In these comments on Roth/Middleton’s very careful and impressive analysis of interviews in the construction of knowledgeability about a topic in science, I want to first discuss their own theoretical and methodological approach, and then pose some questions about broader implications for teaching and assessment in science education.

The initial point and key examples made by Roth/Middleton are typical of the CA (Conversation Analysis) approach and its inheritance from Alfred Schutz’ phenomenological micro-sociology (or “ethnomethodology”): that talk is a way of accomplishing face-to-face interaction in a local, situated, and contingent process. This perspective is a useful antidote to more foundationalist or essentialist perspectives (as in classical macro-sociology) that reify categories like “knowledge” and “expertise” or even “gender,” “intelligence,” or “social class” into pre-given realities.

Phenomenology in general takes reality to be a rather blurry and messy, dynamic flow which natural and social science tries to corral into a deceptive neatness, putting our own cultural and analytical categories in the place of the reality of primary experience. For instance, we substitute metric time for the experience of duration and call our new “time” a reality, when we have no warrant for claiming it is more real (and plenty of historiographical reasons for knowing that it is our own invention) than the experience of duration or persistent doing. All the more so for our social categories like “intelligence,” “race,” “gender,” “age,” or “knowledge.” What then is the experiential reality that underlies these categories?

The CA school, following the early work of Harvey Sacks and now led in the USA by Emanuel “Manny” Schegloff, more or less contends that ordinary talk is at the heart of how we make social reality seem real to us. I have no basic disagreement with this, but I do see this view as in some respects too limited an approach both to the social functions of talk and to useful ways of conceptualizing social reality.

That talk is constitutive of social reality is as sure as anything in social science. But if we look at talk not as an answer to the phenomenologists’
question about the experiential basis for social categories, but as a phe-
nomenon in its own right, we can see that it has a wide range of functions. 
They are all in some sense social, but not in the obvious way that interper-
sonal interaction is social. Michael Halliday, the leading functional linguist 
of the last generation, identified three (or sometimes four) major functions 
of talk: ideational-thematic, interpersonal-attitudinal, and organizational-
logical. Ideational-thematic functions present us with a world, a state-of-
affairs. This is the function of talk in construing the world as consisting 
of events and processes, actions and relationships, in which there are par-
ticipants (human and nonhuman) and for which there are circumstances 
(of time, place, manner, means, etc.). Then we have the interpersonal-
attitudinal functions, which are closest to those of interest to CA: constru-
ing interaction in terms of exchange, speakers and addressees, rhetorical 
functions and speech acts, degrees of intimacy and formality, superiority or 
derence, attitudes to participants and attitudes to the ideational-thematic 
content of what we are talking about, in the course of also construing our 
relations to those we are talking to (or past). The organizational (Halliday’s 
term is “textual”) functions construe wholes and parts, not of the world, but 
of the text itself: its hierarchy of organizational units, its strands of cohe-
sion running across the boundaries of those units. And the logical function 
enables us to define the kinds of relationships among text units and among 
units of things we talk-into-being in the world: additive and concessive, 
superordinate and subordinate, extensions and elaborations, etc. Each of 
these dimensions is very important to how both scientists and science teach-
ers and students make meaning (e.g., Halliday and Martin, 1993; Lemke, 
1990).

Neither is it all just talk. Talk is easier to analyze because of decades of 
research in linguistics. We also make all these kinds of meaning, and also 
make real-seeming our social categories through the artifacts we produce 
(a major point made first, I think by Marx), the typical kinds of actions 
and activities we do (other than talk), and the ways we deploy resources to 
make visual meaning, music and sound effects, dance, architecture, cuisine, 
fashion, etc. Social reality is made multi-modally.

And finally, it is not all just in the moment and at a single scale of 
social organization. What happens here and now in a conversation does 
not make sense totally in its own terms. It also makes sense in relation to 
other conversations and other events in our lives elsewhere and elsewhen. 
It may be one part of an on-going relationship over a long timescale. It 
may re-enzyme common ways of saying in the community that are typical 
of some social category of people – even if that category is itself partly 
a product of our talk (and also of many other sorts of actions, so that 
changing our talk might make it less salient for us, but not in itself change
other material circumstances of persons formerly so classified). There is a macro-social reality. It is not neat, but it is also not just a way of talking. Reality is multi-scaled in space and time, or we may say it lives across our experience, and not just in each moment of it. The phenomenological philosophers were all very clear about this primary fact, but it somehow got lost in the descent of CA analyses of talk, perhaps initially as a prophylaxis against importing unwarranted social categories, but then a bit too much as a focus on the autonomy of the now. More recently, complex systems theory has rediscovered this key fact about all dynamics: you cannot analyze at one timescale (or extensional scale) only, you must analyze across multiple scales. In complex systems and in human experience, the now, the local is never autonomous, because the local unit of experience is itself one of our excesses of neatness. Reality blurs across time and space, across conversations and events. It is not composed of these artificial units; we decompose it into them.

Meaning arises not just intra-textually, but also inter-textually. It arises not just in the give and take of conversation, but also of gesture and posture, of movement and action. And it arises not just here and now, but across the times and places through which longer-term processes or more extended practices are taking place.

In speaking of “interactional achievements” we need to keep in mind that what is achieved is achieved not just in the present interaction. It is achieved by the deployment of resources that were created and renewed in many other, prior interactions, and not only ones in which we ourselves participated. The resources we deploy to achieve something have histories that matter, their meanings and effects are not totally under our control, and certainly not created totally by us in the moment, though we can greatly modulate their effects by our strategies of deploying them. As resources, they have a certain meaning “potential,” and how that potential gets realized as something very concrete and specific on a given occasion is an interactional accomplishment.

