DOCUMENT RESUME

ED 466 728 SE 066 302

AUTHOR	Lesh, Richard
TITLE	Dealing with Complexity: New Paradigms for Research in Mathematics Education.
PUB DATE	2000-10-00
NOTE	20p.; In: Proceedings of the Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (22nd, Tucson, AZ, October 7-10, 2000). Volumes 1-2; see ED 446 945.
PUB TYPE	Opinion Papers (120) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Educational Research; Elementary Secondary Education; Higher Education; Learning Strategies; Learning Theories; *Mathematics Education; *Numeracy; Professional Development; *Program Development; *Research Methodology

ABSTRACT

This paper focuses on some of the most important factors that appear to have strongly influenced the development of the dominant research designs used in mathematics education research. Factors discussed that have influenced research design in mathematics education include a concern about increasing the relevance of research to practice and the recognition of the complexity of all aspects of teaching and learning. Some relevant assumptions about student development, teacher development, and program development are also examined. (DDR)

DEALING WITH COMPLEXITY: NEW PARADIGMS FOR RESEARCH IN MATHEMATICS EDUCATION

Richard Lesh
R.B. Kane Distinguished Professor
Purdue University

During the closing years of the twentieth century, a number of books and articles have been published describing the status of research in mathematics education and discussing possible ways to increase its power and usefulness (Sierpinska & Kilpatrick, 1998; Grows, 1992; Steen, 1999). Most of these publications have focused on summaries of past research, or on descriptions of the authors' views about which problems or theoretical perspectives they believe should be treated as priorities for future research. — Should teachers' decision-making issues be treated as higher priorities than those that confront policy makers or others who influence what goes on in classroom instruction? Should issues of equity be given priority over those involving content quality, or innovative uses of advanced technologies? Should theoretical perspectives be favored that are grounded in brain research, or artificial intelligence models, or constructivist philosophies? Should quantitative research procedures be emphasized more than qualitative procedures? — In this paper, such questions are not my central concerns. Instead, I'll ask: *"What kind of research designs have proven to be especially useful in mathematics education, and what principles exist for improving (and assessing) the quality of these research designs?"*

In general, the preceding publications suggest that: (a) mathematics education research has made far less progress than is needed, and (b) little attention has been given to many of the most important issues that are priorities for teachers or other practitioners. — I don't dispute these claims. But, in general, I'm far more impressed with the achievements than with the shortcomings of mathematics education research. For example, during the two decades that have passed since the first meeting of PME/NA was held at Northwestern University in 1978, it would be striking to anyone who attended that meeting that the mathematics education research community has made enormous progress to shift beyond theory *borrowing* toward theory *building*. At the 1978 meeting, most of us were doing Piagetian Research, or Vygotskian research — or research based on psychometric models, or information processing models, or artificial intelligence models — where both our theoretical models and our research methodologies were borrowed from these other fields. But, today, examples abound where mathematics educators have developed their own distinctive theoretical models, conceptions of critical problems, research literature, research tools and procedures — and, most importantly, communities of inquiry — in topic areas ranging from early number concepts, to rational number concepts, to early algebraic reasoning.

As a result of the preceding progress, rapid increases have occurred in the volume and sophistication of mathematics education research; and, these increases have ushered in a series of paradigm shifts that involve new ways of thinking about the nature of students' developing mathematical knowledge and abilities, as well as new ways of thinking about the nature of mathematics, problem solving, learning, and teaching — and new systemic ways of thinking about program development, dissemination, and implementation. For example, the paper that describes the PME/NA 2000 *Models & Modeling Working Group* gives a concise description of some of the most significant shifts in thinking that apply to research being conducted by members of this group. Such new ways of thinking have provided primary driving forces behind many of the most successful recent attempts at standards-based curriculum reforms; and, they've also created the need for new research methodologies that are based on new assumptions, and that focus on new problems and capitalize on new opportunities.

Unfortunately, researchers are similar to other busy people who know that *when you're up to your elbows in alligators there's not much time to consider innovative ways to drain the swamp.* Thus, the development of widely recognized standards for research has not kept pace with the development of new problems, perspectives, and research procedures in our field; and, as a consequence of this fact, because there is a lack of clarity and consensus about appropriate principles for optimizing (or assessing) the quality of innovative research designs, three kinds of undesirable results are likely to occur when proposals are reviewed for projects, publications, or presentations. First, appropriate methodologies may be marred by avoidable methodological flaws. Second, studies employing appropriate methodologies may be rejected because they involve unfamiliar research designs, or because inadequate space is available for explanation, or because inappropriate or obsolete standards are used to evaluate them. Third, inappropriate methodologies may be accepted because they employ traditional research designs - even though the assumptions that they presuppose may be antithetical to perspectives the researcher wants to adopt about the nature of mathematics, teaching, learning, or problem solving.

To develop productive ways of dealing with the preceding difficulties, I recently served as the co-editor of an NSF-supported *Handbook of Research Design in Mathematics & Science Education* (Kelly & Lesh, 2000). This handbook includes chapters written by more than forty leading researchers in mathematics and science education. It's aim was to emphasize research designs that: (a) have been pioneered by mathematics and science educators, (b) have distinctive characteristics when used in mathematics or science education, or (c) have proven to be especially productive in mathematics or science education.

