A Learning Progression Emerges in a Trading Zone of Professional Community and Identity

Richard Lehrer
Vanderbilt University

Establishing a learning progression is an epistemic enterprise in that the goal is to position students to participate in the generation and revision of the forms of knowledge valued by a discipline. I describe steps taken to establish a learning progression intended to support the development of middle-school students' understandings of concepts and practices of data modeling. Data modeling refers to the invention and revision of models of chance to describe the variability inherent in a wide number of processes, from human- controlled enterprises like measuring or manufacturing to the outcomes of natural processes such as growing. Quantifying chance processes constitutes the cornerstone of statistical inference and hence of experiment, a practice that is often critical to the conduct of scientific inquiry. The construction of the learning progression involved analysis of core concepts and practices of data modeling that were apt to be generative, yet intelligible to students; conjectures about fruitful instructional means for supporting student learning developed in partnership with teachers; and the design and development of an assessment system that could be employed for both summative and formative purposes. Versions of the progression were investigated in a series of classroom design studies, conducted initially by the designers of the progression and later by teachers who had not participated in the initial iterations of the design. The movement from initial conjectures to a more stabilized progression involved both coordination and conflict among disparate professional communities, including teachers, learning researchers, and assessment researchers. Coordination and conflict among these communities were mediated by a series of boundary objects, ranging from curriculum units to construct maps that summarized typical trajectories of student learning. I conclude from these experiences that learning progressions do not accord well with metaphors of ladders or pathways of development. Instead, they are more like what P. Galison (1997) calls a trading zone.

The still-emerging Next Generation Science Standards (National Research Council, 2012) seek to position students to learn science by engaging in the epistemic practices by which scientific knowledge is generated and revised. These practices include the familiar emphasis on inquiry and investigation, but also some that are less familiar in K-12 education, such as generating and revising models and developing quantitative frameworks to sustain modeling and model contest. In the Next Generation Science Standards (NGSS), the work of quantification is spread across practices of interpreting data and in employing mathematics, including the mathematics of measure. The writers emphasize too that practices are intertwined; one does not develop quantities in the absence of legitimate inquiry. To articulate a framework for K-12 schooling, the NGSS rely heavily on the construct of learning progressions. Learning progressions articulate potential means for nurturing the long-term development of disciplinary knowledge and dispositions (National Research Council, 2006). I write this essay with two aims in mind. First, to date, learning progressions tend to represent the long-term development of core ideas in scientific disciplines, but the ways in which practices must co-develop to support the coordination of practice and the generation of core concepts are generally not as well articulated. Because models of phenomena are situated in practices of generating measures, structuring and visualizing measures, and often, making inference in light of uncertainty, I describe an effort to establish a progression of learning that engages students in these practices. I term this set of practices data modeling to emphasize the coordination of posing questions, generating measures and data representations and making inference in light of uncertainty (Lehrer & Romberg, 1996; Lehrer & Schauble, 2007). Models of data are critical intermediaries in the process of developing scientific models, but are often understated in pedagogical designs or are ritualized as procedures to be followed. The work of generating a learning progression contributes to the specification of what is meant by this form of quantification. The second aim follows from the first. I believe that widely conveyed images of learning progressions understate the nature of the disparate communities and forms of expertise that must be drawn together to make the progression pragmatically viable in the current context of schooling. Hence, I describe the arc of the development of the learning progression in data modeling, an effort that involved partnerships with software developers, assessment specialists, a teacher collaborative, and schools willing to serve as test-beds for multiple iterations of instructional design. In this essay I retrospectively view the design and development of the learning progression, a work that is still in progress, and suggest implications for the kind of work that must be accomplished if student participation in practices of science are to guide their emerging conceptions of the nature of the natural and designed worlds.

Getting Started

The initial impetus for the learning progression about data and statistics came about as Leona Schauble and I sought to extend a pre-existing progression of learning about spatial measure to include opportunities for children to develop ideas about sources of error in measure (Lehrer, 2003). Our work engages K-6 students in the invention and revision of models of natural systems (e.g., Lehrer & Schauble, 2006). Involving students in the construction of measures of these systems, and in the structuring of the resulting quantities as data, is fundamental to the practice of scientific modeling (e.g., Lehrer, Schauble, Carpenter & Penner, 2000; Lehrer & Schauble, 2002). Measure and data are motivated by questions that students pose about the systems they are investigating, so that all models of system behavior demand the concurrent development of models of data so that a conceptual pathway from inquiry to inference is made explicit, a process that we term data modeling (Lehrer & Romberg, 1996).