So I am not arguing against the excellent methodology of Roth/Middleton’s paper, but trying to place it into a larger context that shows what CA methods typically leave out or put into background. I certainly agree with the paper’s general stance toward and conclusions about the status of interviews as research data for science education and other fields. Interviews do not produce data about subjects, but data about the whole interview activity itself; they are just as much data about the interviewers and the relationships they co-constitute with those answering their questions (and often enough asking ones of their own). There is more symmetry than asymmetry in interviews, and I think many experienced ethnographic interviewers would agree that the most useful interviews (and
not the ones that are easiest to analyze!) are those where there is the most symmetry: where the interviewees raise topics unanticipated by the interviewers, ask their own questions, and engage in genuine dialogue, not some artificially controlled pseudo-conversation, an oral questionnaire. In fact, as is well known even in sociology, and especially among cross-cultural interviewers, the genre of the interview is something that must belong to the repertory of forms of talk on both sides for even the pseudo-conversation to be possible.

So it is not surprising to me that Roth/Middleton tell us quite bit about the interviewer, Daniel, more really than we know initially about the “subject,” Annemarie. But an emphasis on symmetry might have led us to expect a two-column table giving us parallel and equivalent accounts for each.

The discussion of analytical method makes one particularly important point, I think: that it helps to analyze progressively through the talk, and not read later events or conclusions back into the possible meanings of earlier activity. Only analysts can do this, not the participants. As a participant, every next move is contingent, and over time, things become more and more unpredictable. Of course we have expectations, from other conversations, about likely or typical next moves and even the longer-term shape of a conversational genre. These are part of the meaning potential of the action-up-to-now as a possible instance of some typical longer-term social process. What something means now is usually a bit indeterminate. It is only later that we, as participants, can retrospectively say what it turned out to have meant for us, how it was subsequently construed, how we built on it. Analysts too often take only the retrospective view of an interaction, reading the whole transcript and then analyzing it as we might analyze a play or a novel, in which presumably the author wrote and edited many drafts, so that the end is indeed mixed into the beginning and middle. But as participants we do not know the end, and so the dynamic, or contingent meaning-as-of-now differs from the final, all-done, retrospective meaning of something in the middle. To know the meaning for the participants at the time we analysts must find ways to forego our possibility of knowing how it all turned out (or how any segment turned out).

This is not to say however that dynamic or contingent meaning is all we should pay attention to. There is also synoptic meaning, meaning when-all-is-said-and-done, and that meaning, too, is important to participants, just at a different time. But we must not let synoptic meaning totally replace dynamic meaning in our analyses.

By the time the paper gets to its more detailed and technical analyses (e.g. prosody as interactional resource), I was becoming ever more curious about its potential implications for science education. An analogy can indeed be constructed between interviews and the assessment process,
and then also to the teaching or tutoring process and an IRE (question-answer-evaluation/initiation-response-evaluation) pattern of dialogue. Certainly there are similarities, but also great differences. Neither dialogue partner is actually an expert in ecology, though presumably Daniel knows the standard interpretation and logic for the graph. The questions are being asked by an information seeker and are not “known-answer” questions as in classroom IRE dialogue. The printed instructions set a rather general task, for which there could not be very clear criteria of what constituted a correct answer and for which there would certainly be many sorts of more or less “complete” answers, more or less “sophisticated” answers. In the nearest analogue of an assessment situation, an oral exam, the role of talk in constituting a “level of knowledge” may not end with the interview itself, but continue with the evaluation of the interview tape or transcript by some further assessor, or by the original interviewer. The analogue in this case would seem to be Roth/Middleton themselves. They are the ones offering interpretations about Daniel’s interview with Annemarie.

This points to a potentially important lesson: the total assessment process begins as soon as and keeps going as long as there is talk that is ultimately made relevant to determining the outcome of the assessment. It begins, in this case, with the creation and selection of the ecology graph as an “item,” and ends with the talk in the paper itself (or in this Forum here). In traditional assessment, it begins with the conversations among the assessors (or the curriculum designers, or the state or national standards committees) about what to assess and how, and it only ends with the talk of the assessors about a student’s response to the item (even if in the form of contributions to an oral dialogue), interpreting it and construing it as evidence for some decision about his or her “level of knowledge” or grade.

If I come to this process as a sort of meta-assessor, to assess the validity and reliability of the assessment process that has been conducted, or is being conducted, I need to take this long view of it. I may even need to take into account the possibility or reality of appeals of the grade, of lawsuits and other extensions of talk that continue to construe and re-construe such things as “level of knowledge” or “fairness of process.” As a professional, I had better also be reflexive enough to try to analyze how my own talk is constituting notions such as “validity” and “reliability,” and I had better know where the bodies are buried within the history of my own discipline concerning the limitations and possibly the total arbitrariness of even these named concepts. For they can certainly not be primary experiential realities, nor can the activities of assessment possibly be so neat as to be describable in such terms.

But a reader interested in science education ought to be wondering by now just what the specific applicability of all this might be to our work?
Why is this paper not in a journal devoted to the study of meaning in human behavior, or social psychology, or communication and interaction? Why are Roth/Middleton talking to us about this?

If we skip ahead to the Discussion section, we get a clearer answer to this question. Roth/Middleton are trying to undermine some conventional certainties in our field. Most fundamentally, they are arguing, I think, that it makes no sense to assume that we can assess what students “have in their heads” because (a) it’s not in their heads, it’s created on-the-fly in the interaction situation itself, and (b) any answer we give to the question of what’s “in their heads” is really a statement about a larger system consisting of us, the test, the curriculum, and the conditions under which the assessment takes place. Whatever “it” is, it is a feature of this larger system, and not of the student, and it is not a static feature, permanently present, but a dynamic one that emerges through very subtle management and deployment of all available resources by all participants, in the moment.

This is a claim we have heard before, as of course Roth/Middleton are well aware. We have heard it from the advocates of situated cognition and of distributed cognition, from anthropologists of cognition like Jean Lave, Mike Cole, and Ed Hutchins, from psychologists like John Bransford, from interactional sociologists like Charles “Chuck” and Marjorie Goodwin or Aaron Cicourel, even from social linguists like Ron and Suzanne Wang Scollon. But have we really been listening to all this detailed empirical research?

