

Running Head: PROGRESSIVE REFINEMENT OF HYPOTHESES

Progressive Refinement of Hypotheses in Video-Supported Research

Randi A. Engle

University of California, Berkeley

Faith R. Conant

Five Colleges, Inc.

James G. Greeno

University of Pittsburgh

To appear in R. Goldman, R. Pea, B. Barron & S. J. Derry (Eds.), *Video research in the learning sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Scientific inquiry in the learning sciences often includes hypothesis generation and evaluation together, in a kind of dialectic process. We believe that understanding such inquiry can be enhanced by focusing on the interaction between hypothesis generation and evaluation, rather than separating them as some accounts of science have in the past (e.g., Popper's, 1959 distinction between the "context of discovery" and the "context of justification"). Therefore, much research in the learning sciences involves not just classic hypothesis testing in which scientists specify hypotheses in advance of collecting data and design experiments to provide definitive refutation or tentative confirmation of those hypotheses. At the same time, what many of us are doing is not characterized by a naïve view of ethnographic inquiry in which investigators attempt to begin with a theoretical clean slate, developing hypotheses only in response to patterns in the data. Instead, the kind of research many of us are doing is characterized by the progressive refinement of hypotheses, with theory and data interacting throughout the process.

Specifically, in the process of *progressive refinement of hypotheses*, an investigator begins with a general question and then decides to collect empirical records in a relevant setting with an initial plan for how to use them to learn more. Initial analysis of these records informs more specific hypotheses that then may be addressed in other aspects of the data. Those results then provide even more specific questions that are addressed in some other part of the data, leading to

explanatory hypotheses that can then be evaluated in still other aspects of the data, and modified or made more specific in relation to those data, and so on. The development and evaluation of hypotheses involves a kind of constraint satisfaction, in which the coherence between hypotheses and the consistency of hypotheses with data are both taken into account.

We believe that much progress that is made in scientific research actually follows some version of this pattern. Commonly the pattern plays out over multiple studies, with data and theoretical proposals provided by multiple investigators. However, with the advent of records like video that both are relatively comprehensive and support multiple re-viewings, it is now often possible to take several steps in the progression using the same set of records. Thus, video records can allow a single study to progress through multiple iterations of hypothesis generation and evaluation, making the resulting findings more robust than they might have been otherwise. By more robust, we mean that the findings have been evaluated and modified in relation to more aspects of the data and therefore are more likely to stand up to further investigation.

In this paper, we discuss a research project that illustrates the uses of video records to support the progressive refinement of hypotheses about student engagement and learning. We first discuss how our overall project was structured to support this process, with a focus on how video provided useful affordances for it. In the core of the paper we then discuss a specific study we did (Engle &

Conant, 2002) to show how the process of progressive refinement can work over the course of a video-based study. As part of this, we show how video-based case studies like ours can be used both alone and in concert with later studies to support hypotheses of wider generality. We believe that this study sheds some light on ways that development and evaluation of hypotheses can interact productively, especially in research that includes extensive video records as a source of data.

Support for progressive refinement of hypotheses in the design of a research project

When we (Greeno and Engle) proposed a research project about the role of discourse in conceptual learning, we had not formulated any hypotheses that could be falsified by the data we proposed to obtain. However, we did much more than simply identify a potentially interesting setting for study (Fostering Communities of Learners classrooms) in the hope that worthwhile findings would somehow emerge. We had formulated the issue in theoretical terms, and we had also characterized how we planned to use video records to develop more specific hypotheses and then evaluate them.

Specifically, the goal of the research that we initiated was to advance understanding of the processes of conceptual growth in project-based classrooms. Our theoretical goal was to develop and evaluate hypotheses about how students'

participation in classroom activities could account for the advances in their conceptual understanding. We proposed to combine methods of cognitive developmental psychology to assess students' conceptual understanding before and after a unit of study, with methods of discourse analysis and cognitive representations of information to analyze students' and teachers' discussions during the unit.

Video records of classroom interactions were a crucial component of our research plan. We made video records of groups of students working on assessment tasks before and after the unit, and these were the materials we analyzed to characterize the conceptual growth that we planned to document and explain. We also needed to observe classroom activity across several school weeks in order to be able to examine significant changes in students' conceptual understanding. Thus we also recorded video during the approximately 12 weeks during which two classes worked on the Endangered Species unit of the Fostering Communities of Learners curriculum (Brown & Campione, 1994; Kohl & Wingate, 1995). In particular, we followed two target groups of students in each class, videotaping both their interactions and whole-class activities.

It was through the examination and analysis of these videos that we hoped to develop and evaluate hypotheses that we could propose as explanations of conceptual growth. Our plan, which we followed, was to use the pre/post assessments to identify topics on which students' conceptual understanding had

advanced, then to examine episodes during the unit in which those topics were discussed, and analyze ways in which discourse in those episodes provided occasions in which students could have advanced their understanding in ways consistent with the changes we had observed. Our expectation was that different groups (and students within these groups) would demonstrate contrasting patterns of change on the assessments that would provide us with contrast cases to constrain the explanations we could generate for how the students' conceptual growth could have occurred.