To embark students on the analysis of error, we involved them in the generation and analysis of repeated measure, a context inspired by accounts of the historic development of concepts of statistics in which uncertainty in this form of measure was consequential both celestially, as in magnitudes of star distances, and more immediately, as in commercial transactions (e.g., Porter, 1986). The initial context involved the design of model rockets, in which students inquired about the relation between the design features of their rockets (e.g., rounded vs. pointed nose cones) and the resulting apogees of flight. Students designed displays of data; our intention was to highlight how the shape of the data emerged from the choices about their organization made by their designers, and we challenged students to invent measures of the rocket's apogee and of the spread of the data. The intention was to ground statistics as sensible descriptions of distribution. The apogee served as signal so that statistics of center could sensibly refer to distance traveled. Similarly, the observed spread of the data could be interpreted as errors of measurement (Konold & Pollatsek, 2002). Students described sources of error and discussed why the spread of the data decreased as they became more practiced in their methods of measure. They also used the data to make inference about whether or not the design specifications of the rockets affected their apogees (Petrosino, Lehrer, & Schauble, 2003).

Following this initial design study, we conducted another during the following school year. Students in the fifth grade employed distributions to describe changes in the heights of a population of plants over the course of the plant life cycle (Lehrer & Schauble, 2004). We hoped to help students develop reasoning about population variability to complement their reasoning about individual variability: Coordinating these forms of reasoning is important for developing evolutionary accounts of the diversity of life. Because many of the students had participated in the previous year's research, we anticipated that they would readily make use of statistics of center and spread and that their efforts to organize the sample of measurements of the plants would be guided by appeal to the aggregate. Some students did, but most did not. In fact, they saw little rationale for a measure of center—they insisted that there was no privileged plant and there was certainly no true height (Lehrer & Schauble, 2007). Eventually, we were able to bridge this apparently simple distinction between contexts of measurement and natural variation, primarily by emphasizing growth as an instance of a repeated process and chance as a tool for considering what might happen if the collection of plants were grown again.

Design Study Iterations

We decided to conduct another iteration of the initial design study, but attending more closely to prospective pathways between measurement and natural variation by engaging students in the invention and revision of models of uncertainty in each context. Such an approach would foreground the epistemic function of models of uncertainty, which are often obscured in textbook treatments that do not integrate chance and data.

Our aim in this third study was to revisit the initial trajectory of learning in the first study and to extend it by involving students in the invention and revision of models of measurement error, a form of analysis of variance, which would set the stage for model-based inference. However, we soon found that this effort was more problematic than we anticipated. This was partly because we conducted the work in a new urban setting where the students little prior experience with mathematics beyond memorizing formulas for spatial measure, for operations on fractions, and for nearly all else that was mathematical. In many ways, the students and their teachers appeared emblematic of a procedural approach to mathematics education (Thompson & Thompson, 1994). Our challenge was to elaborate the instructional design in ways that would foreground sense making for these students. We took on the role of teachers to help students make the epistemic shift from relying solely on procedures to coming to see procedures as intimately related to mathematical ideas.

Over three years, we conducted three design experiments (Cobb, Confrey, diSessa, Lehrer & Schauble, 2003) in this setting, first with a small group of students for "enrichment," and subsequently, in one fifth- and one sixth-grade classroom. We employed methods of design research, including documenting forms of student thinking evident in classroom talk and in clinical interviews, to iteratively revise and refine a prospective trajectory of learning. By trajectory, I mean a prototypical sequence of forms of student thinking and of the learning ecology in which these forms of thinking emerge and are sustained. Designing learning ecologies includes specifying the tasks, tools and related material means of production, modes and means of argument, and recurrent forms of activity (e.g., participation structures) that one intends as the context for learning (Lehrer, 2009).

To initiate the learning trajectory, we sought to first pose a challenge or to otherwise create uncertainty about a central concept or idea. We refer to this as creating a problematic relation, often where none is initially perceived to exist by the learner. For example, we began again with repeated measurement, but this time, not of rocket apogee, but of a more readily determined length, that of a body part: the classroom teacher's arm- span and/or the circumference of a researcher's head. We provided tools for measure that we anticipated would influence the variability of measure, such as a 15 cm. ruler and a meter stick. The central concept was that the body-part had a measure, and the initial problematic was that the students did not know its value.