Examples of such research designs include several different types of teaching experiments, and distinctive types of clinical interviews, videotape analyses, naturalistic observations, and action research paradigms in which participant-observers may include not only researchers-acting-as-teachers or classroom teachers-acting-as-researchers but also curriculum designers, software designers, and teacher educators whose aims include both optimizing and understanding mathematics teaching, learning, or problem solving. In general, these new research designs draw on multiple types of quantitative and qualitative information; the knowledge-development products they produce often are not reducible to tested hypotheses or answered questions; and, they often involve cyclic and iterative techniques in which participant-researchers include a variety of interacting students, teachers, and other mathematics educators. Finally, and most importantly, they often involved new ways of thinking about the nature of students' developing mathematical knowledge and abilities, and new ways of thinking about the nature of effective teaching, learning, problem solving.

The purpose of our *Handbook of Research Design* was to clarify the nature of some of the most important experience-tested ways to improve (or assess) the usefulness, power, share-ability, and cumulativeness of the results that are produced when the preceding kinds of research designs are included in proposals for research projects, publications, or presentations at professional meetings. Of course, from the beginning of our efforts, participants were mindful of the fact that, if obsolete or otherwise inappropriate standards are adopted, then the results could hinder rather than help. But, as long as decisions must be made about funding, publications, and presentations, it is not possible to avoid issues related to quality assessments. Decisions WILL be made. Therefore, our goal was to try to increase the chances that appropriate issues will be considered and that productive decisions will be made.

For details about factors that contribute to the quality of specific kinds of teaching experiments, or other research designs emphasized in the handbook, readers are referred to the handbook itself. This chapter will restrict attention to some of the most important factors that appear to have strongly influenced the development of virtually all of the research designs that have developed in distinctive ways in mathematics education.

Two Factors Influenced Research Designs That are Distinctive in Mathematics Education

In the development of the *Handbook*, two factors emerged as having especially strong influences on the kind of research designs that have been pioneered by mathematics educators. First, most are intended to radically increase the relevance of research to practice - often by involving practitioners in the identification and formulation of problems to be addressed, in the interpretation of results, or in other key roles in the research process. Second, there is a growing recognition that students, teachers, classrooms, courses, instructional programs, curriculum materials, learning tools and minds are all complex systems, taken singly, let alone in combination. Therefore, regardless of whether we focus on the developing capabilities of students, or groups of students, or teachers, or schools, or other relevant learning communities, the continually evolving ways of thinking of each of these "problem solvers" involve complex conceptual systems that are dynamic, living, self-regulating and continually adapting – and that have competencies that generally cannot be reduced to simple checklists of condition-action rules. They don't simply lie dormant until being stimulated. They initiate action; and, when they are acted on, they act back. So, interactions often involve feedback loops that lead to second-order (or higher-order) effects that overwhelm first-order effects. Furthermore, among the most important systems that mathematics educators need to investigate and understand: (a) many do not occur naturally (as "givens" in nature) but instead are products of human construction, (b) many cannot be isolated because their entire nature may change if they are separated from the complex holistic systems in which they are embedded, (c) many may not be observable directly but may be knowable only by their effects, or (d) when they're observed, changes often are induced that make investigators integral parts of the systems being investigated. So, there is no such thing as an immaculate perception; and, the behaviors of these systems often cannot be described adequately using simple algebraic, statistical, or logical formulas.

Because of the complex, constructed, and systemic nature of most of the "subjects" and "constructs" that mathematics educators need to investigate and understand, it's become commonplace to hear mathematics education researchers talk about rejecting traditions of "doing science" as they imagine it is done in the physical sciences (where, it is imagined, researchers treat "reality" as if it were objectively given). But, when educators speak about rejecting notions of objective reality, or about rejecting the notion of detached objectivity on the part of the researcher, such statements tend to be based on antiquated notions about the nature of modern research in the physical sciences. For example, in mature sciences such as astronomy, biology, chemistry, geology, or physics, when entities such as subatomic particles are described using fanciful terms such as color, charm, wisdom, truth, and beauty, it is clear that the relevant scientists are quite comfortable with the notion that reality is a construct; and, when these scientists speak of principles such as the *Heisenberg Indeterminancy Principle*, it also is clear that they are familiarity with the notions that: (i) the relevant systems act back when they are acted upon, (ii) the observations researchers make often induce significant changes in the systems they observe, and (iii) researchers often are integral parts of the systems they are hoping to understand and explain. Yet, such realities do not prevent these researchers from developing a variety of levels and types of productive operational definitions to deal with constructs such as black holes, neutrinos, strange quarks and other entities whose existence is related to systems whose behaviors are characterized by mathematical discontinuities, chaos, and complexity.