Thus, although we did not have empirically testable hypotheses at the start, we were committed to developing and evaluating more definite hypotheses in the course of our study. We anticipated findings about the nature of conceptual growth from the assessment part of our study, and we proposed to obtain evidence in the classroom discourse part of the study that we would use to explain the findings of the assessment part. Our plan was plausible, we believed, because it was continuous with methods of successful programs of research in the general domain of our study (e.g., Rosebery, Warren & Conant, 1992; Yackel, Cobb & Wood, 1991), and we had participated in projects that used many of the methods that we needed to carry out in our proposed research (e.g., Engle & Greeno, 1994; Greeno & Engle, 1995; Greeno, Engle, Kerr & Moore, 1993; Linde, Greeno, Roschelle, Brereton, Lewis & Stevens, 1994).

Our plan was consistent with a research strategy of progressive refinement of hypotheses. We needed to conduct and record the pre/post assessments in order to identify aspects of the conceptual domain on which students advanced in their understanding. Having found some evidence for conceptual growth, we needed to have records of teaching and learning to examine in order to develop and evaluate hypotheses about how that conceptual learning occurred.

Consider what we could have done without records of this kind. If we had not made recordings—either audio or video—we would have had to depend on field notes or observational coding sheets that we would have been forced to develop before observing the units or identifying the topics on which students changed their understanding. As it turned out, that would have vitiated our study almost entirely. Although we correctly anticipated the general topics on which the students were likely to make progress (e.g., a better understanding of habitats and adaptation), we had no basis for anticipating the specific contents of their conceptual advances. For example, one of the features of many of the students' conceptual growth was their consideration of birthrates as a quantitative variable that is influenced by multiple causes and has multiple effects (Engle, 2005b). This is not surprising in retrospect, but we had no basis beforehand for giving this topic sufficient priority to specifically code for it in a classroom observation. Therefore, instead we decided to collect an extensive record of classroom activities on video, making it likely that, once we did identify exactly where particular students'

conceptual growth occurred, we would have sufficient material about what happened during the unit to develop justifiable accounts of its development.

More importantly, an observational coding sheet would have presupposed that we could anticipate the specific kinds of interactions that would figure in our explanatory hypotheses. Similarly, for field notes to have been successful, we would have had to recognize potentially significant interactions as they were occurring and make sure to record the most relevant details. We simply were not in a position to anticipate the kinds of events that we have decided, in retrospect, help account for the kinds of conceptual understanding and learning that appear to have occurred. For example, in retrospect we have concluded that teachers attributing authorship of ideas and information to students was probably a critical factor in their learning (Engle, 2005b; Engle & Conant, 2002). Such attribution of authorship also might have been expected, but it is unlikely that we would have given that characteristic, above many other possibilities, a high priority in any coding scheme we could have developed in advance. Even if we had, we cannot imagine how we could have constructed a sufficiently accurate record of the events we now identify as positioning students as authors without having had a fairly complete record of the interactions to refer to. With video, however, all that was possible.

In addition, the video records turned out to enable an analysis of an issue that we could not have anticipated, as it involved analysis of a series of events that

were not a part of the planned curriculum. Furthermore, our attention was drawn to the issue and the events by unanticipated opportunities for collaboration.

Following the collection of the video records, as planned Greeno, Engle, and their colleagues began to analyze records from the assessments and some of the records of interaction from the unit in order to develop accounts of conceptual growth (Benke, 1999; Engle, 2004a; Greeno, Benke, Engle, Lachapelle & Wiebe, 1998; Greeno, Brown & Campione, 1999; Lachapelle, 1997; Wiebe, 1999). Faith Conant and Frederick Erickson were each visiting for the year and interested in working with us. Erickson said he was particularly interested in exploring the question of “what makes a conversation take off?” During our collection of the FCL videos, we had noticed that the target group who was studying whales had become unexpectedly and passionately engaged in an argument about whether orcas (“killer whales”) are whales or dolphins (what we came to call the “orca controversy”). This occurred despite the fact that the issue of how species should be classified was not part of these students’ assigned task, which was simply to explain why whales were endangered.

We had not planned to focus on this series of events as its topic did not correspond with any items in the pre/post assessments, but Erickson’s interest made us realize its potential value for our understanding of conceptual learning. These incidents underscored that conceptual understanding is fundamentally affective and historical as well as informational. For example, when reporting

about their work to a student teacher just joining the class, one student's brief mention of the orca issue caused the argument to re-erupt with great emotional intensity, with this student remarking that the group had had a "big ol' argument" about the topic earlier. This implied to us that our account of the students' conceptual understanding must include such historical and affective aspects along with informational content. We had captured the students' activity on our video records, but we were, in turn, captured by those video records, especially by the intellectual and affective intensity of the students' discourse. Thus, this series of arguments became a major focus of analysis.

