reliability. A low inference measure might ask a question such as "How many desks are in the classroom?". Such a question calls for little inference or judgement on the part of the observer, and reliability is potentially high. However, low inference measures are not necessarily valid, and in this sense reliability and validity tend to operate orthogonally with each other.

We can illustrate these concerns in relation to an HSRC instrument used to evaluate the implementation of C2005. We have selected this particular instrument because it is in the public domain, and because the HSRC, as a publicly funded institution, expects to have its work open to scrutiny. Having said this, this instrument has much in common with other instruments we have perused which were designed for studies funded under the PEI.

This instrument was designed in three parts: one collected brief information about the school and class observed, the second collected information about "learning programme attributes", "learner activity", "learning environment", "motivation", "learning support materials" and "assessment procedures", and the third collected information on "critical incidents".

The second part of the instrument, an extract from which is presented in Figure 2 on the next page, is an example of a high inference measure. It relies on the judgement and skills of the fieldworkers (and hence upon significant funding for training). For example, how does one gather data on whether the learning programme "develops critical thinking skills"? How do "critical thinking skills" manifest themselves in the classroom? What does one look for? What does rote learning look like? We do not deny the existence of rote learning, but because it has tended to become a term of evaluation rather than description, we cannot assume that all researchers mean the same thing by it. Is attending a lecture a manifestation of rote learning? Is the learning of the times tables in the junior school rote learning, and if so, is this necessarily a bad thing? This extract from the observation instrument is an example of what for us is an uninterrogated view of good practice, and hence an untheorised view of pedagogy. The emphasis in the schedule as a whole is upon the regulative features of classroom life – only one item is concerned with instructional discourse in that it indexes the development of higher order thinking skills.

Figure 2:   HSRC instrument used to evaluate the implementation of C2005

NATIONAL FORMATIVE EVALUATION AND MONITORING OF CURRICULUM 2005

SECTION B

		30-minutes blocks of observation						
		1	2	3	4	5	6	
	Time started:							
	Time completed:							
A. Learning Programme Attributes								
1. Learning programme is outcomes driven.								Learning programme is content driven.
2. Outcomes for the learning activity are clear.								Learning activity is taken because it is part of the syllabus or for interest.
3. Programme develops critical thinking skills.								Emphasis is on rote learning.
4. Prior knowledge of individual learners is accommodated.								Teaching aimed at the whole class.
5. Programme is learner-centred.								Programme is educator-centred.
6. Learning facilitation is evident.								Traditional teaching methods are used.
7. Learner activities are sequenced.								Learner activities are not sequenced.
8. Identification and diagnosis of learning difficulties are built into the learning programme.								There is no attempt to identify learning difficulties.
9. Learning support for individual needs is evident.								Individual needs are not accommodated.
10. Enrichment is provided according to individual needs.								There is no enrichment according to individual needs.
11. There is immediate acknowledgement of the responses.								Immediate acknowledgement of the responses does not take place.
12. Learners are actively involved in their own learning.								Learners are passively fed information.

Classroom Interaction Analysis, 2 of 10

Another example of a high inference measure used in a large-scale study is that used in the evaluation of the United Kingdom numeracy strategy (Brown, Askew, Rhodes et al, 2001). In this large-scale study, data were collected using a high inference, closed instrument which was concerned with *mathematical tasks*, (evaluated according to three criteria: "mathematical challenge", "integrity and significance" and "engage interest"); *talk* ("teacher talk", "teacher-pupil talk", "pupil talk" and "management of talk"); *tools* ("range of modes", "types of modes"); and *relationships and norms* ("community of learners", "empathy"). These criteria, taken together, constitute a particular view of good practice, and the instrument is concerned to record the extent of presences and absences in classrooms. The criteria of mathematical challenge under tasks is illustrated below.