Consider the case of the neutrino where huge vats of heavy water are surrounded by photo-multipliers in order to create situations in which the effects of neutrinos are likely to be observable and measurable; and, notice that, even under these conditions, neutrinos cannot be observed directly. They can be known only through their effects; and, between the beholder and the beheld, elaborate systems of theory and assumptions are needed to distinguish signal from noise and to shape interpretations of the phenomena under investigation. Also, small changes in initial conditions often lead to large effects that are essentially unpredictable; observations that are made induce significant changes in the systems being observed; and, both researchers and their instruments are integral parts of the systems that scientists are hoping to understand and explain.— So, mathematics educators are not alone in their need to deal with systems that have the preceding characteristics.

As the diagram in Figure 1 suggests, in mathematics education, just as in more mature modern sciences, it has become necessary to move beyond machine-based metaphors and factory-based models to account for patterns and regularities in the behaviors of complex systems; and, it also has become necessary to move beyond the assumption that the behaviors of these systems can be explained using simple linear combinations of unidirectional cause-and-effect mechanisms that can be accurately characterized using models from elementary algebra, statistics, or logic.

From an Industrial Age using analogies based on *hardware*		**Beyond an Age of Electronic Technologies** using analogies based on computer *software*		**Toward an Age of Biotechnologies** using analogies based on *wetware*
where systems are considered to be no more than the sum of their parts, and where the interactions that are emphasized involve no more than simple one-way cause-and-effect relationships.	⇒	where silicone-based electronic circuits may involve layers of recursive interactions which often lead to emergent phenomena at higher levels which are not derived from characteristics of phenomena at lower levels	⇒	where neurochemical interactions may involve "logics" that are fuzzy, partly redundant, partly inconsistent, and unstable – as well as living systems that are complex, dynamic, and continually adapting.

Figure 1. Recent Transitions in Models for Making (or Making Sense of) Complex Systems

In the *Handbook of Research Design*, the chapter on "operational definitions" describes how scientists in more mature sciences specify ways to cause, recognize, and measure the occurrence of complex systems (or "elements" whose existences depend on complex systems) without reducing these "operational definitions" to checklists of condition-action rules.[1] For example, when devices such as cloud chambers or cyclotrons are used to observe, record and measure illusive constructs whose existences depend on complex systems, it is clear that: (i) the relevant construct does not *reside in* the device,[2] (ii) being able to measure a construct does not guarantee that a corresponding dictionary-style definition will be apparent, and (iii) even when a dictionary-style definition can be given (for a construct such as a black hole in astronomy), this doesn't guarantee that procedures will be available for observing or measuring the construct. Nonetheless, useful operational definitions usually involve three parts that are similar, in some respects, to the following three parts of traditional types of behavioral objectives (of the type that have been emphasized in past research in mathematics education).

Behavioral Objectives have Three Parts

> GIVEN {specified conditions} THE STUDENT WILL EXHIBIT {specified behaviors} WITH IDENTIFIABLE QUALITY (perhaps specified as percents correct on relevant samples of tasks, or perhaps specified in terms of a correspondence with certain criteria for excellence).

Whereas behavioral objectives collapse three different kinds of statements into a single condition-action rule - and thus treat all knowledge as if it consisted of nothing more than condition-action rules - more general types of operational definitions keep these components separate.[3] For example, when researchers in fields such as physics deal with complex phenomena involving such things as photons or neutrinos, a minimum requirement for a useful operational definition is that explicit procedures must be specified for three things:

1. situations that optimize the chances that the targeted construct will occur in observable forms.[4]
2. *observation tools* that enable observers to sort out signal from noise in the results that occur.
3. assessment criteria that allow observations to be classified or quantified.

Although it's beyond the scope of this chapter to give details about principles that mathematics educators can use to deal with the preceding three components of productive operational definitions, examples can be found in some of the best standards-based "performance assessment instruments" that have been developed during recent years (Lesh & Lamon, 1993). For instance, in the research design book's chapter on *Principles for Developing Thought-Revealing Activities*, examples are given of performance assessment instruments that involve: (i) *thought-revealing activities* that require students, teachers, or other relevant "subjects" to express their ways of thinking in forms that are visible to both researchers and to the subjects themselves, (ii) *response analysis tools* that teachers (or researchers) can use to classify alternative responses, and to identify strength and weaknesses, as well as directions for improvement, and (iii) *response assessment tools* that can be used to evaluate the relative usefulness of alternative responses.

What are Some Relevant Assumptions about Student Development, Teacher Development, and Program Development?

In general, research involves getting clearer about the nature of the subjects (e.g., students, teachers, curricula, etc.) or constructs (e.g., students' conceptual systems) that are under investigation; and, current assumptions about the nature of these subjects or constructs have strong influences on decisions about *whom* to observe, *what* to observe, *when* to observe, and *which* aspects of the situation to observe – as well as decisions about what kind of tools to use to filter, denote, organize, analyze, and interpret the information that is gathered or generated. For example, even when data collection involves tools such as video recordings, which sometimes give the illusion of capturing "raw data", researchers' prejudices about the nature of relevant subjects strongly influence decisions to focus on one situation rather than another, to zoom in on the behaviors of one participant rather than another, or to emphasize one "window" of observation (or one aspect of the situation) rather than another.