In the rest of this chapter, we will highlight some of the opportunities that our video records afforded us as we worked with these data to explore Erickson's question of what make a conversation take off, drawing primarily on Engle and Conant (2002), the first paper our orca controversy analysis group has written from our work together.¹ First, we explain how we interacted with the videotapes to capture the phenomenon of a conversation taking off. In particular, in this section, we discuss how having this focusing question shaped our initial selection of video records to analyze, how we used transcripts to represent our interpretations of these records, and how we helped check and enrich our initial hypotheses by both sharing the videos with other investigators and performing formal codings of them. Following this, we then discuss how we worked with our video records to develop an explanation for why the conversation took off. In this

section, we consider how we approached this process with an eye towards generalization; identified possible explanatory factors; restructured our explanation into a simplified general model; drew upon the full set of videotapes to refine and specify it; and finally investigated whether the theoretical concepts in the model could be used to help account for other cases in the literature.

Progressive refinement while capturing and communicating the phenomenon

Selecting video to analyze. As is typical with many classroom-based projects that use video, we had collected a whole bookcase worth of videotapes that we could have analyzed. How did we know what to actually analyze? Our research question drove our selections. We remembered the conversation had first “taken off” the day after the students had returned from a trip to Marine World as they were working on their bulletin board, and that the oral report in which the debate had re-erupted had occurred seven weeks later. We began our analyses with these two incidents and then searched before and after them to find other occasions in which the students discussed the orca issue.

Searching for episodes of discussion of this topic was feasible because we had made content logs of the videotapes in our collection (Jordan & Henderson, 1995). A content log is written by someone watching a tape with only minimal rewinding in order to provide a time-indexed list of topics being discussed. These content logs were used to select segments of video to watch and transcribe.² In

addition, we paid careful attention to references that the students themselves made to potentially relevant past events or anticipated future ones. Using these resources, we found out that the orca issue had been discussed by the students on at least eight occasions over the last eight weeks of the unit. We found that although there were differences in the degree to which particular students continued to be engaged in the orca issue over time, it was clear that it was not a one-time issue for this group, but instead the focus of sustained examination by them.

Representing evidence for interpretations. Compared with fieldnotes and observational coding sheets, video records are a relatively under-interpreted source of information.³ Because of this, working with video invariably involves a second phase in which researchers grapple with how they are going to interpret the events that have been recorded. In addition, since video materials contain so much information, part of this process involves deciding which of the many things that can be seen and heard are relevant to making interpretations about the phenomenon of interest.

In our case, we were interested in understanding how one can tell that an academic conversation has taken off, which we eventually described by developing the concept of “productive disciplinary engagement.” Briefly, productive disciplinary engagement is occurring when: students are deeply engaged in an issue (“engagement”), their engagement with it makes significant

contact with disciplinary ideas and practices (“disciplinary”), and through this engagement, they make progress on the issue in question (“productive,” see Engle & Conant, 2002, pp. 402-403 for more complete definitions). It seemed evident to us from the tapes that the students were, at the very least, intensely engaged in the orca issue, but we needed to become more explicit about the basis for this and other interpretations to be able to examine them more systematically and share them with others. We did this in a multi-step process that involved progressive refinement of our representations of the events and our claims about the nature of the students’ engagement in them.

From viewing the videos we had noticed, for example, that the students often fought for the floor, emphasized their points in a dramatic fashion, built upon each others’ ideas, and celebrated when they had successfully argued a point. However, many of the specific actions we were considering as evidence of the students’ engagement—e.g., overlapping speech, emotional displays, and emphatic stress—had not been systematically included in the rough transcripts we had. If we were to describe the intensity of engagement that we were seeing to others, we would have to record its relevant features. More detailed transcripts were constructed to include these and other features. For example, a section of transcript from the “Big Ol’ Argument” that originally had been transcribed as:

Liana: AND she probably didn't even know, because she was just there for a little while,
and then she just made it up.

Brian: She was only there for two years!

Liana: yeah, so she didn't know that much, so she just thought of it then, at that time 'cause she probably was thinking about DOLPHINS!

Brian: Exactly! (*Brian and Toscan high-five*)

was revised to:

409 Liana: (*waving right hand wildly as Brian nods head and shoulders emphatically*)
 410 OH! . . . oh . . . oh . . .
 411 A:::ND . . . she probably didn't even know .
 412 because she was probably just there for a little whi::le .
 413 and then she just made it [up
 414 Brian: [SHE WAS ONLY THE:RE FOR TWO
 415 YEA::[:RS
 416 Liana [YEAH! . . . so she didn't know that much she/ she just thought of
 417 it that . . . that time cuz she probably was thinking about . . . DO::LPHINS
 418 Brian: (*big grin*) . . . exa::ctly (*Brian "high five's Toscan*)
 419 Liana: yup
 420 Racquel: (*pats Liana on the back and they exchange nodding looks*)

to represent the degree of emphasis, overlapping speech, emotional displays, and non-speaker involvement that were characteristic of this discussion.⁴ In effect, our decisions about what categories of things to record in our transcription helped us specify some characteristic features that might provide evidence of student engagement. Thus, our transcripts became a record of our emerging hypotheses, as embodied in the specifics of a particular interaction (see also Ochs, 1979).