But they seemed to expect to find the same value. After all students had measured, it became evident that the anticipation of common measure was vastly under-stated. The batch was surprisingly variable, and there was some murmuring from students about "mistakes" and the moral character of those making them.

For each problematic, we sought to characterize the resources that students could bring to bear as they attempted to resolve it. For example, body measure relied on students' previous experiences (which we designed) with spatial measures to develop conceptions of the nature of units and of the nature of scales, and of how unit and scale are coordinated to produce measured quantities. Moreover, we anticipated that individual actions of measure, such as iteration of units, would be intelligible and that students would remember these forms of action, creating opportunities subsequently to use the classroom experience as a resource.

To complement problematics and resources, we designed further forms of assistance that we anticipated would not be in students' native repertoires. These additional meditational means often included participation structures with goals and participant roles that we anticipated would make the mathematical foundations of data and uncertainty more visible. For example, students invented displays of the measurements they had recently made, attempting to make visible to others any trends or patterns that they noticed when they looked at the collection of measurements. We designed a participation structure, *Display Review*, in which students other than the designer of the display reported what they thought that display showed about the data and what it tended to hide. "Show and hide" were repeatedly discussed to communicate the principle that representational choices always entail trade-offs. During these discussions, teachers asked all students to consider what choices the designer had made to render some characteristics visible and others less so. The intention was to highlight the role of order, count, class, and interval/scale in the resulting shape of the data. We did not anticipate that students would conduct analyses like these spontaneously and knew they would need explicit support. For each of the central concepts, we articulated how the learning ecology would function to sustain growth of an idea or particular relation among ideas. For example, we anticipated that the surprise engendered by variability of measure would motivate closer inspection of a new emergent—the collective--and the subsequent problematic of shape of the data. We anticipated, too, that by measuring, students would have at minimum a tacit understanding of potential sources of errors. Although from the perspective of the discipline, we might wish that students conceived of these errors as random, we anticipated that instead they would think of the process as deterministic. This attention to students' perspectives differentiates learning trajectories from Gagne-like disciplinary decompositions.

By means of conjecture and test, design research sheds light on the aspects of the design that are most fundamental and those that are extraneous or even deleterious. At this point on our work, we currently represent the learning trajectory as described in Table 1. Beginning with noticing differences among measurements, students structure the data as displays, ideally learning that attention to different mathematical characteristics of the collective, such as order, class, count and interval, result in different shapes of the data. Therefore, we are deliberately fostering representational (Greeno & Hall, 1997) and meta-representational (diSessa, 2004) competencies. We intend as well to support an image of a collection of data as resulting from a repetitive process (Thompson, Liu, & Saldanha, 2007). Students then invent statistics as measures of characteristics of the collection of measures—center or true measure and precision or variability among the measurements. Their inventions are influenced by the shape of the data and by the fundamental problematic of measure: How should the structure of the measure be coordinated with the structure of the distribution of data-measures? Statistics are second- order measures of the batch of initial measurements.

We next turn to the measure of uncertainty by engaging students in the analysis of simple chance devices. With the assistance of TinkerPlots™ tools, students model the behavior of these devices with an emphasis on long-run outcomes. The key is to reconcile trial-by-trial uncertainty with long-run structure and to coordinate two forms of estimation: theoretical and empirical probabilities. Students also develop models of the joint uncertainty of multiple, independent devices. This work sets the stage for modeling the initial batch of measurements obtained with two different measurement tools, a meter stick and a 15-cm ruler. Using the former typically results in less variable measurements when contrasted to the latter. Modeling here involves identifying sources of error, designing chance devices to reflect the probabilities of different magnitudes of error for each source, and combining the probabilities of these errors with a fixed estimate of the true length of measure. Model fit features prominently in comparisons among student- invented models; we are interested in cultivating a modeling stance in which statistics are re-purposed to summarize the behavior of models. This is a new form of explanation: the collection of measurements, which was previously conceived as determined, is now ideally re-conceived as partly reflecting chance. The unit concludes with model-based inference, in which the concept of sampling distribution plays a central role. A sample statistic is compared to a model-based sampling distribution of that statistic, again assisted by TinkerPlots™ tools that support partitioning of this (empirical approximation to a) sampling distribution.