The impact of prejudices about what is important and what is not are especially important to consider when tests or questionnaires are used that give the appearance of being prejudice free – because the assumptions on which they are based often are buried under many layers of psychometric fog. In the research design book, most of the chapters in the section on "assessment design" deal explicitly with these sorts of issues. They emphasize that some of the most important factors influencing the quality of research involve the degree of alignment between: (i) intended assumptions about the nature of the subjects (or the constructs) being investigated, and (ii) assumptions presupposed (implicitly or explicitly) by the procedures that are used for selecting, analyzing, and interpreting information. – Consider the following instances of "subjects" commonly investigated by mathematics educators:

(a) *Common Assumptions about Students*:

When mathematics education research investigates the nature of the evolving conceptual schemes that students use to interpret their learning or problem solving experiences, one common assumption is that "thinking mathematically" involves more than simply computation with written symbols. For example, it also involves mathematizing experiences - by quantifying, dimensionalizing, coordinatizing, or in other ways making sense of them using mathematical constructs (systems). Consequently, to investigate the nature of students' constructs and sense-making abilities, researchers generally need to focus on problem solving situations in which interpretation is not trivial; and, in non-trivial situations, most modern theories of teaching and learning believe that the way students' interpret learning and problem solving experiences is based on interactions between (internal) conceptual systems and by (external) systems that are encountered or constructed. Therefore, when interpretation is not trivial, different students are expected to interpret the situation in fundamentally different ways.[5] Furthermore, when the goal of a task involves developing an interpretation, the description or explanation that's produced often involves significant forms of learning. Therefore, when the products involve learning, students who engage in a sequence of such tasks that are "structurally similar" would not be expected to perform in the same way across all tasks.

The preceding observations raise the following kinds of questions in the design of research. When a given student is not expected to perform in the same way across a series of similar tasks, what does it mean to speak about "reliability" in which repeated measurements are assumed to vary around the student's "true" (invariant) understandings and abilities across all tasks? When different problem solvers are expected to interpret a single problem solving situation in fundamentally different ways, what does it mean to speak about "standardized" questions? What does it mean to speak about "the same treatment" being given to two different participants or groups? – By raising such questions, I do not mean to suggest that constructs such as "reliability," "validity" or "replicability" are not relevant to modern research in mathematics education. Indeed, closely related criteria (such as usefulness, meaningfulness, power, and share-ability) always must be part of productive knowledge development in applied fields such as mathematics education. But, the meanings these constructs must be conceived in ways that are consistent with our best assumptions about the nature of systems and constructs being investigated; and, this means that off-the-shelf definitions borrowed from obsolete theories may no longer be appropriate – especially if they are grounded in machine-based models of teaching, learning, and problem solving.

In the *Handbook of Research Design*, many of the chapters identify common mismatches that occur between the assumptions underlying procedures and theories commonly employed by mathematics education researchers. They also suggest procedures in which these assumptions are better aligned. For example, in a chapter about Multi-Tier Teaching Experiments, several interacting levels and types of "subjects" (students, teachers, curriculum designers, researchers) each are engaged in sequences of *thought-revealing activities* in which the repeatedly express their current ways of thinking in forms that go through multiple testing-and-revising cycles. A byproduct of this process is that the trail of documentation that is produced yields a trace (that's analogous, in some ways, to the trace produced by an electron in a cloud chamber in a physics laboratory) that reveals the nature of the constructs that the student produces.

If one asks *"Where is the mathematics?"* in the preceding kinds of thought-revealing activities, the answer is that it resides in the responses that students generate. It does not reside, as many naïve task analyses have assumed in the past, in the tasks themselves.

(b) Common Assumptions about Teachers:

For teachers just as for other types of problem solvers and decision makers (including students), expertise is reflected not only in what they *do* but also in what they *see*. Alternatively, we could say that what teachers *do* is strongly influenced by what they *see* in given teaching and learning situations. For example, as teachers develop, they tend to notice new things about their students, about their instructional materials, and about the ideas and abilities that they are trying to help students learn; and, these new observations often create new needs and opportunities that, in turn, require teachers to develop further. Also, the situations that teachers encounter are not given in nature; they are, in large part, created by the teachers themselves – based on their current conceptions of teaching, learning, and problem solving. Thus, there exists no fixed and final state of excellence in teaching; and, in fact, continual adaptation (development) is a hallmark of teachers who are successful over long periods of time. Furthermore, no teacher is equally effective for all grade levels (kindergarten through college), for all topic areas (algebra through statistics and geometry or calculus), for all types of students (handicapped through gifted), and for all types of settings (inner-city through rural). No teacher can be expected to be constantly "good" in "bad" situations; not everything that experts do is effective, nor is everything that novices do ineffective; and, characteristics that lead to success in one situation (or for one person) often turn out to be counterproductive in other situations (or for another person). Furthermore, even though gains in students' achievements should be primary factors to consider when documenting the accomplishments of teachers (or programs), it is foolish to assume that the best teachers always produce the largest gains in student learning. One reason this is true is because some great teachers choose to deal primarily with difficult students or difficult circumstances?