Refining hypotheses by making comparisons, enlisting other analysts, and coding. Based on sufficiently detailed transcripts, we were able to be much more explicit about what we were noticing in the orca discussions that led us to the conclusion that the students were strongly engaged in them. In order to provide

evidence that these features were significant, however, we needed some other discussions to which we could compare the orca discussions. To provide this comparative data, we re-examined the video of the students' report to the new student teacher, which also contained presentations about the rest of the students' work during the unit. In this session, we could compare the students' engagement in these other topics with their engagement in the orca issue in order to verify our impression that they were more engaged in the orca issues than their other research topics. As mentioned earlier, there appeared to be a dramatic shift in the students' engagement the moment that one of the students (Brian) brought up the controversy as part of his report on whale features.⁵ We slowed down the tape, sometimes watching it with the sound off, and noticed both how Brian signaled that the topic was coming, and the rapidity with which the other students responded to his signals with evidence of their engagement. For example, as soon as Brian mentioned the orca issue, the whole group squeezed closer together, other members of the group began shaking their heads in agreement or disagreement, and various students began to bid for the floor. (These phenomena, by the way, would have been impossible to observe had we only audiotaped this session.)

Our impression of a shift was then corroborated by a group of colleagues whom we asked to watch the tape and note where they observed a change (if any) in the students' level of engagement. Their proposed locations were close to

where we had identified the shift as occurring, giving us additional confidence that it was there. On this basis, Engle and Conant (2002, pp. 417-419) wrote an analytic description of the shift, supported by a revised transcript of Brian's full report (*ibid.*, pp. 465-473).

To provide evidence that could be abstracted from the video record for readers without access to it, and to illuminate features of interaction that might provide evidence for changes in student engagement, we then coded and statistically compared the 70 seconds of Brian's report before he raised the orca issue and the next 70 seconds afterwards. In particular, we showed that when discussing the orca issue as compared to the rest of Brian's report: (a) speakers were much more equally distributed between the students; (b) students were more likely to address each other with their contributions rather than only the new student teacher; (c) students engaged in activities unrelated to the discussion much less often; (d) there were many more emotional displays by the students; and (e) there were also many more task-relevant spoken overlaps as students fought for the floor or collaboratively completed each others' ideas (see Engle & Conant, 2002, p. 419 for results, and pp. 473-476 for methodological details). This comparative coding provided measurable, behavioral indicators of the students' greater engagement in the orca question than in the rest of their work during the unit. At the same time, it suggested some features that might be useful for assessing engagement in other contexts.

We then participated in similar processes of progressive refinement in order to establish how the students' engagement in the orca issue was also both "disciplinary" and "productive" (see Engle & Conant, 2002, pp. 420-424 and 424-429, respectively).

Progressive refinement while developing a generalizable explanation for the phenomenon

While we were using the videotapes to clarify how the orca discussions were a case of productive disciplinary engagement, we were simultaneously exploring an additional set of related research questions at another level of analysis. First, we wanted to know why this group of fifth graders had become so engaged in this particular, fairly technical question about the classification of orcas. Given it was not part of the students' assigned task, how did it become relevant to them and why did they continue debating it over the course of the unit? At the same time, we wished to see whether there was anything about the design of the FCL learning environment itself that might account for the students' engagement. If we were successful at doing that, this might help us better understand how FCL classrooms work. Last but not least, we were interested in theorizing much more generally about what needs to be in place in a learning environment if it is likely to support such productive disciplinary engagement. If we could specify aspects of the FCL learning environment that could help account for the students'

engagement at a sufficiently broad level of generality, then our research would provide hypotheses about how productive disciplinary engagement could be supported more generally. So we began trying to account for the whale students' engagement in the orca question with an eye to both explaining the specific events that were involved in their discussions while also creating hypotheses that could be adapted for use in other contexts.

Identifying possible explanatory factors through watching and re-watching.

The initial phase of trying to explain the students' engagement involved a lot of collective watching of the videotapes by our group of five while we brainstormed about possible explanations for the students' productive disciplinary engagement.

In much of the video-analysis work that we have done, a phase like this that involves some kind of collective exploration of the videos has proven to be essential. Everyone gets to know the data very well. From watching the same videos on different occasions, multiple interpretations and hypotheses can be specified and then evaluated for agreement both with the data and with each other. If one short-circuits this phase, one can easily make premature, unwarranted conclusions. In our case, we added ideas to a whiteboard as we thought of them, eventually identifying as many as a dozen or so possible factors that might have contributed to the students' productive disciplinary engagement in the orca question.

Enhancing communication and generality through a simplified model. There were many possible factors that might have contributed to what had happened during the orca controversy. How could we possibly describe—let alone explain—all of them in a talk or article? This problem is endemic to video analysis. Because of the richness of video data, the problem is not usually of having something to say, but of choosing among the many things that one could say and fashioning them into a coherent account.