Figure 3:   'Mathematical tasks' item in the UK Numeracy Strategy schedule

Tasks

Mathematical challenge			
All/nearly all pupils are appropriately challenged mathematically, e.g. • most of pupils, most of the time appear to be doing mathematics which challenges them to think mathematically • pupils have some control over level of difficulty	About half the pupils are appropriately challenged all of lesson/all pupils appropriately challenged for a part of the lesson, e.g. • good differentiation in main part of lesson, plenary/intro. not adequately differentiated	Some pupils are doing appropriately challenging work for some of the time.	Some pupils are doing appropriately challenging work for some of the time.

On the face of it this seems like a reasonable request for data. Being able to comment on mathematical challenge, integrity and significance is something mathematics educators want to be able to do. But there are a number of problems with this item, which affect the reliability and validity of the research results. Firstly, what does one mean by "mathematical challenge"? This is not a trivial matter, and requires judgement by the researchers. What forms of behaviour, modes of communication and utterances does one look for

as indicators of mathematical challenge? In our review of classroom observation instruments we found this to be a common feature – it was not apparent to us how the categories would be used to collect data. This is not in itself an insurmountable problem, but in large-scale research we need the assurance that field workers have been adequately trained (as described, for example, by Galton, Simon and Croll, 1980).

Both of the instruments above are high inference, and both pay attention to the issue of the recording of observations over time. However, of the sample of 30 instruments we analysed from the PEI study and elsewhere, only nine specified timing of observations, that is, specified who and what should be observed, and how observations should be spaced over the duration of a lesson. Is the appearance of mathematical challenge, for example, something one expects to see across a lesson, or only isolated instances? In this regard the numeracy instrument guides the data collector in terms of frequency. Many instruments do not do this, however, which undermines their reliability and validity. The following is an example from a PEI study interested in "best practice" amongst underqualified mathematics and science teachers.

Figure 4: Item extracted from observation schedule used in a PEI study

2. Teacher makes the meaning clear

1. Teacher uses a variety of examples, simplification strategies aimed at enhancing learners' grasp of meaning and understanding
2. Teacher uses some strategies to make the meaning clear
3. Teacher uses few strategies to make the meaning clear
4. Teacher focuses on content with slight reference to meaning and understanding
5. Teacher teaches in a manner that does not relate to meaning and understanding

Comment…………………………………………………………………….........................

This extract illustrates our concern about the lack of specification in high inference schedules: how does one recognise "meaning and understanding" in the classroom, and how does one recognize these notions in relation to specific subject areas such as mathematics or English? Furthermore, the extract illustrates the difficulties of under stipulating how the data should be sampled. What is the difference between "some strategies" and a "few strategies" and over what period these data should be collected – at five minute intervals, at the end of the lesson? Is the fieldworker to make a general assessment of the lesson at its end, or in relation to the different activities that make up the lesson?

The foregoing discussion has raised some of the difficulties we have identified in studying classroom observation schedules used both in South Africa and elsewhere. In the next part of the paper we discuss how we have attempted to address some of these issues through our own research and in particular how we bring data collection and analysis together through Bernstein's notion of languages of description.

Generating and analyzing data using languages of description

A language of description denotes the vocabulary and the syntax, the concepts and the ways in which these are woven together, which enable empirical data to be both produced and read. Bernstein describes languages of description as follows:

> Briefly, a language of description is a translation device whereby one language is transformed into another. We can distinguish between internal and external languages of description. The internal language of description refers to the syntax whereby a conceptual language is created. The external language of description refers to the syntax whereby the internal language can describe something other than itself (Bernstein, 1996, pp.135-6).

We can illustrate the relationship between the two languages by considering in the first instance the relationship between a theoretical framework (internal language of description) and classroom data (language of enactment). For the theoretical language to be able to both produce and read data, it requires a layering of categories and subcategories which allow the theory to speak about the empirical world: what is to count as data and how these data are to be read. As Dowling (1998) suggests, an external language of description develops on the basis of deductive and inductive analysis, moving interactively between the internal language and engagement with empirical data. The language of description thus developed provides the basis for establishing what are to count as data and provides for their principled reading.

