Chapter 36
A Dialog on Productive Multivocality

Timothy Koschmann and Claire O'Malley

An Inquiry into Inquiries

We are grateful to the editors for inviting us to comment on this large and fascinating project. Thomas Kuhn (1970), in his well-known treatise on scientific revolutions, argued that science advances when a community of researchers comes to agree upon a research question and a set of methods for addressing it. Taken together, the question and the associated methods represent what Kuhn termed a "paradigm." Despite some early, overly optimistic proposals (e.g., Koschmann, 1996), the research arena that we have come to know as CSCL has yet to develop its own identifying paradigm. It is a scholarly community with a shared interest in learning in settings of collaboration, but with neither an agreement on a focusing question nor a common methodology. The current volume engages this as its central problematic.

Current research in CSCL is incredibly diverse. Some work involves engineering new technologies, some the evaluation of instructional innovations, and some focuses on better understanding practices of sense-making and collaboration. The Productive Multimodality (PM) Project focuses on the latter, but even within this restricted subdomain, a rich variety of research traditions and disciplines are represented. The project organisers sought to engender conversations across these traditions. Their strategy was to invite workers with different theoretical backgrounds to look at a common set of data together. In this way, the PM Project is reminiscent of

Prepared for inclusion in *Productive Multivocality in the Analysis of Group Interactions* (Edited by Dan Suthers, Kristine Lund, Carolyn Rosé, Chris Teplovs, and Nancy Law).

T. Koschmann (✉)
Department of Medical Education, Southern Illinois University, Springfield, IL, USA
e-mail: tkoschmann@siumed.edu

C. O'Malley
School of Psychology, University of Nottingham, Nottingham, UK

D.D. Suthers et al. (eds.), *Productive Multivocality in the Analysis of Group Interactions*, Computer-Supported Collaborative Learning Series 16,
DOI 10.1007/978-1-4614-8960-3_36,

various earlier efforts in the learning sciences to conduct "collaborative analyses" (Koschmann, 2011, p. 9). These would include special journal issues organised around a single set of analytic materials (e.g., Koschmann, 1999; Sfard & McClain, 2002) and book-length projects such as those organised by Cobb (1995), Koschmann (2011), Maher (2011) and Stahl (2009). The current volume reports on five such analytic exercises involving five different data sets and five different sets of analysts. The recruited analysts were charged with locating "pivotal moments" (PP. ch1–4) within the materials provided to them. The project, in this way, is designed to encourage the analysts to focus on the same phenomena and be explicit about their research practices and about the assumptions that inform their work. It is, in short, an inquiry into their methodic inquiries. Is that about the way you read it, Claire?

*Yes. It is interesting to note that in every case, one of the analysts was the person responsible for generating the data and representing it—in several senses—to the other analysts. This seems to have been quite a deliberate strategy, since one goal is to compare (and/or contrast) what would have been the primary researcher's analysis with that of other analysts. So the difference between the primary analyst and the secondary analyst is not merely one of different methodological approaches or theoretical orientations, but also the difference in their relationship to the data (see also Rosé & Lund, Chap. 32). This clearly became both a practical challenge for the project at certain times—and indeed caused frustration, where other researchers disagreed with the way that data were collected or recorded—but also a source of potentially productive multivocality, often for the primary researchers themselves. Another aspect of this difference in relationship with the data surfaces in the observations by the editors about the (usually only tacitly acknowledged if at all) ontological commitments implicit in (a) the data themselves (*i.e.*, what is collected/measured entails some assumptions), (b) how they are recorded (*i.e.*, transcripts already encode certain primary analyses/assumptions) and (c) represented (as in higher order coding schemes). One wonders whether the PM project would have been more or less productive if none of the analysts had a relationship to the primary data?*

The PM Project is very reminiscent of those earlier projects you outline, Tim, but it also connects to certain broader themes in educational research today. Firstly, there is the idea of having multiple researchers all working on a common set of data, often quite large. Over the years there have been many examples of this including, for example, TIMSS and PIRLS,[1] *PISA,*[2] *the US National Household Education Surveys,*[3] *the UK Household Longitudinal Study.*[4] *Secondly, there is a great deal of interest in so-called "big data"—not just in science, but also with respect to social*