To address this problem, we took the dozen or so contributing factors we had identified and organized them into four categories:

- *Problematizing*: encouraging the emergence of open intellectual questions that are considered open and of importance to learners while incorporating issues that learners are supposed to be learning about;
- *Authority*: “authorizing” learners to address the questions in their own way (*agency*), publicly giving them “authorship” over their responses (as *stakeholders*), and encouraging them to grow into local “authorities” (*local experts*) whose ideas might begin to shape those of others (*contributors*);
- *Accountability*: encouraging students to “account for” how their contributions are responsive to shared norms for quality contributions as well as to the relevant contributions of others in the local learning environment and beyond;⁶ and

- *Resources*: providing or helping learners to find the various resources they need to directly support their productive disciplinary engagement (e.g., sufficient time or instruction in key technical skills to be able to engage productively) as well as to realize the other three principles (e.g., training in how to construct evidence-based arguments in a particular discipline in order to foster accountability to the norm of giving evidence).

Originally, the purpose of these categories was to provide an uncontroversial package that could embed the specific factors we were using to explain the orca case.

However, as we reviewed relevant literature, we noticed that these four categories that had emerged in our analysis were often discussed separately and in somewhat different terms by different researchers. At the same time, we came to recognize that these four aspects were often realized together in the designs of many innovative and effective learning environments. Thus, we realized that together these four categories were, in effect, a model for what might be involved in supporting productive disciplinary engagement more generally (see especially Engle & Conant, 2002, pp. 408-410). They had the potential to characterize one possible consensus set of design principles implicitly shared by current designers of many successful learning environments. Therefore, by proposing the four categories as guiding principles that might account for productive disciplinary engagement more generally, we also were taking advantage of an opportunity to

identify some common threads underlying many—in some cases quite diverse—efforts.

Refining and specifying the model with the larger dataset. Now that we had a model consisting of principles whose embodiment might foster productive disciplinary engagement, how could we use our videotapes to specify how this model might work? Here, because we had collected a comprehensive record of the unit, we could draw on videotapes not just of the orca controversy discussions themselves, but also of most everything else this group of students had participated in during the unit, including discussions that the teacher held with the class as a whole that helped reinforce the norms of the classroom. With our content logs and transcripts, we could do a systematic search before the controversy erupted to explore how each principle was embodied in the classroom learning environment more generally. We could then combine this with analyses of how the principles were embodied when the controversy erupted, during the critical Big Ol' Argument discussion, and throughout the rest of the unit (see summary tables in Engle & Conant, 2002, pp. 449 and 452 as well as full discussion on pp. 430-447).

We found our explanation compelling to the extent that we could identify linkages between specific aspects of the students' productive disciplinary engagement and the specific ways that each principle had been embodied in this particular learning environment with these particular students around this

particular topic (see especially Engle & Conant, pp. 448-451). Our theorizing was significantly constrained by the fact that our observations during data collection suggested that this particular issue was the one that had prompted the strongest level of productive disciplinary engagement during the unit among any of the four target groups. So we could not fashion an explanation that only referred to global factors true about the learning environment in general—instead we needed to combine analysis of such global factors with those that applied specifically to this group’s discussions about the orca issue. By coding how the students positioned themselves and others, we were able to demonstrate that it was not just the teacher and learning environment that had positioned the students as accountable stakeholders in the debate, but that the students had also positioned themselves in this manner (see *ibid.*, pp. 447-448 for results and pp. 480-483 for methodology).

Generalization by accounting for additional cases. As mentioned earlier, we sought to identify explanatory factors that could account for productive disciplinary engagement in general. Some critics might maintain that successfully achieving such generality is threatened in a research project such as ours either because of the specificity of our videotaped materials, or because we had not specified falsifiable hypotheses in advance of data collection and analysis. The alternative that we consider more realistic and productive involves using such records to inform the development of concepts and principles capable of providing coherent explanatory accounts of findings in a wide range of activities

and settings. From the start, we viewed the orca discussions as a “case” of something more general (see Shulman, 1992), productive disciplinary engagement. We then sought to explain this case of productive disciplinary engagement by proposing four general principles for fostering such engagement. We combined a theoretical argument for why these principles might help account for productive disciplinary engagement in general with detailed empirical analyses of how the principles were embodied in this particular case to help explain specific aspects of it.