Table 1. Summary of Learning Trajectory for Data Modeling

Problematic	Resources	Supporting Means	Conjecture
Difference among measures.	-Personal agency and access to measure process. -Theory-of- Measure.	-Repeated Measure Signal-Noise interpretation. -Familiar tools/ activity.	-Repeated measures support coordination of data & process via signal-noise interpretation.
Multiple shapes of the same data	-Pattern. -Invention. -Conversational norms.	-Technologies of paper. -*Display Review* activity structure. - Reflections on what would happen if the process were repeated.	-Image of repeated process emerges from participation and reflection on “again?” -Variable invented displays promote closer inspection of the mathematics of order, class, interval and count.

Problematic	Resources	Supporting Means	Conjecture
Statistic-as-measure of a characteristic of a distribution (not merely the result of a computation)	Theory-of-Measure. Shape of the Data. Algorithm. Invention.	-*Measure Review* activity structure. -Norms re measures as quantities, generalization. -Precision of measure focuses attention on center- spread. -TinkerPlots tools for partitioning, percent, display & measure.	-Variable products establish grounds for examination of correspondences between statistics and characteristics of distribution. -Contest about grounds of measure/ comparative goodness of measures promotes generalization.
Measure of uncertainty **Representation of sample-to-sample variability of a statistic**	-Theory-of- Measure. -Shape of the Data. -Statistics-as- measures. -TinkerPlots. -Image of repeated process.	-TinkerPlots sampler as tool for modeling uncertainty. -Investigations of the structure and behavior of familiar and simple random devices.	-Investigation of the structure and behavior of familiar and simple random devices supports coordination of theoretical and empirical estimates of probability.
Modeling measurement process with chance	-Intuitions of model as stand-in. - Resources of previous strand. - Sampling distribution - Use of TP to develop chance devices to model uncertainty.	-Identification of sources of measurement error. - Signed arithmetic. - Representing likelihoods of different magnitudes of error with TP. - Scaffolds for modeling observed measure as true score + error -*Model Review*	-Signal noise interpretation of sample measurements supports analysis-of-variance. -Signal-noise also renders intelligible invention of models as composed of true score and error. - During model reviews, comparisons among models provide grounds for considering properties of models such as intelligibility, and various senses of fit to data.

Having engaged students in this arc of practices of data modeling, we have positioned them in many ways to participate in building and revising models of natural systems, as described more completely elsewhere (Lehrer & Schauble, 2001, 2012). However, I turn now to the pragmatics involved in realizing the vision of the hypothetical learning trajectory, a transformation of trajectory into a viable learning progression.

Interactions with Other Communities, Other Professional Identities

Learning progressions are not merely trajectories in a new costume, but rather, as they are instantiated, they reflect an intertwining of distinct professional communities and practices. The first professional community with which we had significant interaction was that of a community of software designers.

Software Design Community

During the initial design studies, we were content to use hand-held spinners and related simple chance devices to model the variability inherent in processes of repeated measurement, but through collaboration with Cliff Konold (2007), we were members of the advisory group informing the design of what is now termed the Sampler in TinkerPlots 2.0. The upshot was that sample-to- sample variability could be more readily studied, its processes dissected, and the enterprise of modeling could enlist a powerful new technological ally. Over time, we diminished the use of hand-held devices, although we never did so completely, because the hand-held devices typically provoked discussion about the concept of trial in ways that appeared less visible in the digital environment. We know from previous research that the idea of a repeated (i.e., trial) process is not trivial, although it is often overlooked during instruction (Horvath & Lehrer, 1998).

Like us, Konold was also interested in students' propensities to invent the kinds of deviation-based measures of variability then being revealed by student inventions in our design studies. Although deviation measures of variability were represented in TinkerPlots 1.0, they required a process of entering formulas which many of our middle school students found formidable. With some help from us, Konold and his colleagues invented a Measure tool that provided students with means to invent and test their own measures of variability and of other characteristics of distribution as well. For the software design team, our design studies became grounds for pilot-testing new and potentially powerful software features. For us, the emerging capabilities of the software offered the prospect of significantly altering the trajectory of learning that we envisioned. As a result of the emergence of these enhancements to Tinkerplots, we changed the nature of instruction about chance and modeling to exploit new opportunities for supporting students' exploration of models and of sampling distributions.