[1] http://timssandpirls.bc.edu

[2] http://www.oecd.org/pisa

[3] http://nces.ed.gov/nhes

[4] http://www.esrc.ac.uk/funding-and-guidance/tools-and-resources/research-resources/surveys/understanding-society.aspx

science—data collected by various government agencies, companies, and via interactions through social media. The distinction between the first and second kinds of efforts are that with the second kind, the data are usually collected indirectly and are potentially capable of being mined for all sorts of purposes not originally planned for. Thirdly, there is the idea of aggregating diverse bodies of research in the interest of identifying convergent findings. This lies at the heart of the many meta-analyses and systematic reviews conducted in recent years.

There has been a long tradition in educational research of so-called mixed-methods research, usually taken to be a combination of qualitative and quantitative methods. Although some have argued for this approach to be superior to single method approaches (Jonson & Onwuegbuzie, 2004), others have been more critical (Symonds & Gorard, 2010).

The PM Project shares the aim of the first of these themes—several analysts sharing the same set of data. However, it is at a much smaller scale, both in terms of the size of the dataset and in terms of the number of analysts. Perhaps another difference is that it is probably true to say that, for most of the kinds of examples given under the first theme (i.e., datasets), although not all users of those data would come from the same disciplines or theoretical orientations, they are more likely to agree about method than disagree (e.g., for one thing, all these examples are quantitative datasets). The difference in scale also raises a question about how well the PM approach can "scale up," both in the sense of the size of the datasets and in terms of the numbers of analysts. It seems fairly fundamental to the approach that the different analysts needed a great deal of face-to-face interaction in order to achieve the level of shared understanding that might lead to productive multivocality—and one has to remark that 5 years is a long period of time for the development of such shared understandings. So it is worth asking whether the approach could generalise to larger-scale scientific programmes, or, indeed, whether the outcomes of a multivocal analysis endeavour can be generalised—i.e., perhaps you always have to "be there" and go through the process first-hand?

The PM Project shared some common features with the second theme, in that the whole endeavour was to see whether new understanding could be obtained from analyses for which the data were not originally intended, and we see throughout the book various examples of where secondary analysts saw different phenomena in the data, whether it was because they focused on the nonverbal interactions, or because they aggregated data at higher levels of scale, thus revealing emergent properties (see also Lund et al., Chap. 34). One other possible similarity with this second theme, and a slightly different project to the PM Project, would be where different analysts collect different data from the same context, then try to link those data—or even "swap" their data—in order to perform different analyses.

Finally, the PM Project shares some commonality with the aim of the third theme—i.e., to attempt to converge on common findings from diverse analyses. However, what is distinctive about meta-analyses and systematic reviews is that it is essential to compare like with like—similar data, collected from studies employing similar methods. The whole point about the PM Project is at least to deliberately use different methods of analysis and on the same data. This also raises what is

perhaps a missed opportunity of the PM Project, and that is, to use similar data across each of the knowledge domains. As it is, not only do methods of analysis vary but so do knowledge domains, types of data, educational contexts and age groups. This is discussed in the book as a deliberate strategy in order to maximise coverage. The downside is that there is so much diversity it is difficult to know always whether any difficulties in agreement, or even productive disagreement, are due to the primary concern (different methods of analysis) rather than any secondary features (e.g., context of data collection).

Yes, the theme of pursuing mixed-method inquiry is also often described in the literature in terms of "triangulation" and the method receives passing reference here. For example, Suthers et al. in their discussion chapter mention "concurrent triangulation" (PP. ch31-19) and Law and Laferrière discuss "triangulated validation" (PP. ch35-9). Denzin (1970), who is often credited with introducing the term, spoke of four different kinds of triangulation: *data* triangulation, *investigator* triangulation, *theory* triangulation and *methodological* triangulation. In the way in which this Project was structured we can see the latter three all coming into play. Data triangulation, on the other hand, which involves applying fixed precepts to different datasets is an exact inversion of the strategy employed in this project. Central to the PM Project was the notion of holding a set of data in common.