We then evaluated the generality of the principles by seeing whether they could be used to help account for several other published cases of productive disciplinary engagement (see Engle & Conant, pp. 451-459). In the paper, we chose to report two of these contrasting cases from the math and science reform literature: (a) the Japanese Hypothesis-Experiment-Instruction approach in which students become engaged in productive discussions around math and science concepts by debating what the correct answer is to a carefully designed multiple choice question (e.g., Hatano & Inagaki, 1991; Wertsch & Toma, 1995), and (b) the Water Taste Test investigation in which a class of Haitian 7th and 8th graders became deeply engaged in scientific investigations of the taste and safety of their school’s water supply (Rosebery, Warren & Conant, 1992).⁷

Since then we have successfully applied the principles to understand cases that contrast even further from the original three in their institutional settings,

topics of instruction, ages of the students, and degree of success in achieving productive disciplinary engagement. For example, Conant and her colleague Marilyn Webster found the principles useful for understanding preschool students' productive engagement in literacy practices in a classroom influenced by the Reggio Emilia approach (Webster & Conant, 2003). Engle and her colleagues have also found the principles helpful for explaining why a set of practices for orchestrating discussions around high-level mathematical tasks are likely to be effective at promoting productive engagement by students in that discipline as well (Stein, Engle, Hughes & Smith, 2005).

Our model has been deepened still further by applying the principles to understand cases of both strong productive disciplinary engagement and weak productive disciplinary engagement by beginning teachers in mathematics education and educational psychology courses (Engle, 2004b; Engle & Faux, 2004). In these studies, we have found evidence of productive disciplinary engagement falling short when one or more of the principles are not fully embodied. From this, we have developed hypotheses about what it might take to realize the principles in new learning environments. In particular, we now hypothesize that when learners are already enculturated into traditional schooling practices, it is wise to first embody the principle of student authority before attempting to hold students accountable to the discipline (Engle, 2004b; Engle & Faux, 2004). We observed that when attempts were made to try to embody

accountability to the discipline before student authority had been fully established, students seemed to interpret the instructor's efforts to hold them accountable to relevant disciplines as a signal that their own authority did not matter (Engle & Faux, 2004; cf. Hamm & Perry, 2002).

Another use of these ideas was in a study of how students' competence as mathematics learners was constructed in several episodes observed in middle-school mathematics classes (Gresalfi, Martin, Hand & Greeno, 2005). These students were positioned as having different kinds of capabilities, ranging from being able to simply execute problem-solving methods they had been shown, to being able to generate examples of a mathematical idea, to being able to invent symbols to represent properties of an event. Gresalfi et al. (2005) found that one important factor in the construction of such student competence was the nature of students' interactionally-constituted accountabilities: to *whom* and for *what* were particular students positioned as being accountable?

The model was used and extended in another analysis of episodes recorded in a 7th grade biology class (Greeno, 2003). The class was part of a large project by Richard Lehrer and Leona Schauble in which students studied the growth of Fast Plants.⁸ In the episode analyzed by Greeno (and other participants in a workshop on comparative methodologies in interaction research), students constructed representations of a collection of data from measurements of the heights of plants on a single day. The analysis showed a need to differentiate our concepts of

authority and accountability. Students were positioned with authority and accountability for constructing a variety of representations, which led to them generating alternative representations that helped problematize issues of representational practice. In the episodes that Greeno analyzed the students were not positioned with the authority and accountability to reconcile these issues, however. Instead, issues were reconciled through the direction of adults. Greeno concluded that it may be productive to consider reconciling, along with problematizing, as aspects of practice to be included in the model, and to consider the authority and accountability with which students are positioned separately regarding these two aspects of practice.

In addition, we are currently extending these ideas to consider how problematizing, authority, accountability, and resources also might be useful for explaining long-term productive collaborations that are not necessarily focused on discipline specific issues. In particular, Engle and colleagues are seeing ways that these principles might explain the multi-year and generally very productive collaborations that researchers and curriculum developers had with teachers and math-using professionals in designing innovative mathematics curricula as part of the Middle-school Mathematics through Applications Project (Engle, 2005; Engle & Goldman, submitted; Greeno et al., 1999). Using both interviews with participants and videotapes of project meetings, Engle and her colleagues are finding that the project systematically supported all four principles. Here we

briefly outline results about how the project supported teachers' authority. First, the project gave teachers the *agency* to focus their efforts on project activities most valuable for them, many of which were designing curricula or other materials for other teachers. Second, teachers perceived themselves as true *contributors* to the project, with almost all noting that teachers' feedback was incorporated into later versions of the unit. The projects' support for teachers to be active agents and contributors to the joint enterprise were viewed as key markers of professional respect that encouraged these teachers to continue actively participating in a highly substantive and productive manner.

Last but not least, these ideas are beginning to be taken up by other investigators who have generatively used the principles in their own research. Some of these studies have provided additional cases of productive disciplinary engagement that can be understood using the principles (e.g., Cornelius & Herrenkohl, 2004; Leinhardt & Steele, 2005). Others have developed the principles in new directions, like investigating how scaffolding can support problematizing (Reiser, 2004) and how effective interdisciplinary collaborators negotiate their disciplinary accountabilities (Nikitina, 2005).