Psychometric Community

At the outset of our work in the urban school setting, we collaborated with colleagues at Berkeley to develop assessment measures sensitive to forms of student reasoning that we observed as we conducted the design studies. We followed Wilson's (2005) design template to construct the assessment. That template begins with specifying constructs or progress variables that describe states of student knowledge about particular dimensions of learning, such as "Conceptions of Statistics" or "Meta-Representational Competence." For example, initially, students might not use statistics to describe data, preferring instead to rely exclusively on descriptors such as *hills*, *holes*, and *clumps*. Or, they might take a procedural view of data and calculate statistics without referring either to the nature of the data or the question posed. We anticipate that as students learn about statistics as measures, they will expect that minimum and maximum values of the statistic will correspond to different states of the characteristic of the distribution measured by the statistic. And, they may anticipate that the statistic will vary from sample to sample. On the progress variable, these different states of knowledge are conceived as milestones along a continuous dimension of learning. We eventually developed seven of these constructs to describe learning outcomes: (a) Theory of Measure, (b) Data Display, (c) Meta-Representational Competence, (d) Conceptions of Statistics, (e) Chance (f) Modeling Variability, and (g) Inference. Each construct describes milestones of learning and for each milestone, performances that can be taken as evidence of a particular milestone.

In Wilson's (2005) approach, items are designed to elicit student responses in ways that can be located on one or more construct maps. The type of item is up to the designer, but we have a preference for constructed response. Student responses are interpreted according to the differentiations illustrated in the progress variables, assisted by scoring guides that provide examples of prototypical student responses to assist reliable coding. The final step in the process is fitting responses to psychometric models that represent the probability of a correct response to an item as a logistic difference between an unknown parameter of a student's knowledge and another unknown parameter, the item's difficulty. The model centers the data at 0, suggesting that when student ability and item difficulty match, the probability of a correct response is ½.

As one might expect, this process of test design involved two constituencies with different traditions of learning and different approaches to the measurement of learning (National Research Council, 2001; 2006). We as learning researchers were more focused on the analysis of classroom interactions and the results of a particular form of assessment, the flexible interview. We tended to prefer images of learning that were not the step-like constructs favored by the psychometric community. But we reconciled our differences by adopting a position that the constructs defined a multi-dimensional space of learning outcomes, not a learning trajectory, and that our intention was to develop formative and summative components, so that testing could be put in service of learning.

Nonetheless, there remained significant obstacles to promoting this vision. The original design studies were conducted with small samples. Our focus on the particular helped us articulate the broad outlines of the construct space, and item design was informed by encapsulation of important forms of student reasoning revealed during the conduct of the design studies. Occasionally, colleagues from BEAR participated in the design studies. Yet, this process had its limits in that fitting data to psychometric models was not feasible: the sample sizes were far too small. Our solution was to develop a series of calibration studies with large (n tending to 1000) samples of students. Yet, data modeling is not an approach to teaching statistics that is prevalent in schools, so finding sites where the items could be calibrated was a challenge. We tried to solicit participation from some groups of teachers participating in reform curricula, especially CMP, but were stymied as we learned that data and statistics were beginning to be strained out of instruction as teachers were required to refocus their efforts in alignment with the national accountability movement. We were fortunate to enlist a regional collaborative of teachers in Northwest Arkansas as partners in the development of the assessment system. However, this recruitment significantly altered the nature and directions of the research. For our psychometric colleagues, the increased sample size provided opportunities toinvestigate issues of multidimensional item response modeling, statistical tests of growth models, and new approaches to testing relations among milestones of the dimensions suggested by the learning trajectory. These issues were of importance to us, but we are not experts or even especially well informed in these forms of modeling. We responded to model outputs and suggested which aspects of them made good theoretical sense and which appeared to be about tangential or even inconsequential questions.

But the BEAR team's conversations about the mathematics of the models are primarily conducted with other members of the psychometric community, not with us, albeit with occasional exceptions for doctoral students who have significant interest in these approaches to mental measurement. Hence, although there is a space of common interest, each of speaks in that space in a language that serves to communicate with reference to the learning progression but which omits significant details of our respective realms of work.