The preceding observations suggest that expertise (for teachers, for students, or for other problem solvers and decision makers) tends to be plural, multidimensional, non-uniform, conditional, and continually evolving. In general, there is no single "best" type of teacher; every teacher has a complex profile of strengths and weaknesses;

teachers who are effective in some ways and under some conditions are not necessarily effective in others; and, teachers at every level of expertise must continue to adapt and develop. – So, what does it mean to classify teachers into categories such as "experts" or "non-experts" as if it was legitimate to collapse complex profiles of capabilities into fixed positions on a single-dimension "good–bad" scale?[6]

In the *Handbook on Research Design*, one approach for dealing with the preceding issues involves *Evolving Expert Studies* in which thought-revealing activities for *students* often provide the basis for equally thought-revealing activities for *teachers* (and others). For example, using *Multi-Tier Teaching Experiments*, teachers (like students) are put into a series of situations where their views must be expressed in forms that are tested and revised or refined repeatedly and iteratively. Formative feedback and consensus building provide mechanisms to encourage development in directions that participants themselves are able to judge to be increasingly "better" – without basing judgments on some preconceived notions of "best."

(c) Common Assumptions about Programs, Materials, or Classroom Learning Environments:

Because classroom learning environments, schools, and programs are not given in nature, constructs and principles that can be used to construct, describe, explain, manipulate, or control such systems often appear to be less like "laws of nature" than they are like "laws of the land" that govern a country's legal system. Also, researchers who are involved in investigating such systems often are not simply disinterested observers, and, they may be more interested in "what's possible" than in "what's real or typical." Therefore, issues about the truth or falsity of given principles or perspectives may be less pertinent than issues about consistency, meaningfulness, power, and the desirability of outcomes.

Legislated programs, defined curricula, and planned classroom learning environments often are quite different than implemented programs, curricula, and classroom activities; and, complex programs, materials, and activities seldom function like simple functions in which a small number of input variables completely determine a small number of output variables. For example, second-order effects (and other higher-order effects) often have impacts that are highly significant; and, emergent phenomena resulting from interactions among variables often lead to results that are at least as significant as attributes associated with the variables themselves. In particular, tests often go beyond being objective indicators of development to exert powerful forces on the programs, curricula, or activities that they are intended to assess. Consequently, if naïve pretest-posttest designs reflect narrow, shallow, or naïve conceptions of outcomes and interactions, then they often have strong negative impacts on outcomes – so that researchers frequently need to abandon assumptions about their own detached objectivity.

The *Handbook of Research Design* includes a number of chapters that describe alternatives or supplements to pretest-posttest designs. Many of these approaches to research could be called *design studies* – because the products that the research produces consists of (or include) models, conceptual tools, and other artifacts that must be designed to accomplish specified tasks - or to satisfy specific design principles).

Throughout this paper, I use the term "research design" rather than "research methodology." This is because, in mathematics education, the design of research generally involves trade-offs often similar to those that occur when other types of complex products (such as automobiles) need to be designed to meet conflicting goals (such as optimizing speed, safety, and economy). So, whereas the term "research methodology" tends to be associated with statistics-oriented college courses in which the emphasis is on how to carry out "canned" computational procedures for analyzing data, the kind of situations and issues that are most important for math educators to investigate seldom lend themselves to the selection and execution of off-the-shelf data analysis techniques. Combinations of qualitative and quantitative approaches tend to be needed; and, in addition to the stages of research that deal with data analysis, other equally important issues typically arise that involve: (i) developing productive conceptions of problems that need to be solved, products that need to be produced, or opportunities that need to be investigated, (ii) devising ways to generate or gather relevant information to develop, test, refine, revise, or extend relevant ways of thinking, (iii) developing appropriate ways to sort out the signal from the noise in information that is available - and to organize, code, and interpret raw data in ways that highlight patterns and regularities, or (iv) analyzing underlying assumptions and formulating appropriate implications.

It Often Is Said That Good Research Requires Clearly Stated Hypotheses Or Clearly Stated Research Questions

Dealing with complex systems in a disciplined way is the essence of research design in mathematics education; and, it is the central theme of this paper. Relevant perspectives involve cognitive science, social science, mathematical sciences, and a wide range of other points of view. No single means of understanding is sufficient; no single style of inquiry is likely to take us very far; and, relevant research can never be reduced to a formula-based process. Far from being a process of using "accepted" techniques in ways that are "correct", it's a "no holds barred" process of developing shared knowledge about important issues. Doing it well involves *developing a chain of reasoning* that is *meaningful*, *coherent*, *sharable*, *powerful*, *cumulative*, *auditable*, and *persuasive* to a well-intentioned skeptic about issues that are *priorities* to address.