Five Orienting Questions

As explained in Chap. 2 by Lund & Suthers, each recruited analyst was asked to address five questions within their respective chapters: (1) "What ontological and epistemological assumptions are made about phenomena worth studying, and how can come to know about them?," (2) "What is the analyst trying to find out about interaction?," (3) "In terms of what fundamental relationships do we conceive of interaction? That is, what is your 'unit of interaction'?," (4) "What representations of data and representations of analytic constructs and interpretations capture your research purposes and units in a manner consistent with the theoretical assumptions?" and (5) "What are the analytic moves that transform a data representation into successive representations of interaction and interpretations of this interaction? And, how do these transformations lead to insights concerning the purpose of analysis?" These were presented as "dimensions along which to describe analytic methods" (PP. ch2-1). In other words, depending on how individual analysts choose to answer these questions, their contributions can be positioned within a space of possible approaches to studying interaction.

These are interesting questions, but I could see how they might be difficult for chapter authors to answer forthrightly. The first, for example—where do you even start? As learning scientists, I suppose, we might begin with our theories of learning. Learning, after all, is our most cherished phenomenon. But what is *its* ontological status and how do we come to know about it? Our various and sundry theories of learning have lots of assumptions built into them and bringing these assumptions

into the daylight is probably a good start toward building common foundation for the field.

Let's take the first question concerning ontological and epistemological assumptions. Data are "theory-laden." One issue that arises is what counts as the primary data. Indeed what data are captured (talk, video, data logs) is arguably already suffused with theoretical and methodological assumptions (cf., Hall, 2000; Ochs, 1979). Indeed there was at least one instance in the PM Project where there was significant dispute between analysts about whether the data were the "right" data to collect to begin with! The way in which data are represented is also not neutral to the analysis. A transcript already carries framing assumptions related to both theory and method. What is left out of a transcript (e.g., pauses, timings, accompanying nonverbal behaviour) is as significant as what is represented (words). Several authors also had to grapple with this issue. It isn't necessarily that, for productive multivocality one should agree on how to represent the primary data, just that it ought to be a conscious and explicit choice, recognising that this will affect the analysis.

The second question pertains to the purpose of the analysis. But within the current project, the purpose seems to be dictated, at least in part, by the requirement to locate "pivotal moments" within the supplied data sets. As Suthers recounted in his introduction, in their first attempts to produce collaborative analyses, the organisers were disappointed to find that the analysts were "'talking past' each other" (PP. ch1-4) because each analyst was pursuing a different research question. To avoid this in their subsequent efforts at collaborative analysis, the organisers imposed a requirement that the analysts identify "pivotal moments." What actually makes a moment pivotal, however, was left "explicitly vague" (Garfinkel, Lynch, & Livingston, 1981).

Yes, the vagueness of the notion of "pivotal moment" is regarded as a strength, since the authors claim that it is important that these instances bring out what is pivotal in the eyes of the analyst—indeed at times the authors seem to use this concept interchangeably with the concept of a "boundary object," which is not defined objectively, but surfaces at points of difference between analysts/methodologies. However, at other times the use of the term "pivotal moment" seems more to refer to what is seen as a significant "shift" in the interaction, whether it is new forms of learning or changes in talk or behaviour, which is not quite the same thing as a boundary object. So I wonder whether the concept of "pivotal moment" really did succeed in overcoming the "talking past each other" problem.

Several, if not all, the analysts in this book did use the concept of "pivotal moments," which were usually higher level segments of talk or action that for some reason denoted significant or foregrounded activity—e.g., points at which insights were achieved. Some analysts chose common pivotal moments, but others highlighted different ones.