Teacher Community

To prepare teachers to be effective pedagogically and to grasp the wider sense of statistics and data implied by the learning trajectory, we worked with teachers (Year 1, n = 34; Year 2, n = 29) who had not participated in the original cycles of design. Seventeen teachers participated in two years of professional development. The professional development constituted yet another cycle of design, with day-long workshops that were again guided by conjectures about how problematics, resources, and supports might function to advance professional knowledge. Our aims were to engage teachers to (a) further develop their knowledge of statistics and chance; (b) understand the intention of the learning trajectory, including knowledge of expected patterns of student thinking; and (c) co-develop and employ the assessment system to interpret student responses and to use this information to guide instruction in classrooms. We positioned the joint work as negotiated in that the trajectories of student thinking outlined by the constructs were conjectures resulting from initial iterations of the instructional design, and that the items also enjoyed a provisional status. As we worked with teachers, we found it necessary to modify and even substantially transform material and conceptual arrangements that had served well to coordinate the activity of the assessment and learning researchers.

Educative curricula. One such transformation involved re-structuring and revising design documents from the previous studies into more familiar lesson formats. Each of seven units in the instructional sequence helps teachers understand the intention of the instructional design and hence is educative (Davis & Krajck, 2005). To assist teachers, each unit featured:

- *Teacher notes* that describe the educational intention of a task or problem posed to students;
- *Thought-revealing questions* that teachers can pose in their classrooms to instigate fruitful classroom conversation and reveal how students are thinking;
- *Students' ways of thinking* that describe prototypical forms of student reasoning and responses to curriculum tasks;
- M*athematical appendices* that elaborate the mathematical foundations of each lesson;
- *Extensions* that pose additional problems and contexts for students to explore, at the-discretion of the teacher.

Video annotated construct maps. We worked with teachers to revise the construct maps (progress variables) to meet their needs for readily accessible examples of student performances that illuminate milestones of progress. For this purpose we developed multimedia versions of the maps that included classroom-based, video examples of each learning performance. The initial examples were drawn from the design studies classrooms, but over time, teachers contributed episodes from their own classrooms. We also included video of formative assessment conversations that exemplified how teachers might employ discussion of the items to support students' conceptual change. Video examples were, again, initially drawn from researcher design studies, but subsequent examples drew from the classrooms of participating teachers. Video annotations clarified how student talk and activity correspond to particular levels of one or more constructs. These annotations helped teachers view learning performances more dynamically as the formative assessment conversations unfolded in the classrooms.

Formative assessment. Studies of teacher use of the assessment system suggested that there are prototypical patterns of progression in conducting these forms of assessment conversations. Kim (2010a, b) observed and interviewed a group of teacher-workshop participants to characterize how teachers made use of the assessment system to support their students' learning. She observed ten teachers (Year 1, eight teachers; Year 2, five teachers; three of the teachers were observed for two years). She noted that most teachers began by treating student responses to the assessment items as right or wrong, so that formative assessment conversations were characterized by traditional initiate-respond-evaluate cycles of classroom talk. Over time, teachers learned to interpret student responses to items within a mathematical horizon defined by milestones of reasoning as described by the construct being measured by the items. Teachers so inclined tended to select student responses for classroom sharing based on the milestones of the construct map, to invite explicit comparisons among forms of reasoning described by the construct, and to press for understanding how higher level forms of reasoning increased the scope and/or precision of an explanation.

Teachers' professional lives are not the same as those of either the learning or psychometric communities, as the modifications in objects shared by the three communities have clarified. Much has been made of the role of boundary objects like these as vehicles for coordinating the work of disparate communities with minimal disruption. Less has been written about the potential for transforming communities by their joint revision. For some of the participating teachers, images of their own profession began to include authoring representations of their classroom practices and objectifying these practices in the video exemplar construct map. This work transformed their professional identity, spurring some teachers to seek new branches in their own professional trajectories, primarily by assuming roles as math coaches or as math specialists in their districts.

