METHODOLOGIES FOR THE STUDY OF INSTRUCTION IN MATHEMATICS CLASSROOMS

Patricio Herbst* & Daniel Chazan**

RÉSUMÉ

Dans ce texte, nous rendons compte d'une problématique méthodologique concernant l'étude du fonctionnement des classes de mathématiques. Nous considérons l'unité d'analyse des recherches et nous discutons la manière dont elle peut varier en fonction des objets de recherche. Nous posons aussi le problème de l'instrumentation des recherches et de la production d'enregistrements secondaires des interactions en classe. Nous soulevons le problème de la création de représentations des interactions en classe pour leur analyse par le chercheur, que nous considérons comme l'un des problèmes nécessitant un développement théorique.

Mots-clés : méthodologie, recherche sur la pratique de classe, enregistrement de pratiques, unité d'analyse, représentation de données.

RESUMEN

Damos cuenta de una problemática metodológica en la investigación en clases de matemáticas. Incluimos en nuestra discusión atención a la unidad de análisis y como ella varía dependiendo de los propósitos de la investigación. Presentamos también problemas concernientes a la instrumentación de la investigación en clases y a la producción de registros de la interacción en clases. Proponemos el problema de la creación de representaciones de la interacción en clase (para ser analizadas por investigadores) como un problema que requiere un desarrollo teórico.

Palabras-clave: metodología, investigación en salones de clase, registros de la práctica, unidad de análisis, representaciones de datos.

---

* School of Education, University of Michigan, USA, pgherbst@umich.edu

** College of Education, University of Maryland, USA, dchazan@umd.edu

ABSTRACT

We account for a methodological problematique in doing research in mathematics classrooms. We discuss how the definition of unit of analysis varies depending on the purposes of the research. We also pose some problems of instrumentation and production of derivative records of classroom interaction. We propose the problem of creating representations of classroom interaction (for its analysis by researchers) as one that calls for theory development.

Key-words: methodology, classroom research, records of practice, unit of analysis, representing classroom data

During the past 30 years, researchers in mathematics education from English-, French-, and Castilian-speaking regions have been giving increased attention to classroom instruction. Studies have focused on the learning activities in which students engage in classrooms, the way mathematics is represented for study in classrooms, and the work teachers do to facilitate the learning of mathematics by groups of students in schools. Communication among researchers across those language differences has however been scarce. Different theoretical traditions as well as cultural differences in how to write and edit scholarship have often contributed to exacerbate the obvious differences in language competence and thus discouraged mutual acknowledgment.

Globalization has brought with it an impetus for rapprochement. At a substantive level, international studies of mathematics instruction have underscored the important role that comparisons of instruction play in helping researchers of any one country understand their own instructional system. International migration (of people, practices, and tools) have made knowledge of research methodologies and results ever more important. At a formal level, the development of networks of resources for scholars as well as the increased expectation that social science research build a world class scholarly community as connected as the scientific communities in natural sciences are encouraging researchers from both traditions to learn more about what each other is studying and about how they are studying it.

An earlier attempt to deepen the communication of research across the language divide was made in a special issue of *Educational Studies in Mathematics*, volume 59, which two members of the community of French didacticians of mathematics were invited to edit. In that special issue, Colette Laborde and Marie Jeanne Perrin included nine articles written in English by researchers working within the paradigm of French *didactique of mathematics* in four different countries. In one way or another all of those articles addressed issues of classroom instruction in mathematics mediated by theoretical approaches that have emerged in the context of French *Didactique*. Some articles studied situations designed for the learning of specific mathematical ideas in a classroom, others investigated the relationship between organization of mathematical knowledge at the curricular and classroom levels, and others compared the work of teaching in several classrooms.

The journal *Recherches en Didactique des Mathématiques* has been the main outlet for publication by French *didacticiens* since 1980. While the journal also welcomes, and has published, articles written in Castilian or English, it has not been an outlet of choice for researchers who write in English. For this special issue, the journal has been interested in publishing a counterpart to volume 59 of *Educational Studies in*

Mathematics, concentrated on methodological aspects for the study of mathematics classrooms. Through us as guest editors, the journal reached the international community with the invitation to submit articles in English that addressed a part of the question, What methodologies are available to assist the study of what happens in mathematics classrooms as it concerns the mathematical content, students' learning, and the work of the teacher?

All the articles published in volume 59 of Educational Studies reported on empirical studies, yet they differed in the grain size of the analysis done, in the various kinds of analyses they coordinated, and even in the time scale with which they looked at mathematics classrooms. It seems as though each of those articles coming from a common theoretical root in French *didactique* might find fellow travelers of methodological practices in other theoretical traditions that also inspect mathematics classrooms from the perspective of the knowledge at stake, students' learning or the work of the teacher. This special issue of RDM thus intended to create a space for communication across research cultures around the methods that researchers have used to inspect classroom mathematical work. It was expected that the articles in this special issue could address portions of large set of questions that researchers on classroom instruction in all cultures need to contend with. For example,

- What study designs, data reduction, and data aggregation practices are available for the various timescales, research questions, and theoretical registers in which mathematics classroom research is done?
- What are the affordances and limitations of specific analytical techniques that might be used to inspect classroom discourse and interaction?
- Can mathematics education research studies that follow developmental, experimental, ethnographic, or comparative designs be described in terms of characteristics that are specific to mathematical knowledge, the institutional characteristics of classroom mathematics, or the obligations and constraints in which teachers work?
- What tools are available to establish relationships between what different grains of analysis or what different timescales in which classroom mathematical work unfolds?