Bernstein stresses the importance for external and internal languages to be loosely articulated so as to allow the external language, developed in conversation with the data, to challenge the internal language and promote its change and development. Furthermore, this loose articulation allows the researched to insert their own voice, and challenge the claims produced by the research. In this sense the idea of a language of description has both a theoretical and ethical imperative.

In our own work we use Bernstein's sociology as a theoretical framework to research pedagogy. Using any such framework inevitably introduces a systematic 'bias' into the research in that it acts selectively upon classroom life in order to answer certain specific types of questions. In our case we are interested in the processes of apprenticeship – how, in the course of pedagogic interaction, students come to master knowledge, be it knowledge of mathematics, or of the moral order of the school and how to comport themselves as learners. We are interested in two dimensions of variation: classification, which is about specialisation of discourses, spaces and agents (the 'what' and 'who' of pedagogy); and framing, which is about the relative control teachers and learners have over selection, sequencing, pacing, evaluation and hierarchical rules (the 'how' of pedagogy). These are high-level concepts and to be able to set them to work in generating and analysing texts from classrooms, they need to become more fine-grained and brought closer to the data.

This approach can be represented in the following diagram.

Figure 5 :   Languages of description
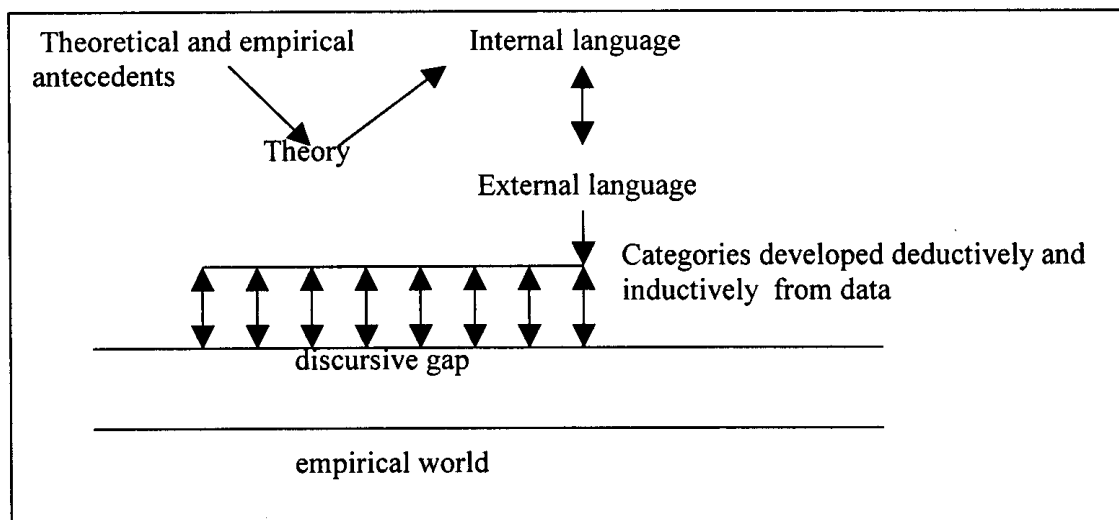
An illustration of how this process is achieved can be provided using aspects from Hoadley's research on teachers' identities and pedagogic practices in diverse social class school contexts (upper middle class and lower working class). Hoadley is conducting her research in Grade 3 classes, and is interested in the teaching and learning of mathematics and literacy. She has developed an external language using the work of Morais and Pires (2002) and Morais and Neves (2001), and more generally the work of the Sociological Studies of the

Classroom project at the University of Lisbon. A coding instrument has been designed to orient the collection of continuous classroom data using video recordings, as well as to analyse the data this generates. The instrument has been developed *a priori* and will be brought into dialogue with data in order to refine and develop it. While the instrument presented here is used as a tool to guide data collection and analysis, it is possible to use this instrument strictly deductively, as a closed instrument, as in the case of Figure 1b, with the associated threats to validity and reliability discussed earlier.