Why are we content to sit silently by while our educational future is determined more and more by the outcomes of testing and assessment policies that are as antediluvian as “scientific eugenics” (which historically is not at all unrelated)? Why do we keep imagining our students heads filled with little filing cabinets to which we, or they, add new knowledge, to be tested by seeing whether they can find the relevant folder up there during the test hour? Why do we imagine that memory is the model for knowing? Or that indeed there is such a phenomenon as “knowing,” as opposed to just being able to produce the goods on demand, one way or another, and maybe in a different way every time? And especially in science education, where we see memory as the very lowest level of scientific understanding (however necessary), and are aiming for deeper forms of understanding?

We are allowing policy decisions to be made on the basis of test results when there is no coherent and defensible conceptual or empirical basis today for any of the principles on which the psychometric testing movement of the last century was based. Tests are not microscopes peering into students’ heads to see how much knowledge is stored in there. Test results have no predictive value whatsoever for what a student can or cannot do in
non-academic settings to which curricular learning may be relevant. There
is no such thing as “what a student knows.” There is only what a student
can do, and how he or she does it, in some complex interactive situation,
with all the resources at hand, including other people, artifacts, and their
own bodies. The future is going to laugh at our testing regimes the way we
laugh at phrenology, if they do not cry over their social consequences the
way we do over scientific eugenics and racism.

Roth/Middleton are rubbing our noses in the gritty details of how human
beings really deal with scientific graphical-conceptual tasks, at the most
immediate level of experiential reality and largely free of so many taken-
for-granted imaginary beasts like “knowledge,” “ability,” “expertise,” and
even (my addition here) “cognition.” Insofar as “cognition” still means to
most of us something that goes on fairly autonomously “inside the head,”
what the data in this paper present to us is not cognition at all (which is
not to deny that something is relevantly going on neurologically in our
cortex), but talk and action, or in their phrase: talk-in-interaction. Talk-
in-interaction is not cognition’s visible outer form. It is the actual pro-
cess itself, which is misnamed and misrecognized as “cognition” largely,
so far as I can see, because psychologists want to go on calling what
they do psychology, rather than interactional sociology or communication
studies.

The implications here for science education are vast. If we give up the
filing cabinet notion of “knowledge,” then we have no grounds for basing
curriculum on a “valued body of knowledge,” the ideal filled filing cabi-
et. We have to think in terms of what real people do to which scientific
ways of understanding are relevant, or potential things they might someday
want to do. Even “ways of understanding” is still verbiage from the old
mental-machinery language. What scientists actually do professionally is
deploy a lot of specialized tools, from ways of talking, to ways of manip-
ulating equations and drawing graphs, to actual apparatus, to interacting
with computer simulations. Doing so fluently and in ways that get results
– and recognition from other specialists – is what we learn to call “under-
standing.” We learn, and improvise, very complex routines for getting some
particular kinds of performances called “understanding,” and in fact no one
can ever tell you, or predict for sure, whether any given performance will
get called that or not. Our elaborate testing regimes are efforts to produce
an artificial predictability of this kind, at the price of misrecognizing and
misnaming the entirety of what we’re doing.

We cannot be living in a world in which what Roth/Middleton are
showing us happens, and where standards-based testing makes sense.

They cannot make that strong a claim in an academic paper, but that is
exactly what I believe, after more than 30 years of research and experience
both in science education and in many of the other fields on which their research draws.

K N O W L E D G E A B I L I T Y A N D T H E P A S T A N D F U T U R E O F R E S E A R C H INTERVIEWING IN SCIENCE EDUCATION

Gregory Kelly

I welcome the opportunity to be part of the conversation around the ideas put forward in Roth/Middleton’s article, “Knowing What You Tell, Telling What You Know: Uncertainty and Asymmetries of Meaning in Interpreting Graphical Data.” The arguments in this paper leave me with much to consider. In this comment, I will try to situate the ideas in this paper in a historical context and then consider what implications this work has for educational research. I will not comment on the specifics of the many cases presented or try to second-guess the already cogent analyses. Rather, I will partially respond to the questions posed by Lemke regarding the specific applicability of the work of Roth/Middleton for science education. I begin by examining where this study occurs in the history of ideas in science education. I will then shift to consider some implications of this work for further research.

The uses of research interviews have had some prevalence in research paradigms in science education. These interviews typically examine students’ or teachers’ conceptions of some(body’s) knowledge regarding science concepts or putative conclusions about the nature of science or something of the sort. The theoretical treatment and careful analysis in Roth/Middleton’s piece raise a number of questions for the future of studies that posit conceptions in other minds and suppose that these conceptions can be accessed through clinical or research interviews.

Cultural studies of learning suggest that analyses occur across timescales and consider the embedded nature of events (Cole and Engestrom, 1993; Lemke, 2000; Wortham, 2003). The conversation of Anнемarie and Daniel occurred in a specific location in time and space, where the history of ideas on graphing were to some extent invoked through the inscription and the interpretation of the conversation partners. Similarly, the analyses by Roth/Middleton occur at a point in the history of ideas in science education. I consider therefore the value of this study in this history with a focus on language and knowledge specific to research interviewing. To examine the frequency of research interviewing in science education research, I conducted a short study using the Web of Science search engine and analyses features. I ran a search from 1986 through 2005 of key word
“interview∗” in three specific journals, *International Journal of Science Education, Journal of Research in Science Teaching*, and *Science Education*. There were a total of 467 studies mentioning interviews in some way, which amounted to 13 percent of the total publications. The results show that there were only a few manuscripts using interviewing prior to 1992, with numbers increasing in the mid 1990s, leading to approximately 37 manuscripts per year from 1996 to present.