Those goals are partially achieved with the two articles included in the issue. The article by Jill Adler outlines methodological aspects of a research program that seeks to outline the mathematical knowledge that teachers use in doing the work of teaching. While the substantive focus of that work is shared with other research programs (e.g., Ball & Bass, 2003; Ball, Thames, & Phelps, 2007), Adler's paper contributes a particular way of framing that problem utilizing Bernstein's theory of pedagogical codes. Of special methodological note in Adler's contribution is the way Bernstein's notion of evaluation is operationalized to segment instruction and within those segments to map the possible representations of mathematical knowledge. The paper by Boero, Guala, Consogno and Gazzolo shows a very different way of conducting research in classrooms, not only in terms of the definition of units of analysis but also in terms of the goals of the research, which appear to be similar in kind to those of French didacticians and American researchers involved in design research. In their contribution, Boero et al. describe how didactical innovation feeds from research in classrooms that targets at students' development of specific competencies.

In this introduction to the special issue we complement those articles with a brief, explicit discussion of methodological considerations of classroom researchers. We also include many reference pointers for readers to deepen their understanding of those methodological issues in the context of specific studies. While nobody will dispute that methodological considerations need to be subservient to the object of study, it is also true that the research accumulated over the years enables now more general discussion of methodologies adapted to the various objects of study of interest to mathematics classroom researchers.

I. THE UNIT AND SUBUNITS OF ANALYSIS IN RESEARCH IN MATHEMATICS CLASSROOMS

An important conceptual element for the establishment of a data set of classroom interactions is the identification of units of analysis. Units of analysis are the units on which observations are aggregated and claims made. Units of analysis or sub-units are also used to segment the data. The papers by Adler and by Boero and colleagues in this issue exemplify different choices of a unit of analysis.

1. Classroom research that characterizes didactical contracts and their development

Some classroom researchers in mathematics have been involved in inquiries that could be described as the design and study of didactical contracts, though in English other words would be used. Often tied to attempts to improve mathematics instruction along the lines of desiderata of visionaries (e.g., Bruner, 1995/1960) or policy documents from professional organizations (e.g., NCTM, 1989, 2000), researchers have attempted to design ways of conducting mathematics instruction alternative to what was deemed "traditional" and engage in studies of classrooms organized according to the alternative. Thus scholars like Ball (1993), Chazan (2000), Chazan & Ball (1999), Chazan & Schnepp (2002), Heaton (2000), Lampert (1990, 2001), and Romagnano (1994) articulated research projects that would inspect the work of a teacher as curriculum constructor and manager of instruction: A common object of study for those researchers has been what it takes to create a mathematics class that reflects some key chosen principles, for example to represent the discipline of mathematics with integrity, to treat students' thinking seriously, and to build a community of learners (Ball, 1993). A key research design consideration for these scholars consisted in deciding to use their own teaching as the setting where to inspect those questions – a research design that has been called "first person studies" (Ball, 2000). A unit of analysis used by researchers who used their own teaching as an instrument to understand practice has been the year of instruction – the assumption being that the object of study calls for aggregating across specific content differences. Thus these researchers have sampled records through a year/course or used comprehensive records of yearlong instructors of mathematics to discover features of the instruction observed, of the pedagogy used to promote particular kinds of mathematical dispositions, or to document dilemmas or problems of teaching they had to confront in doing that work. Analysis of this unit of instruction might also subdivide the timeline of the year to make developmental claims about the epistemological or management features of the contract established.

Some researchers have studied classroom contracts without necessarily engaging in first person studies or in collecting year long comprehensive records. Case studies such as Gregg (1995) exemplify forays into classrooms where researchers have been interested in characterizing the teaching observed and for that purpose relied on sampling lessons from a course of studies, but where the specific characteristics of the mathematics being studied was not relevant to the particular claims being made. Likewise, Boaler's (1998) studies of instruction in two different schools, where she collected among other things field notes of instruction across many lessons within a year, illustrate research that seeks to characterize kinds of instruction. Boaler's (1998) also illustrate that not all year-long studies are first person studies. Other methods of sampling and recording have been used to study course-long instruction: For example Rowan, Harrison, and Hayes (2004) used teacher logs (diaries) completed regularly over the year to describe the quality of mathematics instruction in more than 500 classrooms. We see all these studies as well as first person studies of whole courses of studies as exemplifying how the course of instruction can be the unit of analysis, the unit about which assertions are made. In that kind of classroom research, smaller sized events such as lessons, episodes, or turns of talk may be used as instances to illustrate claims about that larger unit – the course of studies, about which the aim of the research is often descriptive. A related genre of studies has inspected classrooms with the focus of understanding the teaching done by individuals – often within a framework that pays attention to individual traits such as beliefs or knowledge (for a review of this work see Philipp, 2007). The ways in which these studies have looked into classrooms resemble those of classroom researchers interested in describing instructional contracts – see for example Cooney (1985) or Thompson (1984).

2. Design research in classrooms

Classroom research has also been done in connection with curriculum development and the design of opportunities to learn. Simple examples of this kind of classroom research have existed for decades and in connection with curriculum implementation. The aim of that kind of research may range from tracking the implementation of curriculum interventions to describing and explaining transformations of the curriculum materials through their handling in the classrooms (e.g., Confrey, Castro-Filho & Wilhelm, 2000; Remillard, 1999; Arbaugh, Lannin, Jones & Park-Rogers, 2006). More elaborated examples of this kind of classroom research include the investigation of the learning opportunities created for students through the enrichment of curriculum materials through their use in classroom interaction (as exemplified in the design experiments conducted by Cobb, McClain & Gravemeijer, 2003). Some research has used curriculum interventions to examine the response of the didactical contract. An example of this is the work of Herbst (2003, 2006) using the implementation of a problem-based replacement unit on area to examine various kinds of negotiation of the didactical contract. The paper by Boero and his colleagues can be considered a part of this group. Finally, we note that the work initiated by Stein, Grover, and Henningsen (1996) on the analysis of tasks has shown how research on designed mathematical tasks can aggregate across many classrooms through focusing on the changes to the cognitive demands observed along the implementation from task design to students' enactment.