Following Bernstein (2000) the instrument seeks to assign values in terms of framing to the discursive rules of pedagogic practice: the selection, sequencing, pacing and evaluative criteria of educational knowledge. It also examines the hierarchical rules (the extent to which teacher and learner have control over the order, character and manner of the conduct of learners). The instrument also considers discourse relations in terms of the strength of classification (or boundedness) between different subject areas (inter-discursive), between school knowledge and everyday knowledge (inter-discursive), and within the subject area (intra-discursive). The instrument also looks at the classification of spaces and agents. In considering the content knowledge that is transmitted the instrument assigns 'high' and 'low' values to the level of conceptual demand and instructional density (the number of ways in which a concept is represented in the instructional practice of the teacher in order that the learner may grasp a concept). The schedule contains a set of forty indicators for the following conceptual categories:

Figure 6:   Conceptual categories for researching pedagogy

Framing	Discursive rules	Extent to which teacher controls **selection** of content
		Extent to which teacher controls **sequencing** of content
		Extent to which teacher controls **pacing** of content
		Extent to which teacher makes explicit the **rules for evaluation** of learners' performances

	Hierarchical rules	Extent to which teacher makes formal or informal the social relations **between teacher and learners**
		Extent to which the teacher controls interactions **between learners**

Instructional density (the range of ways in which a mathematical concept is represented in the instructional practice of the teacher in order that the learner may grasp the concept)

	Relations between discourses	**Inter-discursive** (strength of boundary between mathematics and other subject areas)
		Inter-discursive (strength of boundary between school mathematics and everyday knowledge)
		Intra-discursive (strength of boundary between different topics within mathematics)
	Relations between spaces	**Teacher – learner** (strength of demarcation between spaces used by teachers and learners)
		Space for learning (strength of between space used for learning
Classification Relations between agents	Teacher – learner (strength of demarcation of pedagogic identities)

Conceptual demand (the level of conceptual demand of the mathematics introduced in the classroom)

One of the forty indicators is presented below to illustrate how the instrument has been designed.

Figure 7:   Indicator 20 (Discursive relations)

Inter-discursive relations (Between school mathematics and everyday knowledges)						
20. In the contents that are used in mathematics teaching	C⁺⁺⁺	C⁺⁺	C⁺	C⁻	C⁻⁻	C⁻⁻⁻
	Extremely high level of abstraction	Predominantly high level of abstraction	Some high level of abstraction	Mostly low level of abstraction	Predominantly low level of abstraction	Extremely low level of abstraction
	90 – 100 % of the content introduced is at a high level of abstraction. Specialised terms and language predominate. All content is different from the everyday experience of learners.	70 – 90 % of the content is abstract and specialised and is different from the local, personal knowledge of the learners. Specialised vocabulary is emphasised.	50 – 70 % of the content is abstract and specialised and more local, personal content is introduced. Some specialised vocabulary is introduced.	50 – 70 % of the content focuses on concrete, local knowledge familiar to the learners, such as me, my body, cooking, shopping, with little introduction of specialised terms and operations.	70 – 90 % of the content focuses on concrete, local knowledge familiar to the learners with very little introduction of specialised terms and language.	90% or more of the content familiar to the learners in their everyday lives is introduced. There is very little or no introduction of specialised terms and operations.

A scheme of this kind has a number of advantages. Firstly, it starts from a clearly stated theory of pedagogy, which is used to develop coding categories. Secondly, and following on from this, it is transparent and relatively open to interrogation. Teachers and fellow researchers can access the criteria by which we analyse classrooms, and can challenge our findings on the basis of these. Thirdly, it provides a language whereby we can look at classroom life in a non-evaluative way. We expect variation in classification and framing relations (what knowledge is transmitted, and how) but we do not set out with a pre-conceived of what constitutes good practice and then go out into classrooms to find it. This allows the theory to go beyond the data collected, and detect both presences and absences. Fourthly, and linked to this, we can use this language to define predominant forms of pedagogy. Rather than allowing terms such as 'learner-centredness' to circulate in a fuzzy and undefined way, we are able to provide a definition using a particular combination of framing relations, usually involving weak control by teachers over micro-sequencing, selection, pacing and hierarchical rules, and sometimes strong framing over the evaluative criteria. This helps us to get away from rather crude equations such as that set up between learner-centredness and group work. Finally, because the schedule is used as an analytic rather than a data collection instrument, the scheme can undergo refinement and change in dialogue with the data.