The third question posed to the recruited analysts had to do with the "units of interaction" that they utilised in their respective analyses. As I will discuss a bit later, there are some fundamental differences between spoken language and textually mediated interaction such as chat or blogging, one of these having to do with how

the interaction is segmented. Segmentation of computer-mediated communication (CMC) is done explicitly by the producer when hitting the "send" button. Face-to-face (F2F) communication, on the other hand, is a little more complicated and it is difficult to formulate any strict rule for what constitutes a turn at talk. If a speaker starts, pauses for a period, and then restarts, is that one turn or two? If one speaker seizes the floor, talking for an extended period without break, is that one turn or several? A number of chapters in the book deal with what have been termed in the CA literature "collaborative turn sequences" (Lerner, 2004). Here, one speaker initiates a statement that is then completed by another. Again, for the purposes of defining a "unit of interaction," does this represent one turn or two? This is not to say that we cannot differentiate turns at talk, because we clearly can. We do so routinely whenever we engage in conversation. But difficulties arise when we have to say exactly how we accomplish this.[5] None of the chapters of this book actually address this problem, but it would appear to be a foundational one for anyone wishing to understand the fundamental organisation of interaction.

At some points in the book the authors talk about "gratuitous differences in data considered." An example they discuss is where analysts differed in terms of whether they analysed "private" as well as public activity. In other words, I take it, some analysts worked with inferred data or units of analysis, whilst others worked only with what could be seen by other researchers in the data presented. This is a related point to yours Tim, in the sense that what constitutes a turn can sometimes be emergent in the interaction. But it is also a slightly different point for me. That is, the authors at some points talk about how it is important to reflect upon differences in what is regarded as the units of interaction, or "appropriate" data—these, after all, provide the "boundary objects" they seek in the PM Project. On the other hand, at times they want to eliminate such differences and it is never really made clear what counts as "gratuitous" versus perhaps what we might call "fortuitous" differences.

The fourth orienting question has to do with representations of interaction and this generally involves transcripts. Let me begin by making a simple observation: transcripts are never complete. This is true for a couple of reasons. First, there is always a certain amount of slippage that occurs in putting spoken language on paper—sometimes there are problems with the quality of the recording, people don't always enunciate clearly, when the subjects are using another language, there may be vagaries of translation, etc. So there will inevitably be varying degrees of certainty attached to every line entered into a transcript. Second, when we begin to include multimodal aspects of conduct such as gaze, gesture, posture, facial expression, etc., the number of things that could potentially be noted is essentially unbounded. This is also true when we just consider all of the possible aspects of delivery (e.g., intonation, stress, pronunciation, dialect, etc.) that could be rendered. Since having a transcript is a prerequisite to doing an analysis, we may not know,

[5] But see Sacks, Schegloff, and Jefferson (1974) for one attempt to provide a formal account.

when making the transcript, how much detail will be required. So, as a practical matter, transcript elaboration and analysis often proceed hand-in-hand. Some things may be initially included, but later dropped when determined to have no apparent value for the analysis. These same elements, however, may turn out to be crucially important to someone else doing some other kind of analysis. This is related to what you were describing previously as the "theory-ladeness" of transcripts. This creates a particular problem in a comparison study in which the secondary analysts are supplied with a transcript readymade. Ideally, one would want to structure the task in such a way that all of the analysts are required to construct their own from scratch. Under such circumstances, the transcripts too would be objects of comparison, just as the analytic findings are. The things these transcripts make visible and the things they elide will determine what eventually gets noticed in the various analyses (see also Suthers' comments on this in Chap. 19).

The point above about all transcripts being incomplete has one exception and that exception arises in computer-mediated, textual interaction. Here what you see is exactly the same as what the subjects saw, with no need for correction or additional annotations. In this case we effectively get a perfect transcript and we get it for free. This, in fact, was the case in two of the five data sets studied in this volume. This seemingly represents a big break for researchers who do this kind of research, but it should also raise a caution—this fundamental difference between transcripts used in studying F2F interaction and CMC interaction ought to give one pause when attempting to make generalisations about and across these very different kinds of interaction. Others (e.g., Garcia & Jacobs, 1999) have made a similar point, but formulated it in different terms.