Why this count of uses of interviews? A central premise of the Roth and Middleton paper is that the interview should be treated as a communicative event (Briggs, 1986; Hymes, 1974). This suggests that theories of discourse and linguistics would typically inform analyses of research interviews to ascertain the jointly constructed meanings. It further raises questions about the extent to which analysts using interviewing as a research method have considered the conversational features of the interviewing process in the research designs of these many studies. To examine the extent of linguistics, discourse or conversational analysis in the history of research interviewing, I did a *Web of Science* search for “interview∗” crossed against, in successive runs, “linguistic∗,” “sociolinguistic∗,” “discourse analysis,” “conversation,” and “conversational,” in these same three journals, finding ten articles at the sum of the intersections. This is a crude measure. Nonetheless, it does show some neglect for the interview as a communicative event. If there were a robust research tradition emerging that considered the interview as a communicative event, the numbers of instances would undoubtedly be greater. I believe this shows that with few exceptions much of the perspective offered by Roth/Middleton has not yet been considered in science education research.

Interestingly, the central premise regarding interviews as communicative events predates much of the research in science education employing interviewing as a method. As Roth/Middleton argue, much research on interviewing assumes the existence of a set of stable features in the head of the interviewee, that are drawn out through the process of probing questions by a skilled interviewer. This assumption has been examined and thoroughly refuted through detailed arguments made by Briggs (1986) and Mishler (1986). These two books offer a set of reasons to be deeply suspicious of the assumption of stable features of the interviewee, and both, in their own way, make the case for making the interviewing event both problematic and a research issue in its own right. Drawing from conversational, discourse, and narrative analysis, Briggs and Mishler make clear that researchers need to develop high degrees of reflexivity and, indeed, concern about the conclusions that can be drawn from interviews. Despite this body of research, much of the research employing research interviews in science education has not paid much attention to the conversational features of the
actual interview; hence, the need for the study offered by Roth/Middleton. I thoroughly agree with these authors that what matters in the study of social interaction are not the supposed content of the mind, but the resources deployed in the social interaction (or transaction). Furthermore, I believe it is clear that the argument made by Roth/Middleton concerning conceptions of knowledge and language use has yet to enter into the discourse community of science education.

Roth/Middleton make the case that the interview itself is a communicative event, that conversational analysis can be a tool for understanding the interview event as such, and there is much to be learned through careful analysis of the interview conversations. Some of this argument echoes the work of Briggs and Mishler. Of course, Roth/Middleton go well beyond this basic assumption of the interview as a communicative event and set out to begin a project of “remaking the concept of aptitude” (manuscript p. 78). This remaking holds much potential for how researchers in science education (and beyond) wish to consider issues such as knowing, learning, and being. The uses of linguistic data to show how conversational cooperation, saving face, seeking acknowledgement, dealing with uncertainty, among other issues, are accomplished through transaction and signaled with prosodic cues. The detailed analyses by Roth/Middleton of pitch and intensity demonstrate the ways that such contextualization enters into the jointly produced meaning. The specific applicability of the work of Roth/Middleton for science education should make life more difficult for science education researchers. These difficulties arise from the following five implications.

First, all the taken-for-granted assumptions about interviewees’ knowledge and researchers’ access to such (putative) knowledge need to be re-examined. This has broad implications for how the field considers competency with respect to a whole set of normative goals, including student understanding of concepts, the nature of science, and inquiry processes. Issues such as the normative measures of “scientific literacy” need to be examined from the point of view of literacy as a situationally defined, and interactionally accomplished practice. This suggests empirical studies that include the multi-modal nature of locally constructed contexts. Furthermore, the work of Roth/Middleton points beyond the specifics of a research methodology to challenge the field of science education to reconsider core concepts such as “knowledge,” “ability,” “expertise,” and “cognition,” as suggested by Jay Lemke.

Second, the concept of aptitude needs to be examined in the specific communicative context in question. Notions such as scientific literacy, for example, will need to take into account the situation of knowledge use. Nevertheless, such notions of aptitude in science may be confounded with
the communicative competence of the speakers and assessors of aptitude. For example, Annemarie is able to co-construct a conversation with Daniel in which she solicits his evaluation of her graphical literacy when he is apparently unwilling to offer it, given his role as interviewer. A less apt interviewee may not be able to engage the interviewer in such a way as to draw out subtle clues about the supposedly correct interpretation of the graph. Thus, assessment of knowledge (really knowledgeability) needs to be seen as occurring in some “situated talk-in-interaction” as the authors suggest.

Third, there are implications for theories of knowledge and language for how research gets done and what can be claimed about other people’s minds, actions, products and identity. Research methods that use language (here I include naturalistic observation, interviews, surveys, testing, and analyses of student work) require careful consideration of the many issues raised by the authors – the uses of uncertainty by the conversational partners is just one example. The semantics of science, but also the spaces between the words, need to be part of the research framework. Face value assessment of the “content” of talk in small group, classroom, or interview discourse should be viewed as seriously suspect, given the authors’ analyses following in the conversational analysis tradition. Such difficulties are not solved (of course) by creating objective tests, surveys, or interview protocols. In such cases, the ambiguities of language are merely glossed and taken away from the interactive context where the research subject has some agency in determining in the joint construction of meaning.

Fourth, the social interactions occurring in science classrooms are considerably more complex than the interview dyad conversation. In such cases, and as indeed it was for the case in this interview, the cultural assumptions about communication of the interlocutors matter. Such assumptions about communication are tied to previous experience and are likely to vary by region, class, ethnicity, gender, and race, and other more specific variations in experiences. The variation in the contextualization (e.g., pitch, stress, intonation, pause structures, physical orientation, proxemic distance, and eye gaze) sends messages, variously interpreted with potential for divergence across ethnicity (Gumperz, 1982a). Given the increasingly diverse student populations (worldwide), the potential for cross-cultural miscommunication in science classrooms is great and increasing. The authors demonstrate just how prosody matters for communication and ways that the conversational partners vary their speaking to send subtle cues open to interpretation (see also, Gumperz, Cook-Gumperz, and Szymanski, 1999). Given the variation in student populations in multicultural societies and the potential for misunderstanding in interethnic communication due to prosodic variation (e.g., Gumperz, 1982a,b; Gumperz and Hymes, 1972),
Roth/Middleton identify the extent of the task for analysts seeking to make sense of interviews, classroom discourse, and other interactional events: subtle variation in pitch, stress, phrasing, and register potentially communicate meaning. Ways that language use hold power, and privilege certain ways of being, speaking, interpreting, and so forth, demonstrate the subtle ways affiliation and discrimination are constructed in situ in and through language use.