The charge has been made against those working with Bernstein's work in South Africa that the use of a strong *a priori* theory such as his removes the possibility for the theory to undergo change. It would appear, from the criticisms made, that we enter the field with categories shaped rather like containers, into which we scoop our data! In the final part of this paper, we discuss the notion of a discursive gap, to show the potential for avoiding circularity in research.

The discursive gap

There are two ways (at least) that the notion of a discursive gap has been used by Bernsteinian researchers. Although they appear to be saying different things, they both, in different ways, point to the loose articulation of three moments of the theory-research process – the internal language, the external language and the language of enactment. Moore and Muller (2002), for example, describe the discursive gap as lying "between the internal language of the theory and the language that describes things outside it" (p.634) suggesting thereby a gap between internal and external languages. Dowling, in contrast, uses the discursive gap to point to a gap between the external language and the empirical world. In spite of differences in interpretation, all three authors set out to illustrate Bernstein's point that theoretical frameworks such as that developed by himself are capable of going beyond the data collected, and hold the potential for data to bring about changes in theory, thereby avoiding circularity and ossification.

The notion of a discursive gap was first raised by Bernstein in a mimeo, *Codes and research* which was subsequently reprinted as a chapter in his final book. In this chapter, Bernstein provides his account of the relationship between theory and research. Theory, he suggests, produces models which provide the means to decide what is to count as data, and how these are to be analysed. Bernstein notes:

> When the model is referred to something other than itself, then it should be able to provide the principles which will identify that something as falling within the specification of the model and identifying explicitly what does not so fall. Such principles we can call *recognition rules* for identifying an external relevant something. However, this something will always generate, or have the capacity to generate, greater ranges of information than the model calls for. The *realisation rules* of the model regulate the descriptions of the something. They transform the information the something does, or can put out, into data *relevant* to the model. However, if the realisation rules produce descriptions which are limited to transforming only that information into data *which at that time* appears consonant

with the model, then the model can never change and the whole process is circular. Nothing therefore exists outside of the model (Bernstein, 2000, pp.125-126, emphasis in original).[2]

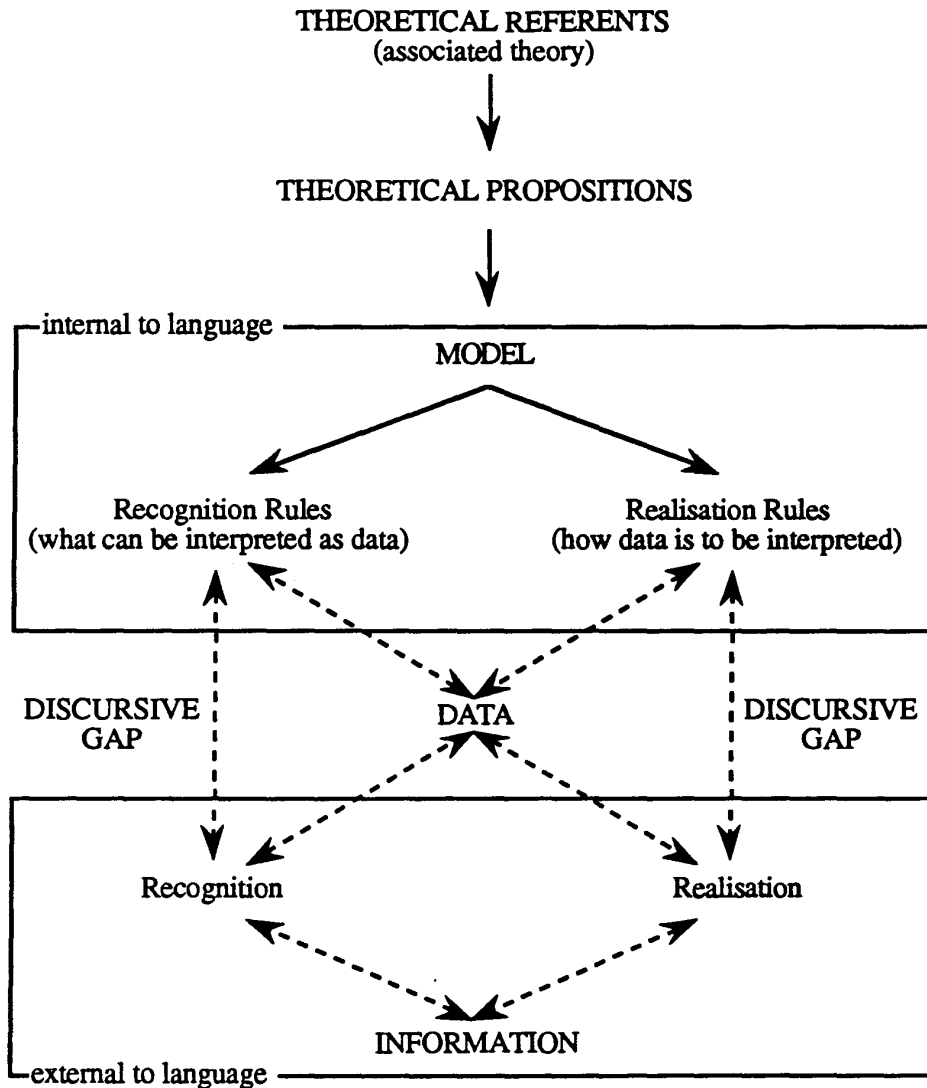Bernstein goes on, with our comments in parenthesis:

> Thus the interface between the realisation rules of the model and the information the something [transmission] does, or can produce, is vital. There then must be a discursive gap between the rules specified by the model *and* the realisation rules for transforming the information produced by the something [transmission]. This gap enables the integrity of the something [transmission] to exist in its own right, it enables the something, so to speak, to announce itself, it enables the something to re-describe the descriptions of the model's own realisation rules and so change. Thus the principles of descriptions of the something [transmission] external to the model must go beyond the realisation rules *internal* to the model (Bernstein, 2000, p.126).

Paul Dowling provides an interpretation of Bernstein's notion of the discursive gap in a PhD thesis which he produced under Bernstein's supervision. This diagram does not reflect the distinction Bernstein makes between internal and external languages of description as Dowling's PhD was produced before this distinction was fully developed.

---

[2]   Bernstein's comments are somewhat elliptical, and we have attempted to illustrate what he means by allowing his comments to speak to our present interest, namely the observation of transmission practices in classrooms. We can paraphrase the above quotation thus:

When Bernstein's model is used to analyse pedagogic transmission in classrooms, then the model should be able to provide the principles which will identify what aspects of transmission fall within the specification of the model as well as identifying explicitly what aspects does not so fall. Bernstein describes transmission as varying along two dimensions, classification and framing. Aspects of classroom life that cannot be captured using these descriptors therefore fall outside of the model. Such principles, classification and framing relations as illustrated in Hoadley's schedule shown above, can be called *recognition rules* for identifying transmission practices. However, transmission practices on the part of teachers will always generate, or have the capacity to generate, greater ranges of information than the model calls for. The *realisation rules* of the model regulate the descriptions of transmission. These realisation rules transform the information that transmission practices produce, into data *relevant* to the model. In other words, realisation rules indicate how the data collected by an instrument such as Hoadley's shown below, is to be analysed. However, if the realisation rules produce descriptions which are limited to transforming only that information into data *which at that time* appears consonant with the model, then the model can never change and the whole process is circular. This means that as Hoadley uses her instrument to orient the process of data collection, and as a basis of analysis, it will undergo transformation. The model, because it is theoretically generated, has the potential to go beyond the immediate data to describe other modalities of transmission, which may be present or absent.