In each of those examples, the scale of the unit of analysis may depend on the scale of the intervention. The nature of the research is often comparative, between an analysis a priori of what the materials might demand or afford and an analysis a posteriori that inspects the details of what happened when the materials were used. Sampling decisions will often depend on the designed materials – for example researchers may record the instruction time taken to cover a particular curriculum, unit of study, or task of interest, no matter whether that is contained in a few minutes within a lesson, over several days of instruction, or several times across the year. The role of the designed materials is to establish a developmental timeline: Thus the sequence of topics or tasks or possible decisions available to the teacher serves as a way to order the data collected. Furthermore it serves as a way to select which data to collect. The data can be examined against the background of the designed materials, parsed according to the anticipated divisions in the designed materials (e.g., the time taken to work on task 1 is separated from the time taken to work on task 2). Thus the continuous development (in the form of microgenetic learning or negotiations of task) that could be observed across the record can be naturally separated by the anticipated divisions in the materials and possibly compared.

3. Classroom research that investigates the nature of mathematics lessons

International comparisons of instruction, such as the TIMSS Video Study (Stigler, Gallimore & Hiebert, 2000; Hiebert et al., 2003) or the Learners' Perspective Study (Clarke, 2006; Clarke, Keitel & Shimizu, 2006) illustrate yet a different unit of analysis – the single day lesson. In contrast with the first kind of study where instances (perhaps lessons) illustrate presumably larger, more stable characteristics of the contract, this kind of research aims at making empirical claims about instruction based on the aggregation of features observed recurrently in lessons. These go from very basic features such as average time spent on mathematics content in lesson in a country to more complicated comparison of kinds of activity structures used in lessons across countries. The notion that there are *cultural scripts* for lessons in different countries (Stigler & Hiebert, 1999) has been proposed to explain the differences found – for example the finding that unlike lessons in other countries, Italian lessons include segments where students are quizzed orally while at the board (Santagata & Stigler, 2000; Santagata & Barbieri, 2005). These cultural scripts describe lessons as patterns of activity types. The notion of activity type (or activity structure), in turn, refers to the different patterns of interaction that participants of classroom activity divide their interaction into. They can be characterized mainly structurally (e.g., triadic dialogue) or also by their purpose (e.g., going over homework). Lemke (1990) contains an inventory of activity types observed in science classrooms, more complex and comprehensive than anything yet done in mathematics education; a comparable inventory of activity structures in mathematics classrooms would be a good basis on which segmentation of lessons could be done for the purpose of comparing lesson structure.

Also focusing on lessons as units but looking deeper into how activity structures are produced, scholars using a cognitive science approach to the study of teaching have created cognitive models of the teachers' action and decision making (Leinhardt & Greeno, 1986; Schoenfeld, 1998). Schoenfeld (2007), for example, sought to create a cognitive model that, when fed with the goals, beliefs, and knowledge of an individual teacher, could reproduce the actions that the teacher did in a lesson, where these actions were described at a grain size commensurate with that of the individual utterance. A comparable modeling grain size done yet from a different theoretical perspective is offered by Simon (1995) who proposed the notion of hypothetical learning trajectory as an operational way of describing the decisions of a teacher within a lesson. Yet other researchers have examined lessons and segments within lessons using sociolinguistic or sociological approaches. An important example is found in Voigt's (1985) work on patterns and routines in classroom interaction. Another example of using sociological methods to examine segments of classroom interaction is shown by Herbst et al. (2009), where the notion of instructional situation is used to segment classroom work according to the role that work plays in instructional exchanges (see also Chazan & Lueke, 2009). Theories from psychology and sociology have brought tools to complexify the way we segment lessons and the way we establish subunits within those segments. We dedicate a special section to this below after we discuss instruments and techniques for data collection.

II. DATA COLLECTION INSTRUMENTS, TECHNIQUES AND THEIR RELATED PROBLEMS

Research on classroom instruction in mathematics has been ongoing since the sixties in mathematics education. The instruments and techniques available for data collection in classrooms have changed dramatically since. Those changes have not been only technological; the changes in available technologies have helped raise issues that are appropriately called methodological – that is, relative to a discourse of method.

The fundamental data collection problem in classroom research is concerned with the adequacy of the fit between, on the one hand, the real-world events that interest us and, on the other hand, the instruments and techniques available to capture them. Of course, to the extent that these "real world events" interest us, they do so through a more or less explicit intellectual apprehension: These can range from mere intuitions carried through language to fully developed theories (Silver & Herbst, 2007). Yet even after such theoretical or proto-theoretical construction of the events of interest, the extent to which our data collection instruments (technologies, procedures, etc.) are adequate to inform us well about the objects of interest requires examination. Often such examination of the adequacy of instrumentation for research can bring important questions to the conceptualization of the object of study.