At times it feels rather difficult to follow the distinctions drawn between data, units of analysis, units of interaction and representations. For some data, for some representations, it is fairly clear (e.g., the data are turns of talk, the representations are codes attached to those turns), but for others, the distinction between primary data and secondary representation is not so clear. Is the transcript the data or a representation of the data, as discussed above. The difficulties in making this distinction are highlighted in the discussion about the use of software tools to "align" different representations. However, software tools also make ontological commitments as much as other (non-software) representations. It is not as if a toolkit is neutral representationally, so if the analysts use a common set of tools to align other data or representations, it behoves the researchers to be explicit about what assumptions are built into the tools.

In the chapter describing the methodological dimensions, Lund and Suthers describe interactional analysis as an iterative process of transforming different one representation of the data into another. The fifth and final question deals with manipulations and analytic moves that transform one data representation into another. This works out in quite different ways, however, in different kinds of research (see also Dyke et al., Chap. 33).

The transformation of one representation into another is most evident in research that entails coding of interactional conduct, research that I will term, for simplicity

of reference, *discourse analytic* (DA).[6] Here units of text or talk are extracted to form a tabulation vector, the transcript giving way to the coded tabulation, which, in turn, may be further worked into additional representations such as correlation tables or comparison charts. These transformations are essentially irreversible because certain information (e.g., sequence, temporality, speaker attribution) is discarded at each turn. DA methods are used in many of the analyses presented here including the chapters written by Chiu (Chaps. 7 and 23); Sawyer, Frey and Brown (Chap. 10); Howley, Mayfield, Rosé and Strijbos (Chap. 11); Law and Wong (Chap. 22); and Howley, Kumar, Mayfield, Dyke and Rosé (Chap. 26). Rather than simple coding, another strategy is to represent the interaction as a graph or network. Here again we see a procedural operation being used to transform one representation (i.e., a transcript or log of posts) into another. Examples of analyses that utilise graph-based representations can be found in the chapters by Oshima (Chap. 12); Teplovs and Fujita (Chap. 21); and Goggins and Dyke (PP. Chap. 29). Again, the transformation is irreversible.

But the view of analysis as iterative transformations of various forms of representation receives its biggest challenge from studies that seek to give direct accounts of the organisation of collaborative interaction. Here the final account is derived through some sort of hermeneutic processing of the transcript and primary data materials. Examples would include Shirouzu (Chap. 5) and Trausan-Matu's (Chap. 6) analyses of the origami classroom, the various "Group Scribbles" studies (Looi et al., Chap. 15; Jeong, Chap. 16; Medina, Chap. 17; and Lund and Bécu-Robinault, Chap. 18), and the analyses of the "cell models experiment" prepared by Stahl (Chap. 27) and Cress and Kimmerle (Chap. 28). As I mentioned earlier, rather than simply serving as a starting point for an analysis, transcripts here tend to evolve in concert with it, serving as a storage medium for all the useful noticing made along the way. In this way, it is not simply used and discarded, but is rather continuously improved and brought into better alignment with the goals of the analysis. The way in which the representation is transformed, therefore, is fundamentally different from that seen in DA or graph-based forms of analysis. This constitutes a tension between these two very different ways of approaching the task of understanding interaction in collaboration. The nature of that tension lies at the heart of what needs to be examined in this "inquiry into inquiries."

A related point is that, for most, if not all analytic methods, the experience and the skills of the analyst are key to producing "good" research. This is especially so for the more qualitative methods that rely on interpretation and a good deal of sophistication in terms of sensitising constructs—it is probably a hallmark of the more critical theoretical approaches, including for example "discursive psychological" approaches. It is also true of the more quantitative and positivistic end of the

[6] My choice of term is admittedly one that invites confusion. 'Discourse analysis' is used in a wide variety of ways in the literature (cf., Brown & Yule, 1983; Cicourel, 1980; Gee, 1999; Fairclough, Mulderrig, & Wodak, 2011; Potter, 2004; Sinclair & Coulthard, 1975). Here I am using it specifically to denote those methods for studying interaction that apply categorization reductionistically.

continuum. There is a skill—arguably an art—certainly a craft, in designing a good experiment, in spotting patterns in data, in knowing which differences are meaningful and which are not. Hence the frustration often with less experienced researchers who are slave to "significant differences" and don't take account of effect size—and even where there are measurably large effects sizes, knowing what is "important" as opposed to "substantial." This sense of the analyst as professional tends to be overlooked in the PM Project. An assumption underlying the whole endeavour is that it is the analysis (method) that is important, and not the analyst.