Extending the Learning Progression by Engaging in Scale: Design Experiment

We are now continuing the development of the learning progression by conducting a design experiment involving approximately 40 schools and 80 teachers, with schools as the unit of random assignment. (I depart here from the usual nomenclature in that the term design experiment does not usually refer to an experiment but instead uses the term to signal case comparisons within and between successive iterations.) Half will adopt the progression-centered approach while other teachers continue their usual instructional practices. This modest effort at scale poses another set of challenges for development. One is professional development, led now by teachers and mathematics coaches who participated in the original professional development that we conducted. This effort has required more formal specification of the rationale and conduct of the professional development workshops. It has also introduced a new element of mathematics coaching that is still unfolding. A second challenge is the development of a measure of the fidelity of implementation, one that allows variation on themes established by the learning trajectory but that detects the extent to which teaching practices provide opportunities for students to learn in the manner anticipated by the trajectory. Further refinements of lessons have proved necessary, especially generation of activity structures related to supporting formative assessment conversations. We continue to revise the video annotated construct maps and expand them to encompass supporting curriculum and assessment. And, we are now engaged with yet another professional community, one centered about experimental design and statistical models that account for multiple levels of variability. In this world, many of the processes and problems inherent in disrupting pre-existing institutional practices of teaching are reduced to issues of "treatment fidelity." Yet, for those of us seeking to maintain and enhance the viability of the learning progression, these issues must be addressed productively, enlisting teachers as partners, with opportunities for enhancing professional identity. Else, the experiment will founder on the shoals of coercing teachers and students.

Discussion

Learning progressions include prospective pathways of conceptual development and specific realizations of these that we call learning trajectories. But this view of a learning progression tends to privilege conceptual development and understates the wrenching work of aligning disparate communities and interests in its service. When I reflect on the development of the data modeling progression, the need to align the communities of software design, psychometrics, curriculum design, and professional development leads me to believe it is more profitable to consider a learning progression as a *trading zone* (Galison, 1997) in which different realms of educational practice intertwine, much as a cable is constructed. Galison's (1997) term *trading zone* encapsulates how knowledge in particle physics gets constructed and revised. His study of knowledge-in-practice suggested that knowledge was constructed at the intersection of three distinct sub-cultures of physics, composed of instrumentalists, experimentalists and theorists. These sub-cultures interacted to produce knowledge, but did so in the manner of a type of commercial trade—each community understood its roles and obligations, but no theorist bothered to understand the pragmatics of every aspect of instrumentation, no instrumentalist bothered with the latest nuances in string theory, and so on. Galison suggested that each sub-culture enjoyed its own trajectory and professional identity, in contrast to the sweeping, vertically integrated periods of science suggested by Kuhn's now classic work on periodization in science. Like the physics subcultures, which work in partial encapsulation from each other with some information sharing across the boundaries, several decades of advances in psychometric modeling appear to have entirely missed the cognitive revolution and ways of characterizing knowledge generated by a group, like a class of students. Similarly, theories of learning appear to have ignored the emerging work in psychometrics and seem to have considered, if anything, only classical approaches to test theory. Thus, there may be good reason to consider the educational design enterprise as a trading zone, as well.

I have considered strands of learning theory, software design, assessment, instruction, and professional development as essential elements of the learning progression that we are attempting to establish for teaching and learning statistics. I have not highlighted other relevant strands, such as school districts and national standards boards, in which this and other learning progressions are embedded. Each of these communities has its own set of issues, questions, and history, and at different points of the development of a learning progression, each may take the leading role. Hence, it seems that the long-term prospects of a learning progression are forecast by its ability to harness disparate communities to its goals, each of which must view the learning progression as a meaningful site for achieving whatever it defines as accomplishments of value and as opportunity for crafting professional identity.

References

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in education research. *Educational Researcher, 32*(1), 9-13.

Davis, E. A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher, 34*(3), 3-14.

diSessa, A. (2004). Metarepresentation: Native competencies and targets for instruction. *Cognition and Instruction*, *22*(3), 293-331.

Galison, P. (1997). *Image and logic. A material culture of microphysics*. Chicago: University of Chicago Press.

Greeno, J. G., & Hall, R. (1997). Practicing representation: Learning with and about representational forms. *Phi Delta Kappan, January*, pp. 361-367.

Horvath, J. K., & Lehrer, R. (1998). A model-based perspective on the development of children's understanding of chance and uncertainty. In S. P. LaJoie (Ed.), *Reflections on statistics: agendas for learning, teaching and assessment in K-12* (pp. 121-148). Mahwah, NJ: Lawrence Elrbaum Associates.

Kim, M.-J. (2010a, April). A case of collaboration between researchers and teachers mediated by boundary objects. In R. Lehrer & M. Wilson (Chairs), *Assessing a multidimensional learning progression: Psychometric modeling and brokering professional development.* Symposium conducted at the annual meeting of the American Educational Research Association, San Diego, CA.