Mathematics educators who are interested in classroom activity are not merely interested in the outcomes of classroom activity, say for example the amount of student learning, but in the activity itself – in what goes on in those classrooms. The mathematics being transacted among students and with their teacher is carried (and shaped) by the interaction among those people. And classroom interaction happens over long, uninterrupted intervals of time; within those intervals it tends to be complex (that is, many related actions by many actors happen at the same time), multimodal (communicative action uses speech, gesture, body position and movement, etc.), and fast paced (instruction-relevant events happen on a timescale of the order of the fraction of a second; Lemke, 2000). Walter Doyle (1986) has described classrooms as characterized by

> 1. *Multidimensionality*, which refers to the large quantity of events and tasks in classrooms….
> 2. *Simultaneity,* which refers to the fact that many things happen at once in classrooms.…
> 3. *Immediacy,* which refers to the rapid pace of classroom events. …
> 4. *Unpredictability,* which refers to the fact that classroom events often take unexpected turns….
> 5. *Publicness,* which refers to the fact that classrooms are public places and that events, especially those involving the teacher, are often witnessed by a large portion of the students….
> 6. *History,* which refers to the fact that classes meet for *5* days a week for several months and thus accumulate a common set of experiences, routines, and norms, which provide a foundation for conducting activities for the rest of the term or year…. (p. 394-395)

Techniques that might be appropriate for collecting data of other activities (e.g., the collection of material artifacts produced over time by a culture in archaeology, time-lapse photography for plant growth) might imply serious data losses in this context; other techniques might occasion serious disruptions or denaturalizations of the activity itself (e.g., asking participants to stop action and repeat what they did so that it can be observed again, or asking participants to share their thinking out loud as they do their work).

Some mathematics classroom researchers have used live observers furnished with coding sheets in which they would note the presence or tally the number of occurrences of events of interest (see for example Carpenter, Fennema, Peterson, Chiang & Loef, 1989). This tradition has used techniques derived from the study of animal behavior such as 0/1 sampling, time sampling, and event sampling to at the same time collect events and reduce them into data (see Pellegrini, 2004). Alternatively, researchers who were interested in producing more comprehensive accounts of classroom life would use live observers writing field notes, sometimes aided by agendas that identified things to pay attention to (see Boaler, 1998). Both in the case of observation forms and checklists and in the case of field notes, the observer is required to code data at the same time that events of interest are happening.

Recording technology, in the form of audio and video tape or digital video has fundamentally enabled classroom researchers to slow down the process by which the events of interest are transformed into data. They have done that by way of enabling the collection of "records of practice" (Lampert & Ball, 1998) that capture events as they happen in a timeline. To be clear, these records capture reductions of those events, since the technology and its instrumentation make some features of the event visible and others invisible. Audio records, for example, potentially enable the capture of all audible speech but impose the *bias of the microphone location* in distinguishing signal from noise and in relating signal to signal: Thus, speech produced closer to the microphone is likely to be heard better than speech produced farther, regardless of its relevance to the communication recorded. Likewise audio records collapse all concurrent communications onto the same timeline, creating the fiction that utterances spoken *close to each other in time* may be turns of talk incident in *the same* conversation. The inferences thus made possible from records may or may not be true, but the point is that while actual participants in conversation have other means of deciding on the fly whether concurrent talk belongs in the same conversation (for example they may check on body orientation and facial expression to ascertain whether an utterance is directed to them), those means are lost for the listeners of an audio recording. Technical sophistication can help reduce some of those biases and eliminate artifacts – for example, the use of multiple microphones feeding separate tapes and careful mapping of the classroom and its participants' displacements can help ascertain which utterances go together in the same conversation. Yet the general point persists – any record can enable artifactual inferences.

Video records, likewise, introduce the *biases of the camera* (position, zoom, and movement). Rogers Hall's (2000) essay "Videorecording as theory" contains a compelling critique of these biases. As an example note that the extent to which the dominant communication modality (say, the speech of the teacher) feeds from subdominant, concurrent modality (say, students' facial expressions) can be lost by a camera that models the behavior of a standard student (who, say, pays attention to the same object that students in the class are expected to pay attention to). On the positive side, video records enable the capture of multimodal communication and an appropriate array of video cameras can help not only capture the complexity of classroom events but also reduce the data collection bias imposed by just one camera. In this regard it is illustrative to compare the videorecording protocol of the TIMSS video study (Hiebert et al., 2003) with that of the Learners' Perspective Study (Clarke, 2006). Technology products like DIVER (which includes a 360º videorecording equipment; Pea, 2006) are promising though they bring with them some serious questions about the feasibility of recording classroom events in naturalistic settings – considering that deploying such complex equipment in a regular school classroom runs into serious constraints of space and time.

Classroom research in mathematics education around the world seems to have appropriated the idea of collecting video records. As noted above, however, records are not yet data. While good record collection follows protocols that are designed having in mind the ulterior use of those records and often such design entails an embryonic use of theory, records are not data. Rather, data comes out from the systematic inspection of those records with instruments. The instruments used to "read" those records and produce data can be more or less sophisticated and more or less theory-based. Hill, Blunk, Charalambous, Lewis, Phelps, Sleep & Ball (2008) account for one such instrument, the "mathematical quality of instruction" protocol, which has been developed, validated, and used to collect data from video records.

Recording equipment and routines are also limited by budget, deployment resources, personnel skills, and managerial capacity – thus it is hard to sustain standards for recording that approach those of TV studios. And absent good ways of integrating multiple records together and reducing them into data, high expenses in recording may not be matched by large gains in reporting. Thus, while current industry standards and access do recommend some kind of rich media recording, researchers need to be very deliberate in designing recording procedures based on the way the records will be reduced into data, aggregated, and analyzed.