Fifth, Roth/Middleton demonstrate that assumptions about roles and institution power need to be interpreted within specific (socially constructed) contexts. The variation over time of who was the person “in the know” between Annemarie and Daniel demonstrate how previous experience, knowledge, and familiarity with the interactionally recognized topics compete with assumptions about competence derived from institutional standing and (supposed) relative expertise. This suggests that analysts will have to examine in detail the interactionally accomplished nature of privilege and power as constructed in specific situations and consider how such interactions may aggregate over time to pose political problems and inequalities.

RESPONSES TO KELLY’S COMMENTS

Jay Lemke

I met Eliot Mishler early in my research work for the Talking Science project (introduced to him by Courtney Cazden, who was a wonderful “matchmaker” for me in those days), and so was aware early of some of the issues Kelly raises, which I could also see in my own classroom discourse data. In my current qualitative methods course, students read in some depth about the “revisionist” view of interview data in sources such as Fontana and Frey (2000), Heyl (2001), and the work of Jan Nespor (1997). In general, I think that the anthropological tradition, more so than the sociological (except for the ethnomethodologist minority) has come to see interviews as generating data about the local system they create by happening, within broader cultural norms, rather than as probes into interviewees’ minds. An interesting angle here is questioning the extent to which interviewer and interviewee (or assessor and student, as below) have the same or compatible (or in some cases, any) expectations about the speech genre of the interview itself. In your response, I particularly like the point: “notions of aptitude in science may be confounded with the communicative competence of the speakers and assessors of aptitude.”
In this *Forum* one important topic we might focus on is the issue of to what extent communicative competence and science understanding are *inevitably* confounded (i.e. in principle inseparable), and to what extent research and especially assessment may make invalid claims about “student understanding” or “student aptitude” because it does not normally take into account issues of communicative competence – and match or mismatch of communicative norms and practices – on the part of both assessors and students. For example, we do not teach students very explicitly a lot of the communicative norms and strategies that are necessary to do well in our assessments, and we certainly do not regularly test “understanding” across genres (say across interviews, “objective items,” problem-solving, narratives, explaining things to a younger friend), across dialect differences (e.g. testing in both standardized forms and in ones that are designed to be culturally responsive, or dialect specific) and across language differences (testing all bilingual students or English-language learners in both English and their dominant language).

The significance of the use of uncertainty and ambiguity as resources in the social construction of knowledgeability is another point in Kelly’s comments that might be worth having more discussion of among all of us – especially if we imagine the implications of the analyses in the original article for valid assessment practices. I am thinking here of the uses by both students and assessors, and the uses even in written tests. For instance, I and many other students with very good communicative skills and strategies often wrote answers, or gave oral answers, that in effect invited our more knowledgeable assessors to “read in” to our answers more than we actually “knew.” We were at the same time making ourselves seem more knowledgeable and both trading on, and simultaneously ratifying, the superior expertise of the addressee/assessor. But there is a lot more to be said about this!

Finally, your last point about institutional power relations vs. locally constructed statuses “in the moment” is another potentially very significant one for issues of assessment, as well as for the validity of research/interview data. (Nespor also discusses this issue.) There is a long tradition in discourse analysis of looking at “known-answer questions,” and classifying types of exchanges in terms of roles taken on or assigned by the discourse moves as to who is the “primary knower” etc. (esp. work by Margaret Berry in the UK, but also by many, many others). So one can think of expertise or knowledgeability as something that is constructed across multiple timescales: there is what is constructed in a particular turn at talk or exchange, what is constructed over the course of a brief interaction, and what is constructed over much longer institutional timescales (credentialing, appointments to official positions, etc.) as well as at the scales in between. But what is
not often enough analyzed is the way in which institutional-scale power relations and locally constructed power-differentials (Foucault is always alluding to this, but never analyzes local interactions) get mixed in with constructions of (and expectations about) knowledgeability.

Again, we all know cases where the lower-power/status participant tries to make the superior appear more knowledgeable, when the former is actually more the primary knower. This is a sort of reversal of the “known-answer question” strategy in classroom dialogue, where teachers try to make it seem as if students are knowers (they are asked, they answer), but where the institutional norms are that the teachers are the primary knowers, and the result of this contradiction is that teacher questions are seen as assessments, not as requests for information. It can also work the other way around, but because of the power asymmetry we get different cultural interpretations of the resulting interaction (e.g. that students are making a fool of the teacher, sucking up to the teacher, “playing” the teacher, etc.).

In terms of assessment, of course, there is also a moral issue here. The usual argument is that power asymmetries that arise inevitably from knowledge asymmetries are not immoral constructions of human relationships, and that teachers are just “helping” students to learn. But in fact the institutional-scale power relationships exist independently of locally constructed knowledgeability, and are at their most extreme in assessment activities. If we take the analysis of the local construction in discourse of relationships of knowledgeability and power out the context of the institutional setting, we miss the morally questionable dimensions of what is happening.

Annamarie and Daniel are an unusual dyad in many ways, in terms of their institutional statuses and the complex relations they have in terms of expectations about knowledgeability-in-context, which is partly why the article can show us so many revealing phenomena in their interaction. There is not a clear, unambiguous long-timescale institutional-level status/role/power relationship to define expectations, norms of behavior, etc. as part of the context. In fact, on the relevant longer timescale, we are missing some interesting information: how did Annemarie come to agree to this interview in the first place?