You comment, Tim, that the transformations of one representation into another are "essentially irreversible," but perhaps one goal of the PM Project should be to develop methods for translating between representations so that it is always possible to backtrack, zoom in and out of levels of data analysis, so that a researcher encountering a higher level of representation might be able to recover the intermediate stages and assumptions that led to the abstraction. This would only be possible if, in the process of translating between representations, every step would be made explicit to a third party. The project has taken a significant step along the way to being able to do this, but some further work is needed to make more explicit the translation steps between the various representations employed by the analysts in the different case studies presented here.

Being Multivocal "Productively"

Claire, a couple of decades ago we were both active in helping to form a community of scholars focusing on issues related to fostering learning in settings of collaboration. The research questions in CSCL drew on different kinds of expertise—pedagogical, psychological, sociological, technological—and, so, the enterprise was from the git-go an interdisciplinary one. The challenge in any such undertaking is achieving some sort of synergy across the different disciplinary perspectives represented. Fundamentally, it is a challenge of creating a conversation in which all parties participate as full members, no disciplinary contribution beholden to another. Over the last 20 years CSCL has grown into a vital research community, one with its own scholarly journal and a biennial international conference. Nevertheless, the problem of managing conversations across disciplinary and theoretical boundaries remains. The PM Project is to be applauded for attempting to address this problem.

It could be observed that all research is to some degree multivocal in that within any publication the voices of many (co-authors, reviewers, past collaborators, former mentors) can often be heard. But in seeking "productive" multivocality, the organisers of the current volume seek to achieve a more radical form of conversation. Here participants are obliged to work with materials that they did not personally gather, to be articulate about the assumptions underlying their methods, and, in some cases, to adopt methods that are foreign to their usual research practices. But there seems to be some ambiguity with regard to how the PM Project is to be taken,

as a template for how research should be conducted in the future, as a tutorial for methodology, or just as an instructive exercise.

I agree. There is also an issue about just what is productive about multivocality—this is never really fully articulated in the book. How does one recognise success? Does it matter as long as you, the researcher, feel it is productive? It could be that the multiple forms of analysis reveal new insights. This could be due to the second analysis providing additional findings to the first, or that two different methods combined reveal emergent findings at some second order level. Or it could be that a second method exposes gaps or problems in the first or reveals underlying assumptions that would otherwise be tacit.

There are clearly some practical issues related to adopting this as a model for future research in CSCL. The researchers involved in this volume are all concerned with documenting aspects of interaction in learning settings. While this is clearly a focus of some work in CSCL, it does not span the full scope of research being done under the CSCL banner, much less the wider range of educational research. Much of the work in education focuses not on process, but rather, following in the Thorndikean tradition (Koschmann, 2011), on learning outcomes. This remains the prevailing paradigm in education research and somehow we need to find a way of bringing it into the conversation as well. There are other issues, as well. Is it feasible to expect every project to take on a team of analysts? Would all principal investigators be willing to expand their projects in this way?

Yes, it isn't entirely clear whether the organisers are truly being prescriptive about this approach and arguing for a new agenda in educational research (as in the design-based research initiative, or those who have advocated mixed-methods research as the gold standard). Some seem to be a little more tentative (e.g., Chap. 30).