A good instrumentation of classroom research will design protocols for recording that depend on a theoretically grounded model of the events being recorded – this model should anticipate the body behaviors, communication modalities, and interaction patterns that need to be captured and it should acknowledge those that can be lost. The model should be used to specify recording equipment (e.g., cameras, audio recorders), their locations and recording protocols (what to frame, when to rotate the camera around the tripod, when to zoom, etc.), and what else to collect (e.g., individual students' written productions at the end of the lesson, time lapse photographs of their productions through the lesson, etc.). The model should also be used to anticipate how records might be complemented, using for example a database or multimedia presentation that enables the combination of different records and the production of derivative records (e.g., transcripts, descriptions of non verbal behaviors, etc.). And such model should also anticipate how possibilities and needs for later analysis (e.g., whether videos would be scored as complete units, or segmented into activity-structured units and then examined in detail) are related to earlier decisions in recording. For the time being it is safe to say that the field knows a lot more about what goes on in classrooms that might be relevant for research on mathematics instruction than the market has the capacity to record faithfully, unobtrusively, and with rapid deployment. It is also safe to say that as of today and in the developed countries the field has better access to equipment and technical expertise than we have access to good protocols for instrumentation and aggregation of records. Thus there is a lot of room for progress in elaborating recording methodologies. A really important technical problem for classroom researchers is that of turning classroom theories into reasonable and effective instrumentation and aggregation standards. While recording equipment has created the opportunity to postpone the design of precise instruments for data coding to be used as the classroom events unfold, it has not and will unlikely ever relinquish the need for purposeful design of record collection.

III. TRANSCRIPTION AND OTHER DERIVATIVE RECORDS – RETURN TO A DISCUSSION OF UNITS OF ANALYSIS

The availability of continuous records of classroom interaction as opposed to data artifacts collected on the fly has made it possible to delay data reduction to after the events have passed. This in particular allows for the production of derivative records out of primary records – including in particular *representations* of those records, such as transcriptions of verbal content and descriptions of nonverbal behaviors. Important questions arise in regard to the units for which such derivative records are to be produced and in regard to the standards for those derivative records. We will not attempt to exhaust these issues, just provide the elements that define them as problematic and suggest possible ways of thinking about those problems. The first issue takes us back to the question of units of analysis.

1. Producing derivative records of what?

The lesson (the time spent in mathematics class for a given day) is a commonly used unit for the segmentation of records, and it illustrates in a minor way why the definition of units of analysis requires policies for data collection. One of these policies used by many researchers takes the mathematics lesson as defined by the length of time that students spend in the mathematics class (or with the mathematics teacher). This policy is useful in secondary or university classrooms because students take their instruction in different disciplines from different teachers and hence their official schedules already include a well-defined segmentation – marked for example by the sound of a central bell or the move to a particular room. The sociological approach known as ethnomethodology has provided the notion that participants of complex social practices not only engage in activities but also ordinarily produce accounts of the activity they are doing, thus sending signals to each other about what is expected to happen; this notion can help delimit units for data collection. In elementary or preschool classrooms, where sometimes the boundaries between lessons on different subjects are not clearly demarcated, those divisions can be derived from looking at the boundary markers that are created by the participants themselves. These boundary markers may include the teacher's movement from one place to another in the room, the request to put some resources away and gather others, etc. While these boundary markers may occur at various times, the concurrence of several of these boundary markers is a likely signal that major changes in the nature of activity are looming; these can be used to parse the stream of real world events into units (see Leinhardt & Steele, 2005, for more on boundary markers).

An important decision to make when producing derivative records is the extent that these derivative records will cover. The transcription of speech and description of nonverbal behavior in a whole lesson may be an expensive proposition. Depending on the focus of the study, it may not be a requirement either. But in case that there are moments in that lesson for which more detailed records of speech and action are required, or moments where precise details of blackboard content are needed, one question to answer is what the boundaries are of the segment that includes the moment of interest and for which derivative records need to be produced. While scholars interested in mathematics teaching and learning may think it natural to make such decision on thematic grounds, the possibility to engage in any kind of thematic analysis depends on a reliable and rich data source. And that rich data source is likely to include those detailed records. To avoid that circularity and have a defensible starting point, researchers can use the same heuristic noted above to establish lessons in elementary school settings. Namely, they can identify boundaries for segments on the basis of aggregation of changes in the interactional patterns observed: changes in the set of key participants, in the median duration of interventions, in the declared topic of discussion (by the participants), in the artifacts being attended to, etc. An aggregation of changes of that sort indicates a change in activity and suggests a possible boundary for a segment of instruction. Transcripts and descriptions can be developed of segments so delimited and then if later analysis shows those as thematically dependent of moments not yet transcribed, one has a firm basis to call for their transcription.

2. Standards for the production of derivative records

In terms of the standards for transcription, the ethnomethodological tradition has again provided with tools to produce transcripts that map onto printed text many of the prosodic and nonverbal elements of human interaction. The transcription conventions used by conversation analysts, such as those developed by Gail Jefferson (see Sacks, Schegloff, & Jefferson, 1974) provide the means to encode pauses, interruptions, increases or decreases of intonation, overlapped speech, and so on. Forrester & Pike (1998) illustrate how transcripts of that nature can be useful to mathematics educators who do research in classrooms. Transcripts of that nature are important in order to examine how utterances derive meaning from their immediate context; they can be particularly useful to discern whether an utterance that is audible for the observer has actually been heard (and perhaps considered) by the other participants – an important problem in deriving classroom data from continuous records. But producing these transcripts is costly, and it is not at all clear that all studies of mathematics classroom interaction benefit from such transcripts. On the other hand, classroom interaction contains much nonverbal information, in particular the development of blackboard content over time and the gestural interaction between participants and inscriptions (such as diagrams or equations). These are usually required in order to make sense of the data (see Rasmussen, Stephan & Allen, 2004 for an example of how gesturing in a differential equations class was analyzed).