I am raising these points because I think a key issue for discussion in the Forum is how the article’s findings may apply to assessment and teaching as well as to research methodology. So we have to have a sense of in what relevant ways the interview situation was unlike that of assessment or the classroom, as well as how it is like them. On the short timescale of discourse processes, there are a lot of similarities. On the longer timescale of institutional contextualization, we may find important differences.
Both Lemke and Kelly point out that there are precedents to viewing interviewing as a situated, indeterminate event, brought about by the moment-to-moment contributions of interaction participants. Each contribution constitutes a new interactional resource that in effect changes the world in the way that it is available to others currently present. That is, the interview – as any other social event – does not constitute a box into which actors with predetermined roles are placed, but which is a social event that produces social reality as much as it reproduces it. I am sure that neither respondent wanted to suggest that our article merely reproduced particular findings or methods of research. For us, the main purpose to write this article is the disciplinary context of science education and educational psychology, where interviews continue to be taken as a pathway to the knowledge of individuals, their beliefs, attitudes, life stories, views on science and scientific knowledge, intentions, and so on, irrespective of the purposes of the interview. To make this point, we practice those research methods that allow us to construct the necessary data. That is, we use a conversation analytic approach, because it allows us to do what we needed to do. It also allows us to analyze other interactional resources produced, such as prosody. In this, the conversation analytic approach appears more appropriate than Michael Halliday’s approach in its almost exclusive focus on language.

Any piece may be framed in a number of ways. Ours raises questions concerning how to get at knowledge or rather aptitude, which is the concept of interest to the late Richard E. Snow, who, after decades of conducting research on aptitude-treatment interactions came to realize that aptitude cannot be measured using one or more (written) instruments, taken as statically given, without any effect on the person who responds to the tests, and, in doing so, also changes. Snow therefore suggested that we needed investigations that find out what happens at the interface between the person whose knowledge is to be evaluated and the evaluating instrument, which may be a researcher. It is this interface that was of primary interest to us. Thus, whereas previous research in conversation analysis or discourse analysis was concerned with understanding interactions, our primary interest was to contribute to an understanding of how the act of measurement – here the think-aloud/interview situation – shapes the phenomenon to be measured. In the natural sciences, it is easy to deconvolve the contributions from measurement instrument and the “real” phenomenon if you know the former. In the situation of interest to us, the same approach does not work, as we show, because the event and whatever it produces is of a new order.
That is, even the roles of interviewer/interviewee – who is in the know, who legitimately asks questions, who believes what, and so on – all are outcomes of the event as much as they are resources to the participants for conducting the event.

Lemke’s comments about long-range (temporal, societal) effects are important. This point may have not received sufficient attention in our already very long piece, but has been treated elsewhere (Roth and Lee, 2004). Marxist scholars in particular have been concerned with and made the topic of their research the “determinations” that derive from social-societal entities (language, tools, relations) not currently thematic in the experienced lifeworlds of interaction participants. If I was what Bourdieu (1992, p. 226) calls a “monomaniac of …” research method, this important aspect would have eschewed us, for conversation analysis has no means to get at those aspects of social reality that lie outside what conversation participants make available to one another. This is why the work of the feminist sociologist Dorothy E. Smith (e.g., 1987, 1990) is of some interest, for she articulates the attendant issues very well. Even within ethnomethodology, a field allied with conversation analysis, scholars have realized the shortcomings of analyzing the lifeworld and conversations and neglecting to attend to the long-range effects. Some of these therefore called for a critical ethnomethodology, that is, a marriage between the field and (neo-) Marxist approaches (e.g., Chua, 1977; McHoul, 1994).

The German critical psychologist Klaus Holzkamp (e.g., 1984) also made salient the problems inherent in analyzing the lifeworld alone. Addressing issues salient to the world as we experience it therefore may be like, to use a medical example, treating the symptoms instead of the causes. If the determinations or interactional resources are from outside the consciousness of the interaction participants, then, of course, simple conversation analysis or discourse analysis will not allow us to understand why participants do what they do.

An issue that readers of scientific articles and even many scholars forget is that when we talk about knowledge, interactions, cultural presuppositions, and all the other concepts that Lemke, Kelly, and our own article raised are concepts that – at best – are resources participants use. Many scholars forget that interactions are conducted in real time, that “to know” or “knowing” or “applying knowledge” therefore cannot mean to have homunculus go to the storehouse and retrieve something to be used in the next action. There exist some explicit modeling in the cognitive sciences that show if we had to reflect on the next action by processing information we would always be too late (e.g., Kirsh and Maglio, 1994). Our article is also a contribution to how participants in knowledge assessment interviews and think-aloud protocols conduct their interactions in real time – though
because of our medium of communication, language, this very important point of the constraints of acting in real time often get lost.

Lemke makes an important point about the relationship between scholarly knowledge productions and everyday politics. It is one of the inner contradictions of society that division of labor constitutes a continual process of diversification splitting domains into different domains that subsequently no longer interact or influence one another. Thus, politicians making decisions about education have to know little about knowing and learning in the same way that architects have to have very little practical knowledge of placing and cementing two stones, framing a house using two-by-four and two-by-six building lumber, or laying electrical wiring. Historically, there was a time when there were no architects and all knowledge to build and knowledge about building could be found in one and the same mason. Similarly, there may have been a time – ancient Greece – when politicians and kings were philosophers. Lemke also asks, “Why are we content to sit silently by while our educational future is determined by . . . ?” But sitting silently by we do not have to. There are many individuals in our societies who do not sit silently by concerning education or anything else but become what we often refer to as activists. With respect to education, when Ken Tobin and I go into schools and teach inner-city kids, we are not sitting by but actively get into the midst of events and contribute. True, these are small influences and minor events – but so is each human–human interaction. The interesting thing is that societal life consists of such interactions, all minor and minute, but some of which come to have determinations far beyond their immediate situation without that we could know just why this is the case – Why might the mother of a fallen soldier camping in front of President George W. Bush’s summer residence bring about more change in American politics than all the social critics that have talked out against it over the three years prior to the mother’s protest?