Kim, M.-J. (2010b, April). Collaboration with teachers: Brokering the assessment system.

In R. Lehrer & M. Wilson (Chairs), *Constructing a multidimensional learning progression for data modeling: Design studies, psychometric modeling and brokering professional development.* Symposium conducted at the meeting of the National Council of Teachers of Mathematics Research Presession, San Diego, CA.

Konold, C. (2007). Designing a data analysis tool for learners. In M. C. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 267-291). Mahwah, NJ: Lawrence Erlbaum Associates.

Konold, C., & Lehrer, R. (2008). Technology and mathematics education: An essay in honor of Jim Kaput. In L. D. English (Ed.), *Handbook of international research in mathematics education* (2nd ed). Philadelphia: Taylor & Francis.

Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal of Research in Mathematics Education, 33*(4), 259-289.

Lehrer, R. (2003). Developing understanding of measurement. In J. Kilpatrick, W. G.

Martin, & D. E. Schifter (Eds.), *A research companion to principles and standards for school mathematics*. (pp.179-192). Reston, VA: National Council of Teachers of Mathematics.

Lehrer, R. (2009). Designing to develop disciplinary knowledge: Modeling natural systems. *American Psychologist*, *64*(8), 759-771.

Lehrer, R., Kim, M-J., Ayers, E., & Wilson, M. (in press). Toward establishing a learning progression to support the development of statistical reasoning. In J. Confrey and A. Maloney (Eds.), *Learning over Time: Learning Trajectories in Mathematics Education.* Charlotte, NC: Information Age Publishers.

Lehrer, R., & Romberg, T. (1996). Exploring children's data modeling. *Cognition & Instruction, 14*(1), 69-108.

Lehrer, R., & Schauble, L. (2002, Eds.). *Investigating real data in the classroom: Expanding children's understanding of math and science*. New York: Teachers College Press. (Spanish Translation, Publicaciones M.C.E.P., Sevilla, Spain)

Lehrer, R., & Schauble, L. (2004). Modeling natural variation through distribution. *American Educational Research Journal*, *41*(3), 635-679.

Lehrer, R., & Schauble, L. (2006). Cultivating model-based reasoning in science education. In R. Keith Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 371-387). Cambridge: Cambridge University Press.

Lehrer, R., & Schauble, L. (2007). Contrasting emerging conceptions of distribution in contexts of error and natural variation. In M. Lovett & P. Shah (Eds.), *Thinking with data.* (pp. 149-176). New York: Taylor & Francis.

Lehrer, R., Schauble, L., Carpenter, S., & Penner, D. E. (2000). The inter-related development of inscriptions and conceptual understanding. In P. Cobb, E. Yackel, & K. McClain (Eds.), *Symbolizing and communicating in mathematics classrooms: Perspectives on discourse, tools, and instructional design* (pp. 325-360). Mahwah, NJ: Lawrence Erlbaum Associates.

National Research Council (2001). *Knowing what students know. The science and design of educational assessment*. J. Pellegrino & R. Glaser (Eds.). Board on Testing and Assessment, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, D.C.: National Academy Press.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council (2006). *Systems for state science assessments*. Committee on Test Design for K-12 Science Achievement. M.R. Wilson & M. W. Bertenthal (Eds.). Board on Testing and Assessment, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Petrosino, A., Lehrer, R., & Schauble, L. (2003). Structuring error and experimental variation as distribution in the fourth grade. *Mathematical Thinking and Learning*, *5*(2&3), 131-156.

Porter, T. M. (1986). T*he rise of statistical thinking 1820-1900*. Princeton,NJ: Princeton University Press.

Thompson., A. G., & Thompson, P. W. (1994). Talking about rates conceptually, part I.: A teacher's struggle. *Journal for Research in Mathematics Education*, *25*, 279-303.

Thompson, P. W., Liu, Y., & Saldanha, L. A. (2007). Intricacies of statistical inference and teachers' understanding of them. In M. C. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 207-231). Mahwah, NJ: Lawrence Erlbaum Associates.

Wilson, M. (2005). *Constructing measures*. Mahwah, NJ: Lawrence Erlbaum Associates.

Footnote

The research reported here was supported by the National Science Foundation and by the U.S. Department of Education, Institute of Education Sciences. The positions represented here are not endorsed by either agency.