When we emphasize that research is about the development of knowledge, it should be clear that what we know consists of a great deal more than *tested hypotheses* (stated in the form of "if ... then ...." rules) and *answered questions* (using standardized tests, questionnaires, or other techniques leading to quantitative measures of relevant variables). For example, some of the most important products of research also include:

- *Descriptions* and *explanations* (e.g., models and conceptual systems) for constructing and making sense of complex systems. So, truth and falsity may not be at issue as much as fidelity, internal consistency, and other characteristics similar to those that apply to quality assessments for painted portraits or verbal descriptions.
- *Demonstrating possibilities* that may involve existence proofs (with small numbers of "subjects") and that may need to be expressed in forms that are acompanied by (or embedded in) exemplary software, informative assessment instruments, or illustrative instructional activities, programs, or prototypes to be used in schools. So, the quality of results depends on the extent to which these products are meaningful, sharable, powerful, and useful for a variety of purposes and in a variety of situations.
- *Developing tools* such as those that are intended to be used to increase (or document or assess) the understandings, abilities, and achievements of students, teachers, programs, or relevant learning communities. So, again, the quality of such tools depends on the extent to which they are sharable, powerful, and useful for a variety of purposes and in a variety of situations. (Note: These tools may or may not involve measurement or quantification.)

Similar products of research are familiar in the natural sciences. For example, in fields such as physics, chemistry, or biology, some of the most important products of research involve the development of *tools* or *explanatory models* that involve references to phenomena such as waves, fields, and black holes, that provide different ways of describing, explaining, constructing, manipulating, and predicting the behaviors of complex systems. Yet, these tools often generate information that goes beyond characterizing complex systems using a single number; and, they often go beyond comparisons that collapse all relevant attributes onto a single-dimension number line. Similarly, the models often are iconic and analog in nature, being built up from more primitive and familiar notions - so that the visualizable model is a major locus of meaning for relevant scientific theories. They are not simply condensed summaries of empirical observations.

It Often is Said that Math Education Research has not Answered Teacher's Questions

If the point of the preceding statement is to emphasize that projects emphasizing the development of knowledge should make a difference in mathematics teaching and learning, then I concur. But, the view that *"teachers should ask questions and* *researchers should answer them"* is naïve and counterproductive to the point of being a large part of the disease for which it purports to be the cure. Consider the following observations.

- In mathematics education, no clear line can be drawn between researchers and practitioners. There are many levels and types of both researchers and practitioners; and, people who are known as "researchers" often have equally strong reputations as teachers, teacher educators, curriculum developers, or software developers. Similarly, many people who are best known in these latter areas also are highly capable researchers. Also, practitioners whose voices should be heard include not only teachers but parents, policy makers, administrators, school board members, curriculum specialists, textbook writers, test developers, teacher educators, and others whose decisions strongly influence what goes on in schools. So, the process of knowledge development is far more cyclic and interactive than is suggested by one-way transmissions in which teachers ask questions and researchers answer them (see figure 2).

- In mathematics education, productive knowledge development projects often involve some form of curriculum development, program development, or teacher development; and, productive curriculum development, program development, or teacher development projects also should involve knowledge development. Such endeavors shouldn't be artificially separated. In any of the preceding areas of development (including the development of presentations at professional conferences such as NCTM, AERA, or PME), it's obvious that, if progress has been made, it is precisely because *we know more*. Similarly, where less progress has been made, the knowledge base has tended to be weak.

- What people ask for isn't necessarily a wise statement of what they need; and, useful tools and conceptual systems usually need to be developed iteratively and recursively. – For many of the practitioners mentioned on the preceding page, the "problems" that they pose often focus on "symptoms" rather than on underlying "diseases"; or, they sound more like "ouches" (expressions of difficulty or discomfort) than they do like well formulated problems. Consider the politician who says: *"Show me what works?"* Need I say more? Whereas small innovations seldom lead to large results, large innovations seldom get implemented completely. Yet, nearly every educational innovation works some of the time, in some situations, for some purposes, in some ways, and for some students. So, unless it's known which parts work when, where, why, how, with whom, and in what ways, the pseudo-information that *"This program (or policy) works!"* is likely to be misleading to educational decision

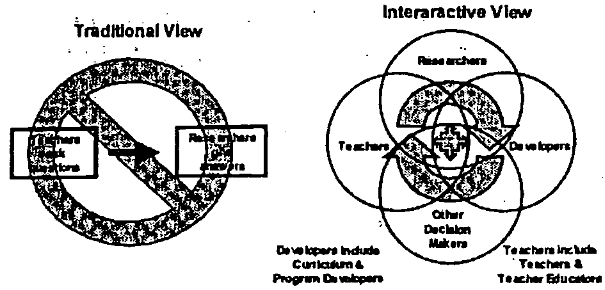

Figure 2

makers. Implementations of sophisticated programs and curriculum materials generally involve complex interactions, sophisticated adaptation cycles, iterative developments, and intricate feedback loops in which breakdowns occur in traditional distinctions between researchers and teachers, assessment experts and curriculum developers, observers and observed.

- Among the challenges and opportunities that mathematics educators confront, most are sufficiently complex that they are not likely to be addressed effectively using results from a single research study. Therefore, rather than thinking in terms of a one-to-one match between research studies and solutions to problems, it would be more productive to insist that results from research studies should contribute to the development of a theory (or a model) – and that this theory should have powerful implications over a reasonable period of time (see figure 3). This is why community building is important; and, it's why, in addition to factors such as usefulness, and share-ability, cumulativeness is another factor that should be considered when assessing the significance of research results.