CLOSING COMMENTS 1

Jay Lemke

Roth is quite correct that his and Middleton’s contribution in this article is to bring these analytical perspectives to bear on important questions of interview research methodology and by wider implication, the claims of even our best forms of knowledge assessment.

As Roth mentioned in passing, his reasons for using CA rather than Michael Halliday’s functional linguistics approach (which I often use myself), I thought I should mention that Halliday’s original approach does take prosody and intonation into account in an integral way, but not
many of those who have used his approach in education have done so. It does not, however, include the many aspects of nonverbal communication that are often tightly integrated with verbal talk in face-to-face interaction. In this area I also particularly admire the work of Chuck and Marjorie Goodwin, and Ron and Suzanne Wong Scollon, all cited by Roth/Middleton.

Of course, as Roth notes, he and Middleton were well aware of the limitations of the methods they use, particularly with regard to including the macro-social factors that contextualize face-to-face interaction. What is still, I think, for all of us in the education research community a difficult task is finding fruitful ways to trace the passage back and forth between the productive kinds of micro-analysis which, as Roth/Middleton show, tell us so much about real-time processes, and the broader sociological frameworks that so often define for us our deeper worries about educational practices, valid assessments, etc. I have been exploring a “timescales” - based approach to this classic micro vs. macro dilemma (Lemke, 2000), trying to frame local events within longer-term processes taking place over longer and longer characteristic timescales, in a sort of embedding and contextualizing (but not controlling) hierarchy. This is one reason I suggested the usefulness of looking “outwards” (or up the timescales) toward issues like how Annemarie came to agree to being interviewed in the first place, or how the focus of the interview questions fits into the Roth/Middleton work, or their concern with Snow’s aptitude-treatment interaction work, etc. One virtue of a forum such as this one is that we can hear more about such “behind the scenes” dimensions, for which there is often neither space nor definite expectations in the standard academic genres of most journals.

Finally, I agree of course with Roth that many of us in science education practice the local activism of working collaboratively toward change in schools, and often against the political grain of the moment. What I myself have not done enough over the years, I think, is to speak out in larger public forums about broader policy and theory issues: Why I doubt the ultimate success of the current science education reform movement; why I do not agree that standards-based curricula make for good education; why, like Roth/Middleton, I have serious reservations about both how we do assessment and especially the kinds of claims about individual aptitude or competence that we justify by it. The careful scholarship of Roth/Middleton may persuade us on one such issue, but we then need to go out into other forums to persuade educators, policymakers, and our fellow citizens at large. Why do we so often leave educational policy to people whose ideas, and even motives, we doubt?
Michael Roth is quite correct that my comments regarding the research interview should not be construed to suggest that the study by Roth/Middleton merely reproduced particular findings or a method of research. This is clearly not the case. Rather, my point is to note that while one central premise of the study (communicative nature of the interview) was made available prior to much research in science education using research interviewing, this work has, for the most part, been ignored. While there are probably institutional reasons for this, such as disciplinary boundaries in academia, I believe that the Roth/Middleton paper helps understand the issue differently. One plausible reason for the lack of consideration of the conversational features of interviews in the field of science education is that the co-construction of (say) a counseling event (as in Mishler, 1986) can be interpreted to be the consequence of a social interaction in ways that the interpretation of the earth, forces, molecules, or even a graph, may not be. Science education research (using research interviewing) has focused primarily on student concepts, misconceptions, and views of the nature of science. Such cognitive artifacts have a (seemingly) tangible nature. After all, among the 6,000 studies of student conceptions (Duit, 2002, cited in Roth/Middleton), there is appreciable redundancy regarding common (supposed) misconceptions. While I have not (thankfully) read all 6,000 of these articles, I am familiar with a few dozen focused on students’ conceptions of force. Given common interview protocols, common patterns emerge in student responses. However, as Roth/Middleton point out, “No matter how standardized, an interview remains a linguistic and transactional event (Maynard and Schaeffer, 2000)” (p. 5). Roth/Middleton start with this premise and examine in detail what can be understood about one interview conversation from this point of view. As I noted earlier, the implications are vast; I offered five examples. But, to return to the issue of this study merely reproducing previous work, the reason for my attention to this matter concerns the extent to which the Roth/Middleton article can communicate the seriousness of the conversational features of interviewing. I believe they have shown that the seemingly unproblematic interpretation of a “simple” graph creates ambiguities regarding the putatively tangible physical events referenced by the graphical relationships in ways that make the science of the issue as interactionally accomplished as the issues that arise in the counseling interviews of Mishler. Therefore, this careful work by Roth/Middleton should require serious review of the uses of interviews in science education as a means to examine pre-existing student conceptions.
Roth and Lemke both identify ways that close examination of in-the-moment social interaction can be hitched to projects concerned with improving the human condition. A common critique of conversational analysis has been that detailed analyses of micro-moments of life fail to consider the ways in which sociocultural practices, institutional norms, and other long-range effects enter into such micro-moments and aggregate to form barriers to emancipation – although Roth notes CA and EM responses to such criticism. I agree with Roth that Dorothy Smith offers some ways to consider how people in everyday life exist in social organization. These everyday situations involving the coordinating or concerting of actual people’s activities are appropriate places to engage in social inquiry as Smith (1996) suggests. Roth proposes engagement in local activism as one way to influence our educational future. In addition to direct influences on the social setting, such local engagements may provide contexts for research supporting alternatives to the narrowing of educational aims and means currently prevalent. Alternative forms of “claiming knowledgeability” may emerge from such engagement, although as Lemke notes there is much work to be done, and done in different forums. Lemke suggests that researchers need to enter more often into the public arena to persuade fellow citizens about the merits and detriments of educational policies that do influence the lives of others. I believe this is our greatest challenge, particularly starting from a research base focused on the discourse processes of interaction. Such research does not (presently) generate ideas that thrive in the policy ecology populated by terms such as accountability, incentives, testing, and standardization.