An open problem in research in mathematics classrooms is that of developing or adapting standards from conversation and multimodal transcription to serve the purposes of analyzing classroom interaction. The problem includes not only the design of signs to map many communication modalities into print (or multimedia), but also the design of conventions for translating the timescales and timing of real world events into the sequencing of print materials (or the timed sequencing of multimedia materials). In the context of project ThEMaT (Thought Experiments in Mathematics Teaching; Herbst & Chazan, 2003a, 2003b; Herbst & Miyakawa, 2008), we have been involved in developments that contribute to such design. We have used cartoon characters to narrate classroom episodes in the form of comic books and animations (these can be seen in ThEMaT's Researchers' Hub at http://grip.umich.edu/themat). The idea that a classroom episode could be transcribed using a comic book format enables the researcher to attend to gestural expression and artifact production (diagrams, equations, or written text) over time in a less complicated way than if those had to be described using natural language. While our development has focused in the production of the content of those episodes, the design of the graphics and of the sequencing of content have emerged as substantive area for research.

The research on standards for representing classroom interaction requires the design of signs that can be reliably employed to transcribe dense records of classroom interaction (e.g., digital video files of classroom work) into sparser, simpler representations of teaching (e.g., cartoon or slideshow representations). Problems on this front are not just graphic design ones but theoretical as well – how to justify the preservation of some semiotic differences while one collapses other differences among objects in the real world? For example, one of our character sets (which we call *ThExpians B*) represents humans (teacher and students) as blue, two dimensional, abstract figures. Facial expressions are limited to combinations of very few geometric shapes used to represent eyes, mouth, and eyebrows, and characters don't have arms but just a big hand with no fingers. While such characters are useful to represent action over time in secondary school interaction (where a lot of what is there to represent is verbal), they would hardly be adapted to represent younger students' nonverbal activity such as counting or manipulating concrete representations. The analysis of embodied mathematical activity, where gesture and bodily expression may be used to think with, requires preserving those by way of semiotic choices. If a graphics-based system of representation of classroom interaction were to replace transcription and description and take good care of such multimodality, it would be important that standards for representation balance the need for semiotic economy and reliability with the need to allow wide and valid usability. For that purpose, it is required to find a cartoon model human capable of representing economically and validly the multiple kinds of mathematical semiosis in which humans may engage in a classroom. Of course to further this methodological agenda, more basic research needs to be done to inventory the ways in which students may use body position, gestures, movement, and other modalities to make mathematical meanings in class.

In addition to their limitations in accounting for the multimodality of classroom interactions, transcripts of speech often differ from written text in regard to the ways they maintain cohesion (including how speech establishes and preserves reference). Spoken clauses often elide words or trail off to the point of being inaudible. Spoken clauses also abound in indexicals, which are often resolved through gesture or facial expression. Elided and indexical speech are common in classroom conversations, and transcripts of speech content are not too useful unless those indexicals and elisions can be resolved. This is an especially pressing need for mathematics classroom researchers who are usually interested not in *how* participants produce classroom conversations but rather in *the meanings* that they create with those conversations. But how can vagueness and ambiguity in classroom conversation be resolved? In our work we have found it useful to rely on the tools of systemic functional linguistics (SFL; Halliday & Matthiessen, 2004) for heuristics (particularly the notions of cohesion analysis and participation analysis) to resolve indexicals and elided speech. Halliday's linguistics is designed to account for the linguistic choices that speakers have available to make meanings; in particular, it postulates that one of the metafunctions of language, the ideational metafunction, is to represent the world (the two other metafunctions are interpersonal and textual, which allude to language as a set of resources to construct human relationships and as a set of resources to construct genres of text). In particular, the theory examines the *linguistic clause* as a building block in representing the world. Insofar as the clause functions to represent the world, it can be modeled by attending to what Halliday calls Process and Participants. According to systemic functional linguistics, speakers' linguistic choices encode "processes" of six possible kinds (material, mental, behavioral, verbal, relational, and existential). Each of these processes, the theory argues, is tied to a particular set of relationships with other elements in the clause (which the theory calls Participants and Circumstances). In particular, for each kind of process there are one or more sets of participants that ordinarily go with them. Thus a *relational* process, for example, can be involved in a clause along with Carrier and Attribute participants for an “attributive” clause such as “This triangle is acute” or with Token and Value participants for an “identifying” clause such as “Those hash marks tell you the sides are congruent.” A survey of SFL theory is beyond the scope of these comments, but we provide these pointers to justify our claim that the theory helps resolve elided speech or indexicals in transcription by suggesting a heuristic. The heuristic we propose consists of identifying the kind of process involved in the transcribed clause, then using the SFL theory to identify what kind of participants could be involved in the clause along with that kind of process, and then using those participants to code the verbal content present in the transcribed clause – the kind of participants left to identify from nonverbal or intertextual information could then be a sufficiently narrower problem. Clearly, it is often easy to make those identifications intuitively (by asking oneself what they meant to say when they elided speech or what they meant to point to when they use an indexical such as “this”) and in that case the heuristic presented could be useful as evidence to justify why such interpolations have been made. The result is often something that does not resemble the actual speech content, but something that establishes it – and it underscores the role of theory in method.

IV. IN CONCLUSION

It is beyond our capacity and beyond the allowances of this introduction to deal in any comprehensive way with the methodology of research in mathematics classrooms. We have, however, offered a few pointers and exposed a few problems in this area. While we continue to depend on other disciplines and industries for the technology and the basic concepts to employ in the instrumentation of classroom research, there also are methodological questions of particular interest to mathematics educators that call for attention.