Mathematics Education is an Exceedingly Young Field of Scientific Inquiry

Many of the most influential leaders who attended the first meeting of PME/NA are no longer among us. Merlyn Behr, Bob Davis, Jack Easley, Nick Herscovics, and Claude Janvier, in particular, were influential in shaping the spirit of PME/NA. Others have retired, or will be retiring soon, who provided leadership through early significant stages in the development of the mathematics education research community. So, with this turn-over, it's not surprising that emerging new leaders sometimes find it difficult to appreciate how brief our history has been.

Many indicators exist to suggest that mathematics education is in its infancy as a field of scientific inquiry. In fact, some cynics might even claim that we need to move beyond Piaget's "unconscious play stage" – where only primitive processes have evolved for planning, monitoring, and assessing our own activities. Instances supporting such claims include (and are partly caused by) the following kinds of commonplace events.

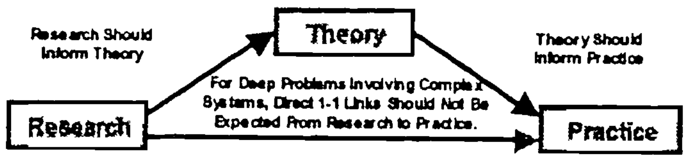

Figure 3

- At professional meetings, it's not difficult to find "research" sessions that consist of little more than stories told about a videotaped episode of teaching, learning, or problem solving.
- In proposal reviews for projects, publications, or presentations, it's not uncommon to hear "quantitative methods" being treated as if they were automatically "scientific" – even when: (i) the "control groups" don't control anything significant, (ii) the pretests and posttests ignore the most important characteristics of the constructs (or subjects) being investigated, and (iii) the analysis procedures presuppose psychometric models whose assumptions are not consistent with modern conceptions of mathematics, problem solving, learning, or effective instruction.[7]
- In promotion reviews for universities with no active researchers on their faculty, reviewers frequently find themselves answering questions like: (i) Could a chapter in a book possibly be as significant as an article published in a journal?[8] (ii) Are all journals equally significant? (iii) Is collaboration or co-authorship a sign of weakness? (iii) Should investigations employing qualitative methods be discounted?[9]

David Johnson didn't edit the first issue of the *Journal for Research in Mathematics Education* until 1970. The first *International Congress of Mathematics Education* (ICME) wasn't held until 1968; and; the *International Group for the Psychology of Mathematics Education* (PME) emerged from Efriam Fischbein's working group on the *Psychology of Mathematics Education* at ICME-II in Exeter, England, in 1972.

One of the main factors leading to the formation of PME/NA was, of course, the formation of PME-International. But, two other significant forces included: (i) activities associated with the *Georgia Center for Research in Mathematics Education* (Steffe, 1974), and (ii) a series of small annual mathematics education research conferences, sponsored by Northwestern University's *Center for the Teaching Professions* (Lesh, 1973), beginning in 1972.

Distinctive characteristics of all three of the preceding initiatives was that they focused on: (i) *stimulating* research (to focus on priority problems), (ii) *facilitating* research (by sharing tools, resources, and perspectives), (iii) *coordinating* research (to form a community of researchers whose work would be more cumulative and mutually supportive), and (iv) in other ways *amplifying* the power and utility of research that already is going on – not on *buying* isolated research projects (through subsidies for salaries, equipment, or supplies), and not on simply providing a venue for people to *talk about results from past projects*.

One reason why these initiatives were so powerful, and why their second-order and higher-order effects have continued to be forceful even today, is that the people who initiated them were active researchers who were intimately in touch with the strengths and needs of the emerging community. In particular, it was clear that a large share of the research that goes on in mathematics education is conducted by:

- doctoral students as part of their degree requirements,
- "researchers" who often have even stronger identities as curriculum developers, program developers, or teacher developers – or as excellent teachers.
- people who do not have large funded projects but who view research as an important part of their professional lives in order to get jobs, promotions, or tenure, and most of all to continue to learn and develop by thinking critically about their teaching and other educational endeavors.
- researchers from fields outside of mathematics education (such as developmental psychology) where mathematics education has a long history of serving as an unusually productive site to explore and test emerging theories.

To harness these resources, PME/NA was intended to be a place that's dedicated to the development of a community of researchers - where participants could: (i) plan future activities as well as report results from past activities, (ii) hear divergent views from colleagues in other fields, and (iii) report research that isn't several years old due to cumbersome review procedures.

I believe that we're currently at a very exciting time in the development of the mathematics education research community. On the one hand, if we compare what is known today with what was known during the "new math" movements of the 1960s, it's clear that a great deal has been learned about children's ways of thinking about elementary-but-deep constructs in topic areas ranging from early number concepts, to rational numbers and proportional reasoning, to geometry and measurement, to early algebra or statistics. On the other hand, as we move from one of these research topics to another, or from one cluster of researchers to another within a given topic, it's equally clear that the accounts of conceptual development often are based on radically different assumptions about the nature of mathematics, problem solving, learning, and teaching. So, mismatches need to be reconciled; and, significant new paradigm shifts can be expected. For example, during the next decade, due to results from fields such as brain research, due to the explosion of ways that mathematics is used beyond schools in a technology-based *age of information*, due to the emergence of new ways to document these abilities, and due to the availability of many new types of modeling tools that are especially well suited to describing the kind of complex, dynamic, interacting, and continually adapting systems that characterize so many of the subjects and constructs that we need to understand in mathematics education, I expect that mathematics educators will need to rethink many fundamental assumptions about the nature of mathematical ability (Lesh & Lamon, 1993; Doerr & Lesh, in press).