CLOSING COMMENTS 3

Wolff-Michael Roth

Interactions such as those that occur in this Forum are important to move our field ahead. Too often, scholars have to work on their own without the opportunity to discuss the similarities and differences in the way they talk and talk about theory and method. This Forum provides both cultural-historical context to our lead article, but also exhibits differences in the way we, the participants, prefer to talk about issues of interest. Although we covered a lot of ground, there are many finer points that could be discussed, but for which there is not sufficient journal space and which may require face-to-face discussions to develop the attendant issues more rapidly and with fewer possibilities of misunderstanding one another. One of these issues is that of “meaning,” what exactly its nature is. For Lemke, it appears to arise with text. For me, taking a more Heideggerian stance, it already
predates the encounter and any new signs, words, concepts, etc. accrue to existing meaning. Meaning therefore is something grounded in our practical and tacit understanding of how the world works, which we can articulate and elaborate when the local circumstances require it. Meaning is mediated by text, but cannot be reduced to text – here, Vygotsky (1986) has already made important but often overlooked contributions by subordinating speech and gesture to the same meaning units. Language and other semiotic resources that interaction participants produce serve to make this event similar to other events known to and previously experienced by the participants or to make this event similar to expectations about how such events ought to look and feel like.

There are methodological questions as well, for example, what we may assume and what we may not assume in the analysis of a stretch of talk. Thus, it is not important what Daniel “presumably” knows, whether it is more or less about the right answers; what is important to the interaction is how the participants use whatever the resources they have at hand to establish who is in the know. I explicitly did not draw on other knowledge to explain the events – as Daniel’s employer, I know how he talked about graphs and about how his talk was at times inappropriate. I know how he has acted and talked in situations other than the interviews. But in the analysis, I have not let enter any of this knowledge: What I allow myself to analyze are the same resources that he made available to Annemarie. This way, I cannot use any assumptions about hidden intentions, about psychological traits (e.g., to appear knowing more than he actually does, to be pretentious), and about his mathematical knowledge more generally. To understand the interview as an interactional phenomenon, all this other information is irrelevant, as the interaction participants themselves provide all we need to know about the interaction.

Another issue is the question of the “long-range” effects, that is, how structures outside the immediate situation, produced and reproduced by the participants, have determinate influence on the events. One important source for the theoretical and methodological problem appears to lie in our culture and the language we use. Thus, in English, we refer to interviews as social or socially mediated events. Thus, when dialectical and critical psychologists and philosophers (Karl Marx, Lev S. Vygotsky, Alexei N. Leont’ev, Klaus Holzkamp) use an adjective “societal” to make salient the fact that interactions are mediated by society and its division of labor, this gets translated into English as “social.” But the result of using “social” versus “societal” is not the same. The former adjective (adverb) may be used to describe the mediation that occurs whenever two or more people interact; the latter adjective (adverb) explicitly brings aspects of society – class, inequality, subcultural differences, gender, and race – into the
analysis. Thus, as soon as I think in terms of societal mediation, I need to conduct analyses that take a cultural-historical perspective on the resources interaction participants use, and how these structures (e.g., Smith (1990) provides the gendering that is inherent in the usage of such terms as “mental illness” and “suicide”) mediate our relations with others.

My ultimate goal is to understand how we, human beings, make the world work, how we produce and reproduce the social events we are familiar with in real time, where there is no time out and where there is no recourse to knowing more than what is or is made available in the situation at hand. My goal is to understand real classrooms, and how, for example, a teacher comes out of one eleventh-grade physics class with a sense of being a “good teacher” but exists from another with a sense of failure. The article we wrote is one step towards this goal, not the ultimate answer. In forums such as this, problems with this as with any approach and contribution to collective knowledge have to be worked out to take us further down a road of better understanding about how the world works. I also hope that our understanding about eventually becomes part of our practical understanding of the world – and in my own life it does to the extent that my social scientific knowledge has changed how I live, experience, and interact with others.

Jay Lemke is Professor of Educational Studies at the University of Michigan with affiliations in Science Education, Learning Technologies, and Literacy Language & Culture. He is the author of Textual Politics: Discourse and Social Theory (Taylor and Francis, 1995), Talking Science: Language, Learning, and Values (Ablex, 1990) and Using Language in the Classroom (Oxford, 1989). He is also co-editor of Critical Discourse Studies and formerly co-editor for ten years of Linguistics and Education. He earned his Ph.D. in theoretical physics from the University of Chicago, and his current research is focused on analyzing the ways people make meaning across multiple media, attentional spaces, and timescales, most recently for interactive-immersive multimedia such as computer/video games.

Gregory Kelly is a professor of science education at Penn State University. He is a former Peace Corps Volunteer and physics teacher. He received his Ph.D. from Cornell in 1994. His research focuses on classroom discourse and science learning. This work has been supported by grants from Spencer Foundation, National Science Foundation, and the National Academy of Education. Greg teaches courses concerning the uses of history, philosophy, sociology of science in science teaching and teaching and learning science in secondary schools. He serves as the section co-editor for Learning in the journal Science Education.

Wolff-Michael Roth is Lansdowne Professor of Applied Cognitive Science at the University of Victoria. After teaching middle and high school science, computer science, and mathematics for over a decade, he began his university career as a statistics professor prior to taking his current position. He now conducts research into knowing and learning mathematics and science from kindergarten to professional practice, which is published in linguistics, sociology, and education journals. His recent publications include Toward an Anthropology of Graphing (Kluwer Academic Publishers, 2003), Talking Science: Language and Learning in Science Classrooms (Rowman and Littlefield, 2005), and, with A. C. Barton, Rethinking Scientific Literacy (Routledge, 2004).
REFERENCES


---

JAY LEMKE

*University of Michigan*

GREG KELLY

*Pennsylvania State University*

WOLFF-MICHAEL ROTH

*University of Victoria*