At times, however, it seems as though the primary goal of the project was to provide a useful source for research methods training. One way in which this might be taken forward is in further articulating the differences in the different methods and analytic traditions employed in this project. As you say, it is clear that what is represented here in terms of range of approaches is not exhaustive or even representative of the whole range of traditions in the fields of learning sciences more broadly, and CSCL in particular. The authors do not claim that they have done this and are clear that they worked with what they had, for whatever reasons. Nonetheless, it does represent a diverse set of perspectives. It would be interesting to develop this with a crisper overview and categorisation of the methods employed here. The authors enumerate a variety of methods/traditions in both Chap. 1 (PP. ch1-7) and Chap. 31 (PP. ch31-5) (content analysis, conversation analysis, polyphonic analysis, semiotic and multimodal analysis, social network analysis, statistical discourse analysis, computational linguistics, uptake analysis, knowledge building analysis, systemic functional linguistics and so on). Perhaps a useful next step would to build upon this project in order to develop dimensions along which these analyses vary, which methods are convergent, which divergent.

At first blush, several of these approaches might be amenable to grouping into a smaller set of categories. However, it does raise the question—why is there such a

seeming diversity of labels for what might, to outsiders, seem to be just minor variants of the same approach? Here I am tempted to comment upon the broader issue of traditions of professional practice in the learning sciences community. So, for researchers whose primary focus is research methodology—i.e., that is their object of study—there is perhaps a disincentive to hone a particular method and a positive incentive to differentiate one's own method from that of other researchers. After all, originality and distinctiveness is a key criterion for professional advancement in the field. If you are a researcher into methods, there are fewer rewards for incremental advances and more rewards potentially for achieving a paradigmatic shift (as in more citations, associating your name with method x, etc.) So, the goal then is not so much contributing to the field by improving method y, but showing that method x is "better." Perhaps this in part explains the proliferation in the field of analytic methods that have different labels but actually, underneath, may well be just minor variants of the same class of methods. (The same might be said for software tools for analysis—a case of "not invented here.") It would have been interesting to see some reflection of this kind in the book and, in practice, there are hints at this issue at times where the discussants comment upon points of disagreements between analysts about what constitutes as "proper data" or appropriate representations. To use this text as a source book on method, however, would require careful scaffolding. The analyses presented here are somewhat complex and the conclusions are not always clear-cut. Added to this, there is the issue that the primary data are, for all intents and purposes, not available; so evaluating each approach is somewhat difficult.

Yes, that will be a problem for all readers, both novice and expert. Not having access to the primary materials upon which the analyses were made really made it difficult for me to sort out the differences between the various approaches. I am reminded of McDermott, Gospodinoff, and Aron (1978) dictum that all descriptive accounts must minimally provide sufficient access to the primary analytic materials to enable the reader/auditor to evaluate the claims being made. Given that they wanted to include five datasets and seventeen analyses, space was no doubt at a premium, but leaving the data sources out will result in a great loss to readability.

Perhaps, in the end, the greatest contribution of the PM Project will be as an instructive exercise. The editors argue that what is different in multivocality (compared with mixed-methods research) is that it is a reflexive practice (Chap. 32), where the researcher or analyst gains insight into methods themselves. So, one wonders whether it is important to actively engage in multivocal research first-hand and that such insights cannot be gained second-hand—for example through reading this book. On the other hand, what the book contains is a very rich resource for other researchers—whatever their level of experience—for reflecting seriously and systematically on how their approaches, from the framing assumptions underlying their approach to data collection to the assumptions underlying their analyses and interpretations, compare and differ with others in the field.

In conclusion, I feel that this has been a hugely interesting project to undertake and should serve to stimulate much discussion in the field. Although I might be accused of relying on positivist assumptions, if I were to take this project forward,

I would articulate more clearly the range of methods, and take care to ensure that they are representative of the diverse traditions of the learning sciences. I would adhere to the policy of holding the dataset constant, but I would make more explicit what counts as criteria for success (of the multivocality enterprise)—for example, the effectiveness of the different methods for revealing insights into learning outcomes and effective teaching strategies, efficiency of methods (perhaps in terms of resources and in terms of generalisability). Or, apply the same methods to different datasets in order to test the assumptions of the methods.