REFERENCES

ARBAUGH F., LANNIN J., JONES D., PARK-ROGERS M. (2006), Examining instructional practices in core-plus lessons: Implications for professional development. *Journal of Mathematics Teacher Education,* 9, 517-550.

BALL D. L. (1993), With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *The Elementary School Journal*, 93(4), 373-397.

BALL D. L. (2000), Working on the inside: Using one's own practice as a site for studying mathematics teaching and learning. In Kelly, A. & Lesh, R. (Eds.).

Handbook of research design in mathematics and science education, (pp.365-402). Mahwah, NJ: Erlbaum.

BALL D.L., BASS H. (2003), Toward a practice-based theory of mathematical knowledge for teaching. In Davis, B. & Simmt, E. (Eds.), *Proceedings of the 2002 Annual Meeting of the Canadian Mathematics Education Study Group*, (pp.3-14). Edmonton, AB: CMESG/GCEDM.

BALL D.L., THAMES M.H., PHELPS G. (2008), Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389-407.

BOALER J. (1998), Open and closed mathematics: student experiences and understandings. *Journal for Research in Mathematics Education*, 29(1), 41-62.

BRUNER J. (1995), On learning mathematics. *Mathematics Teacher,* 88(4), 330-335. (Original article published 1960).

CARPENTER T. P., FENNEMA E., PETERSON P. L., CHIANG C., LOEF M. (1989), Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*, 26(4), 499-531.

CHAZAN D. (2000), *Beyond Formulas in Mathematics and Teaching: Dynamics of the High School Algebra Classroom*. New York: Teachers College Press.

CHAZAN D., BALL D. L. (1999), Beyond Being Told Not to Tell. *For the Learning of Mathematics,* 19(2), 2-10.

CHAZAN D., LUEKE H. M. (2009), Exploring tensions between disciplinary knowledge and school mathematics: Implications for reasoning and proof in school mathematics. In Stylianou D., Knuth E. & Blanton M. (Eds.), *Teaching and Learning Mathematics Proof Across the Grades: A K-16 perspective*. Routledge.

CHAZAN D., SCHNEPP M. (2002), Methods, goals, beliefs, commitments, and manner in teaching: Dialogue against a Calculus backdrop. In Brophy J. (Ed.), *Advances in Research on Teaching, Vol. 9: Social Constructivist teaching* (pp.171-195). JAI Press.

CLARKE D. (2006), The LPS research design. In Clarke D., Keitel C., and Shimizu Y. (Eds.) *Mathematics classrooms in twelve countries* (pp.15-36). Rotterdam : Sense Publishers.

CLARKE D., KEITEL C., SHIMIZU Y. (2006), The learners' perspective study. In Clarke D., Keitel C. & Shimizu Y. (Eds.) *Mathematics classrooms in twelve countries* (pp.1-14). Rotterdam : Sense Publishers.

COBB P., MCCLAIN K., GRAVEMEIJER K. (2003), Learning About Statistical Covariation. *Cognition and Instruction*, 21(1), 1-78.

COONEY T. (1985), A Beginning Teacher's View of Problem Solving. *Journal for Research in Mathematics Education,* 16(5), 324-336.

CONFREY J., CASTRO-FILHO J., WILHELM J. (2000), Implementation Research as a Means to Link Systemic Reform and Applied Psychology in Mathematics Education. *Educational Psychologist,* 35(3), 179-191.

DOYLE W. (1986), Classroom organization and management. In Wittrock M. C. (Ed.), *Handbook of research on teaching* (3rd ed. pp.392-431). New York: Macmillan.

FORRESTER M., PIKE C. (1998), Learning to estimate in the mathematics classroom: A conversation-analytic approach. *Journal for Research in Mathematics Education,* 29(3), 334-356.

GREGG J. (1995), The Tensions and Contradictions of the School Mathematics Tradition. *Journal for Research in Mathematics Education,* 26(5), 442-466.

HALL R. (2000), Video recording as theory. In Kelly A. E. & Lesh R. A. (Eds.), *Handbook of research design in mathematics and science education* (pp.647-664). Mahwah, NJ: Erlbaum.

HALLIDAY M.A.K., MATTHIESSEN C. (2004), *An introduction to functional grammar* (3rd edition). London: Hodder Arnold.

HEATON R. (2000), *Teaching Mathematics to the New Standards: Relearning the Dance*. New York: Teachers College Press.

HERBST P. (2003), Using novel tasks in teaching mathematics: Three tensions affecting the work of the teacher. *American Educational Research Journal,* 40(1), 197-238.

HERBST P. (2006), Teaching geometry with problems: Negotiating instructional situations and mathematical tasks. *Journal for Research in Mathematics Education,* 37(4), 313-347.

HERBST P., CHAZAN D. (2003a), ThEMaT: Thought Experiments in Mathematics Teaching. *Proposal to the National Science Foundation, Education and Human Resources Directorate, Division of Elementary, Secondary, and Informal Education, Teachers' Professional Continuum Program.* Awarded as NSF grant ESI-0353285.

HERBST P., CHAZAN D. (2003b), Exploring the practical rationality of mathematics teaching through conversations about videotaped episodes: The case of engaging students in proving. *For the Learning of Mathematics*, 23(1), 2-14.

HERBST P., CHEN C., WEISS M., GONZÁLEZ G., NACHLIELI T., HAMLIN M., BRACH C. (2009), "Doing proofs" in geometry classrooms. In Blanton M., Stylianou D. & Knuth E. (Eds.), *The teaching and learning of proof across the grades : A K-16 perspective*. Routledge.