Figure 8:   Structure and Application of a Language of Description (Dowling, 1995)

THEORETICAL REFERENTS
(associated theory)

↓

THEORETICAL PROPOSITIONS

↓

internal to language

MODEL

Recognition Rules
(what can be interpreted as data)   Realisation Rules
(how data is to be interpreted)

DISCURSIVE
GAP   DATA   DISCURSIVE
GAP

Recognition   Realisation

INFORMATION

external to language

The solid lines in this diagram show lines of deductive argument. In Dowling's interpretation of Bernstein's commentary, he notes

> the 'discursive gap' is between that which is internal to the language of description and that which is external to it. Data is shown within this gap. Data can be understood as the product of the recognition and realisation rules of the language, but there will always be an excess in terms of possible interpretation. The 'discursive gap' is the region of the 'yet-to-be-described' (1995, p.88).

Dowling notes further that with the methodological inclusion of the notion of the discursive gap, "it is not necessary to introduce a formal condition for reflexivity" (1995, p.88) such as that used by Bourdieu.

Our use of the discursive gap has much in common with Dowling and we invoke it here in order to signal two crucial aspects of our work. Firstly, 'the gap' signals an acknowledgement that the empirical world can only be grasped via theory, and that the empirical world is, as Bernstein emphasised, 'always ideologised'. In both diagrams of Figure 1 above a gap is indicated between the theoretical framework and the categories this gives rise to, and that which is termed the empirical world (in this case classroom life). This gap suggests that we can only get at classroom life through a theory about it, and different theories potentially generate different descriptions. These theories are cultural arbitraries in the sense that they are historically and contextually contingent, and our knowledge of the world stands removed from its "objective" materiality.

Stating the problem in this way commits us, like Moore and Muller, to a "sociological realist" position (Moore and Muller, 2002, p.635), one which admits that the world is unknown but potentially knowable, and that the material, the social and the cultural worlds are dialectically interlinked. As Moore and Muller comment: "against constructivism it [sociological realism] acknowledges the ontological discipline of the discursive gap – reality 'announces' itself to us as well as being constructed by us" (p.636).

Positivists and postmoderns might suggest that they can dispense with the notion of a discursive gap; the first because the world is deemed to be unproblematically graspable through data collection and analysis, and the second because the world is deemed to be as we construct it through the production and analysis of texts, where any artifact at all can signify as a text. All researchers ultimately are called upon to resolve the question of the status of their data texts, as to whether they reflect unproblematically and transparently upon a 'real' world (as positivists would argue), whether they should be regarded as 'events' produced by social or psychic structures, as (post)structuralists might argue, or whether they should be regarded, as postmoderns would have it, as nothing but text, so that our interest is not with what the text means but how it means, and how it constitutes rather than reflects a social or psychic world. Our argument here is that while these commitments are different, and important, they do not alter the requirement for the rigorous collection and analysis of texts, the issues with which this paper is concerned.

One reason we invoke the notion of a discursive gap, then, is to signal a particular epistemological commitment but at the same time to suggest that the requirements of making strong claims about pedagogy stand somewhat independently of such a commitment. An additional reason we invoke 'the gap' is to recognize the hiatus that inevitably occurs when developing theoretical constructs are brought into conversation with data and the potential for theory development.

Conclusion

In this paper we set out to address the question: how do we make trustworthy claims about pedagogy? We set about addressing this issue in the first instance by scrutinising classroom observation schedules used both in South Africa and abroad, in both small and large-scale studies. From this study we highlighted two key issues which for us potentially threaten our ability to make robust claims about what goes on in classrooms. Firstly, many of the studies which have been undertaken do not emerge from strong theories about pedagogy, either in general or in relation to specific subject areas. Secondly, and related to the first point, many of the schedules we studied exhibit difficulties in relation to both validity and reliability. To highlight these difficulties, and at the same time to provide a productive way forward, we have suggested an alternative approach. This approach draws on a strong theory of pedagogy and attempts to address some of the threats to validity and reliability through the development of an external language of description from the internal language constituted by Bernstein's sociology. An analytic device designed by Hoadley to analyse pedagogy was introduced to show how a valid instrument might be developed from a robust theory of pedagogy.