The authors and the many researchers involved in this project should be congratulated for their courage in embarking on such an ambitious project—one which has already generated several papers in high quality outlets—and in their patience in seeing it through over a long period of time, involving many research teams internationally, and in synthesising a very complex body of work. The field is very much in need of such an ambitious programmatic enterprise and, even if the project has revealed as many problematic issues as "successes," it is still to be counted as productive in my view, for opening up these debates as much as anything else.

References

Brown, G., & Yule, G. (1983). *Discourse analysis*. New York: Cambridge University Press.

Cicourel, A. V. (1980). Three models of discourse analysis: The role of social structure. *Discourse Processes, 3*, 101–130.

Cobb, P. (1995). Mathematical learning and small-group interaction: Four case studies. In P. Cobb & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures* (pp. 25–130). Hillsdale, NJ: Lawrence Erlbaum Assoc.

Denzin, N. K. (1970). Triangulation: A case for methodological and combination evaluation. In N. K. Denzin (Ed.), *Sociological methods: A sourcebook* (pp. 471–475). Chicago: Aldine.

Fairclough, N., Mulderrig, J., & Wodak, R. (2011). Critical discourse analysis. In T. A. van Dijk (Ed.), *Discourse studies: A multidisciplinary introduction* (2nd ed., pp. 357–378). London: Sage Publications.

Garcia, A., & Jacobs, J. (1999). The eyes of the beholder: Understanding the turn-taking system in quasi-synchronous computer-mediated communication. *Research on Language and Social Interaction, 32*, 337–368.

Garfinkel, H., Lynch, M., & Livingston, E. (1981). The work of discovering science construed with materials from the optically discovered pulsar. *Philosophy of Social Science, 11*, 131–158.

Gee, J. (1999). *An introduction to discourse analysis: Theory and method*. New York: Routledge.

Hall, R. (2000). Videorecording as theory. In E. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 647–664). Mahwah, NJ: Lawrence Erlbaum.

Jonson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14–26.

Koschmann, T. (1996). Paradigm shifts and instructional technology: An introduction. In T. Koschmann (Ed.), *CSCL: Theory and practice of an emerging paradigm* (pp. 1–24). Mahwah, NJ: Lawrence Erlbaum Assoc.

Koschmann, T. (1999). The edge of many circles: Making meaning of meaning making. *Discourse Processes, 27*, 103–117.

Koschmann, T. (2011). Theorizing practice. In T. Koschmann (Ed.), *Theories of learning and studies of instructional practice* (pp. 3–17). New York: Springer Science + Business Media.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago, IL: University of Chicago Press.
Lerner, G. (2004). Collaborative turn sequences. In G. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 225–256). Amsterdam, The Netherlands: John Benjamins.
Maher, C. A. (2011). The longitudinal study. In C. A. Maher, A. B. Powell, & E. B. Uptegrove (Eds.), *Combinatorics and reasoning: Representing, justifying and building isomorphisms* (pp. 3–8). Dordrecht: Springer.
McDermott, R., Gospodinoff, K., & Aron, J. (1978). Criteria for an ethnographically adequate description of concerted activities and their contexts. *Semiotica, 24*, 245–275.
Ochs, E. (1979). Transcription as theory. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental pragmatics* (pp. 43–72). New York: Academic.
Potter, J. (2004). Discourse analysis. In M. A. Hardy & A. Bryman (Eds.), *Handbook of data analysis* (pp. 607–624). London: Sage Publications.
Sacks, H., Schegloff, E., & Jefferson, G. (1974). The simplest systematics for the organization of turn-taking for conversation. *Language, 50*, 696–735.
Sfard, A., & McClain, K. (2002). Analyzing tools: Perspectives on the role of designed artifacts in mathematics learning. *Journal of Learning Sciences, 11*(2/3), 153–138.
Sinclair, J. M. H., & Coulthard, M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. London: Oxford University Press.
Stahl, G. (Ed.). (2009). *Studying virtual math teams*. New York: Springer.
Symonds, J. E., & Gorard, S. (2010). Death of mixed methods? Or the rebirth of research as a craft. *Evaluation & Research in Education, 23*, 121–136.