Thus, the test for generality that we applied and that will continue to be applied to our principles and similar ones deriving from this kind of video-based research is whether they support coherent accounts in a wide range of relevant activities and settings, what is sometimes referred to as the criterion of

“fruitfulness” or “fertility” in philosophy of science circles (e.g., Kuhn, 1977; McMullin, 1976). If they do, they will become established in the field’s explanatory discourse. If they do not, they will cease to be attended to. Granted, this is not as clear-cut as the method of falsifying hypotheses. However, that apparent clarity is probably illusory, as analyses of actual scientific practice have shown (e.g., Lakatos, 1970).

References

- Benke, G. (1999). The relevance of discourse: A relevance-theoretic approach to learning as discourse practice. In J. G. Greeno, A. L. Brown, A. L. & J. C. Campione, (Chairs). *Cognitive and social-interactional aspects of conceptual growth*. Symposium presented at the annual meeting of the American Educational Research Association, Montreal, Quebec.
- Cornelius, L. L. & Herrenkohl, L.P. (2004). Power in the classroom: How the classroom environment shapes students' relationships with each other and with concepts. *Cognition and Instruction*, 22(4), 467-498.
- Engle, R. A. (2004, April). Revisiting previous discussions to deepen teachers' engagement with mathematics and pedagogy. In M. S. Smith (Chair), *Developing a knowledge base for teaching: Learning content and pedagogy in a course on patterns and functions*. Symposium at the annual meeting of the American Educational Research Association, San Diego, CA.
- Engle, R. A. (2005a). *Engaging a diverse group of stakeholders in innovative curriculum design: The case of the Middle-school Mathematics through Applications Project (1990-2002)*. Stanford, CA: The MacArthur Network for Teaching and Learning.
- Engle, R. A. (2005b). *Framing interactions to foster generative learning: A situative account of transfer in a community of learners classroom*.
Manuscript under revision.

- Engle, R. A. & Conant, F. C. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. *Cognition and Instruction*, 20(4), 399-483.
- Engle, R. A. & Faux, R. B. (2005). *Towards productive disciplinary engagement of beginning teachers in educational psychology: Comparing two methods of case-based instruction*. Manuscript submitted for publication.
- Engle, R. A. & Goldman, S. R. (submitted). Fostering collaborations over time: The role of teacher authority in the Middle-school Mathematics through Applications Project. In E. S. Charles & R. A. Engle (organizers), *Characterizing collaborative learning environments: Incorporating multiple time scales and participant perspectives*. American Educational Research Association conference, San Francisco, CA.
- Engle, R. A. & Greeno, J. G. (1994). Managing disagreement in intellectual conversations: Coordinating social and conceptual concerns in the collaborative construction of mathematical explanations. In A. Ram & K. Eiselt (Eds.), *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 266-271), Hillsdale, NJ: Erlbaum.
- Greeno, J. G. (2003). *A situative perspective on cognition and learning in interaction*. Paper presented at the Theorizing Learning Practices conference, University of Illinois at Urbana-Champaign.

- Greeno, J. G., Benke, G., Engle, R. A., Lachapelle, C., & Wiebe, M. (1998). Considering conceptual growth as change in discourse practices. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 442-447). Mahwah, NJ: Erlbaum.
- Greeno, J. G., Brown, A. L. & Campione, J. C. (1999). *Cognitive and social-interactive aspects of conceptual growth*. Symposium presented at the annual meeting of the American Educational Research Association, Montreal, Quebec.
- Greeno, J. G. & Engle, R. A. (1995). Combining analyses of cognitive processes, meanings, and social participation: Understanding symbolic representations. In J. D. Moore & J. F. Lehman (Eds.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 591-596), Hillsdale, NJ: Erlbaum.
- Greeno, J. G., Engle, R. A., Kerr, L. K., & Moore, J. L (1993). Understanding symbols: A situativity-theory analysis of constructing mathematical meaning. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 504-509). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gresalfi, M. S., Martin, T., Hand, V. & Greeno, J. (2005). Constructing competence: An analysis of student participation in the activity systems of mathematics classrooms. Unpublished manuscript.

- Hall, R. (2000). Video recording as theory. In D. Lesh & A. Kelley (Eds.) *Handbook of Research Design in Mathematics and Science Education* (pp. 647-664). Mahwah, NJ: Erlbaum.
- Hatano, G., & Inagaki, K. (1991). Sharing cognition through collective comprehension activity. In L.B. Resnick, J.M. Levine & S.D. Teasley (Eds.) *Perspectives on socially shared cognition* (pp. 331-348). Washington, DC: American Psychological Association.
- Jordan, B. & Henderson, A (1995). Interaction analysis: Foundations and practice. *Journal of the Learning Sciences*, 4(1), 39-103.
- Kuhn, T. S. (1977). Objectivity, value judgment, and theory choice. In T. S. Kuhn (Ed.), *The essential tension* (pp. 320-339). Chicago: University of Chicago Press.
- Lachapelle, C. (1997). Looking at changes in student understanding using a situation model analysis of discourse. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (p. 978). Mahwah, NJ: Erlbaum.
- Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91-196). Cambridge, UK: Cambridge University Press.