The charge has often been made of those using strong theoretical frameworks such as that of Bernstein, that research becomes non-reflexive, circular and incapable of change and development. We have invoked Bernstein's notion of the discursive gap to index the hiatus that inevitably arises as theoretically driven descriptors are brought into dialogue with data, and the redescription and development that should arise from this. The discursive gap signals the potential for the theory to incorporate reflexiveness.

We acknowledge that the notion of a discursive gap positions us as sociological realists, but want to suggest further, that irrespective of epistemological commitment, the challenges we face in making robust claims

about pedagogy remain shared. At issue are the steps we take to produce and analyse classroom data in order to make trustworthy claims about pedagogy. Trustworthiness ultimately is a matter of rigour, and the establishment of clear criteria of worth, rather than taking up epistemological positions and asserting that particular data collecting strategies or modes of analysis necessarily fall into line behind them.

References

Bernstein, B. 2000. *Pedagogy, symbolic control and identity: theory, research and critique.* Revised edition. Oxford: Rowman and Littlefield.

Brown, M., Askew, M., Rhodes, V., Denvir, H., Ranson, E. and Wiliam, D. 2001. *Magic bullets or chimeras? Searching for factors characterising effective teachers and effective teaching in primary numeracy.* (mimeo).

Bryman, A. 1988. *Quantity and quality in social research.* London: Routledge.

Croll, P. 1986. *Systematic classroom observation*. London: Falmer.

Delamont, S. and Hamilton, D. 1993. Revisiting classroom research: a continuing cautionary take. In Hammersley, M. (Ed.). *Controversies in classroom research.* Second Edition. Milton Keynes: Open University Press, pp.25-43.

Dowling, P. 1995. A language for the sociological description of pedagogic texts with particular reference to the secondary School Mathematics Scheme SMP 11-16. *Collected original Resources in Education* 19, 2.

Evertson, C.M. and Green, J.L. 1986. Observations as inquiry and method. In Wittrock, M. C. (Ed.). *Handbook of research on teaching.* Third edition. New York: Macmillan, pp.162-213.

Galton, M. Simon, B. and Croll, P. 1980. *Inside the primary school*. London: Routledge and Kegan Paul.

Galton, M. and Delamont, S. 1985. Speaking with forked tongue? Two styles of observation in the ORACLE project. In Burgess, R.G. (Ed.). Field methods in the study of education. London: Falmer, pp.163-189.

Hammersley, M. 1993. Revisiting Hamilton and Delamont: a cautionary note on the relationship between 'systematic observation' and ethnography. In Hammersley, M. (Ed.). *Controversies in classroom research*. Second Edition. Milton Keynes: Open University Press, pp.44-48.

Morais, A. and Neves, I. 2001. Pedagogical social contexts: studies for a sociology of learning. In Morais, A., Neves, I., Davies, B. and Daniels, H. (Eds). *Towards a sociology of pedagogy: the contribution of Basil Bernstein to research*. New York: Peter Lang.

Morais, A. and Pires, D. 2002. The w*hat* and the h*ow* of teaching and learning: Going deeper into sociological analysis and intervention. Presentation to the Second International Basil Bernstein Symposium: Knowledges, Pedagogy and Society. University of Cape Town: July.

Moore, R. and Muller, J. 2002. The growth of knowledge and the discursive gap. *British Journal of Sociology of Education,* 23(4), pp.627-637.

NCES, 1999. The TIMSS videotape classroom study: methods and finding from an exploratory research project on eighth grade mathematics instruction in German, Japan and the United States. http://nces.ed.gov/timss.

Silverman, D. 1993. *Interpreting qualitative data: methods for analysing talk, text and interaction.* London: Sage.

Stigler, J.W. 1997. Understanding and improving classroom mathematics instruction: an overview of the TIMMS video study. *Phi Delta Kappen.* http://www.kiva.net/~pdkintl/kappan/kstg9709.htm

Acknowledgements

Paula Ensor
Ursula Hoadley
University of Cape Town

uh@humanities.uct.ac.za
pe@education.uct.ac.za