HERBST P., MIYAKAWA T. (2008), When, how, and why prove theorems? A methodology for studying the perspective of geometry teachers. *ZDM Mathematics Education,* 40, 469-486.

HIEBERT J., GALLIMORE R., GARNIER H., GIVVIN K., HOLLINGSWORTH H., JACOBS J., CHUI A., WEARNE D., SMITH M., KERSTING N., MANASTER A., TSENG E., ETTERBEEK W., MANASTER C., GONZALES P., STIGLER J. (2003), *Teaching mathematics in seven countries: Results from the TIMSS 1999 Video Study.* Washington, DC: National Center for Education Statistics.

HILL H. C., BLUNK M., CHARALAMBOUS C., LEWIS J., PHELPS G., SLEEP L., BALL D. L. (2008), Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction,* 26(4), 430-511.

LAMPERT M. (1990), When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal*, 27, 29-63.

LAMPERT M. (2001), *Teaching problems and the problems of teaching.* New Haven, CT: Yale.

LAMPERT M., BALL D. (1998), *Teaching, multimedia, and mathematics: Investigations of real practice*. New York: Teachers' College Press.

LEMKE J. (1990), *Talking science: Language, learning, and values.* Westport, CT: Ablex.

LEMKE J. (2000), Across the scales of time : Artifacts, activities, and meanings in ecosocial systems. *Mind, Culture, and Activity,* 7(4), 273-290.

LEINHARDT G., GREENO J. G. (1986), The cognitive skill of teaching. *Journal for Educational Psychology*, 78(2), 75-95.

LEINHARDT G., STEELE M. (2005), Seeing the complexity of standing to the side: Instructional dialogues. *Cognition and Instruction,* 23(1), 87-163.

NCTM. (1989), *Curriculum and evaluation Standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

NCTM. (2000), *Principles and Standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

PEA R. (2006), Video-as-Data and Digital Video Manipulation Techniques for Transforming Learning Sciences Research, Education, and Other Cultural Practices. In Weiss J., Nolan J., Hunsinger J. & Trifonas P. (Eds.), *International Handbook of Virtual Learning Environments*. Dordrecht: Springer.

PELLEGRINI A. D. (2004), *Observing Children in Their Natural Worlds: A Methodological Primer* (2nd ed.). Mahwah, NJ: Erlbaum.

PHILIPP R. (2007), Mathematics Teachers' Beliefs and Affect. In Lester F. (Ed.), *Second Handbook of Research in Mathematics Teaching and Learning*. New York: Information Age.

RASMUSSEN C., STEPHAN M., ALLEN K. (2004), Classroom mathematical practices and gesturing. *Journal of Mathematical Behavior,* 23, 301-323.

REMILLARD J. (1999), Curriculum Materials in Mathematics Education Reform: A Framework for Examining Teachers' Curriculum Development. *Curriculum Inquiry,* 29(3), 315-342.

ROMAGNANO L. (1994), *Wrestling with Change: The Dilemmas of Teaching Real Mathematics*. Portsmouth, NH: Heinemann.

ROWAN B., HARRISON D., HAYES A. (2004), Using Instructional Logs to Study Mathematics Curriculum and Teaching in the Early Grades. *The Elementary School Journal,* 105(1), 103-127.

SACKS H., SCHEGLOFF E., JEFFERSON G. (1974), A simplest systematics for the organization of turn-taking for conversation. *Language,* 50(4), 696-735.

SANTAGATA R., BARBIERI A. (2005), Mathematics teaching in Italy: A cross-cultural video analysis. *Mathematical Thinking and Learning*, *7*(4), 291-312.

SANTAGATA R., STIGLER J. (2000), Teaching Mathematics: Italian Lessons From a Cross-Cultural Perspective. *Mathematical Thinking and Learning,* 2(3), 191-208.

SCHOENFELD A.H. (1998), Toward a theory of teaching-in-context. *Issues in Education*, 4(1), 1-94.

SCHOENFELD A.H. (2007), On modelling teachers' in the moment decision making. In A. Schoenfeld (Ed.), A study of Teaching: Multiple Lenses, Multiple views. *Journal for Research in Mathematics Education*, Monograph Number 14. Reston, VA: NCTM.

SILVER E., HERBST P. (2007), The role of theory in mathematics education scholarship. In Lester F. (Ed.), *Second Handbook of Research in Mathematics Teaching and Learning* (pp.39-67). New York: Information Age.

STEIN M. K., GROVER B., HENNINGSEN (1996), Building Student Capacity for Mathematical Thinking and Reasoning: An Analysis of Mathematical Tasks Used in Reform Classrooms. *American Educational Research Journal,* 33, 455-488.

STIGLER J., GALLIMORE R., HIEBERT J. (2000), Using Video Surveys to Compare Classrooms and Teaching Across Cultures: Examples and Lessons From the TIMSS Video Studies. *Educational Psychologist,* 35(2), 87-100.

STIGLER J., HIEBERT J. (1999), *The Teaching Gap: Best Ideas from the World's Teachers for Improving Education in the Classroom*. New York : Simon and Schuster.

THOMPSON A.G. (1984), The relationship of teachers' conceptions of mathematics and mathematics teaching to instructional practice. *Educational Studies in Mathematics*, 15, 105-127.

VOIGT J. (1985), Patterns and routines in classroom interaction. *Recherches en Didactique des mathématiques*, 6(1), 69-118.

Author Note: The authors acknowledge valuable suggestions by Gloriana González. The work reported has in part been supported by a National Science Foundation grant ESI-0353285. Opinions expressed are those of the authors and do not necessarily represent the views of the foundation.