- Leinhardt, G. & Steele, M. D. (2005). Seeing the complexity of standing to the side: Instructional dialogues. *Cognition and Instruction*, 23(1), 87-163.
- Linde, C., Greeno, J. G., Roschelle, J., Brereton, M., Lewis, J., & Stevens, R. (1994, April). *Explanations in discourse: Considering explanation as a social process*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- McMullin, E. (1976). The fertility of theory and the unit for appraisal in science. In R. S. Cohen, P.K. Feyerabend, & M.W. Wartofsky (Eds.), *Essays in memory of Imre Lakatos* (pp. 395-432). Dordrecht, Netherlands: Reidel.
- Nikitina, S. (2005). Pathways of interdisciplinary cognition. *Cognition and Instruction*, 23(3), 389-425.
- Ochs, E. (1979). Transcription as theory. In E. Ochs & B. Schieffelin, (Eds.). *Developmental pragmatics* (pp. 43-72). New York: Academic Press.
- Popper, K. R. (1959). *The logic of scientific discovery*. New York: Harper and Row.
- Reiser, B. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The Journal of the Learning Sciences*, 13(3), 273-304.
- Rosebery, A., Warren, B., & Conant, F. (1992). Appropriating scientific discourse: Findings from language minority classrooms. *The Journal of the Learning Sciences*, 2, 61-94.

- Shulman, L. S. (1992). Toward a pedagogy of cases. In J. H. Shulman (Ed.), *Case methods in teacher education* (pp. 1-30). New York: Teachers College Press.
- Stein, M. K., Engle, R. A., Hughes, E., & Smith, M. S. (2005). Orchestrating productive mathematical discussions: Helping teachers learn to better incorporate student thinking. Manuscript under revision.
- Webster, M. & Conant, F. R. (2003). "It says 'soap' now": Understanding young children's approaches to text in signs. Technical Report, Pioneer Valley Teacher Research Collaborative, Amherst, MA.
- Wertsch, J. V., & Toma, C. (1985). Discourse and learning in the classroom: A sociocultural approach. In L. P. Steffe & J. Gale (Eds.), *Constructivism in education* (pp. 159-174). Hillsdale, NJ: Erlbaum.
- Wiebe, M. (1999). Learning as transformation of participation: One student group's changing practice of doing research. In J. G. Greeno, A. L. Brown, A. L. & J. C. Campione, (Chairs). *Cognitive and social-interactional aspects of conceptual growth*. Symposium presented at the annual meeting of the American Educational Research Association, Montreal, Quebec.
- Yackel, E., Cobb, P. & Wood, T. (1991). Small-group interactions as a source of learning opportunities in second-grade mathematics. *Journal for Research in Mathematics Education*, 22(5), 390-408.

Footnotes

¹ The core group included Conant, Engle, Erickson, Greeno, and Muffie Wiebe Waterman.

² We could also make use of the fact that some rough transcripts had already been made as part of another analysis of this group's learning (Engle, 2004a).

³ However, although specific interpretations might not be attached to particular video records, the types of interpretations one can most easily make can be profoundly affected by choices about where to focus the camera and what kinds of audio information to record (Hall, 2000). Nevertheless, these constraints and affordances on interpretation are dwarfed by those characteristic of observational coding sheets. For them to work, it is necessary to pre-specify one's interpretive categories. This means that a key part of the interpretation of events is done before they even occur with other aspects done in-the-moment, as one fills out one's observation sheet. The degree and types of interpretative moves made in a given set of fieldnotes depend on the skills and methodological commitments of the person making them, but it is neither possible nor desirable to eliminate all interpretation from them. As it is only humanly possible to record a small subset of what one might be able to observe when making such notes, at the very least, one's selection of what to record ends up reflecting one's emerging interpretations about what is going on. Given this selectivity, a relatively "raw" record might not

be understandable. Such issues are well-recognized by experienced ethnographers and observational researchers so we do not wish to further belabor them here, in a book focused on the use of video.

⁴ In this transcript, colons represent elongation of vowels; periods between spaces indicate the lengths of pauses; underlining and capitalization each represent additional vocal emphasis; and paired left brackets indicate the beginning of overlapping speech. We are grateful to Fred Erickson for making this version of this transcript as well as for leading the effort to create transcripts that would more systematically record theoretically-relevant features of the discussions.

⁵ The orca controversy was relevant to whale features as, during their earlier debate, the students had tried to resolve the issue by comparing various features of an orca (especially, its dorsal fin) to those characteristic of different species of whales and dolphins.

⁶ Notice the contrast between this notion of accountability and the more externally-oriented one discussed more commonly in education in which outsiders evaluate the degree to which students have met standards. Here we are discussing a more internally-oriented notion of accountability in which students account for how their contributions make sense within the context of their learning environment.

⁷ Because Conant had participated in videotaping and analyzing the Water Taste Test discussions, we were required to account for a wider range of data than is usually available in published papers.

⁸ Fast Plants complete a life cycle in about 40 days, which helps support inquiry-based investigations of plant biology and related topics. For more information, see <http://www.fastplants.org/>.