Finally, new communication technologies are making it possible for close collaboration to occur among researchers representing multiple perspectives at remote sites. As an example where these new kinds of communication are being used to facilitate the development of a research community, Purdue University, Indiana University, Purdue-Calumet, IU/PUI (Indianapolis) have created a new *Distributed Doctoral Program* (DDP) [10] in which many of the key courses are co-taught by faculty members representing any of the four campuses, and students also can participate from any of the four campuses. Similarly, a loosely knit federation of leading research institutions also have participated in several of the shared courses that are part of the DDP. For example, during the Spring of the 1999-2000 academic year, a course was taught on *Research Design in Mathematics and Science Education.*[11] It included participants from all four Indiana campuses and also Arizona State University, SUNY-Buffalo, Queensland University of Technology in Australia, and the University of Quebec at Montreal in Canada. Similarly, during the Fall of the 2000-2001 academic year, a course is being taught on *Models & Modeling in Mathematics & Science Education,*[12] and, again, participating campuses will include all of the preceding institutions plus Syracuse University. Finally, early in the Spring Semester of the 2000-2001 academic year, the PU/IU Distributed Doctoral Program will play host to the first annual *Distributed Doctoral Research Conference in Mathematics Education.* Like courses in the DDP, this conference will use internet-based videoconferencing and other communication tools to enable participants to interact from remote sites.[13]

One reason why the preceding kinds of collaboratively taught courses are important is because doctoral students in mathematics education are perfect examples of students with highly specialized needs, that do not occur in sufficient numbers on most campuses so that their needs can be addressed effectively. Another reason is that multiple-campus collaboratively taught courses can provide ideal ways to promote the kind of community building that will be needed to address many of the most important issues that mathematics educators need to understand.

Notes

1. Even in everyday situations, thermometers measure temperature; yet, it's obvious that simply causing the mercury to rise doesn't do anything significant to change the weather. Clocks and wrist watches measure time without leading us to believe that they tell what time really is. Symptoms may enable doctors to diagnose a disease;

yet, it's clear that eliminating the symptoms is different than curing the disease.
2. Whereas behavioral objectives treat mathematical ideas if they *resided in* specific problems or tasks, modern mathematics education researchers have turned their attention beyond analyses of "task variables" to focus on analyses of "response variables" – because mathematical thinking resides in students' interpretations and responses, not in the situations that elicited these mathematical ways of thinking.
3. In many respects, the development and assessment of complex conceptual systems is similar to the development and assessment of complex and dynamic systems that occur in other fields - such as sports, arts, or business - where coordinated and smoothly functioning systems usually involve more than the simple sums of their parts. For instance, it may be true that a great artist (or athlete, or team) should be able to perform well on certain basic drills and exercises; but, a program of instruction (or assessment) that focuses on nothing more than checklists of these basic facts and skills is not likely to promote high achievement. If we taught (and tested) cooks in this way, we'd never allow them to try cooking a meal until they memorized the names and skills associated with every tool at stores like Crate & Barrel or Williams Sonoma; or, if we taught (and tested) carpenters using such approaches, we'd never allow them to try building a house until they memorized the names and skills associated with every tool at stores like Ace Hardware and Sears. But, in education, it's common to treat low level *indicators* of achievements as if they *embodied* or *defined* the understandings we want students to develop.
4. Even if it's impossible to reduce Granny's cooking expertise to a checklist of rules for others could follow to duplicate here abilities, it may be quite easy to identify situations where her distinctive achievements are required – and where many of the most important components of here abilities will be apparent.
5. This is because a variety of levels and types of interpretations are possible, a variety of different representations may be useful (each of which emphasize and de-emphasize somewhat different characteristics of the situations they are intended to describe), and different analyses may involve different "grain sizes", perspectives, or trade-offs between factors such as simplicity and precision.
6. Is a Ford Taurus better or worse than a Jeep Cherokee? Clearly, answers depend on purpose, context, and other factors that apply to assessments of any complex system that is intended to function in complex situations.
7. The Handbook of Research Design gives a number of detailed examples of each phenomena.
8. In mathematics education, a great deal of the best work of the most productive researchers has never fit the constraints of the *Journal for Research in Mathematics Education*. Consequently, at least since the early 1970's, there has been a healthy "black market" of research publications – such as those associated with early years of the "Georgia Center" (Steffe, 1974).

9. My experiences coincide with others who write large numbers of review letters for people being considered for jobs, promotions, or tenure. That is, the frequency of such questions appears to be inversely related to the quality of the institution asking them.
10. For information about the Distributed Doctoral Program, contact Terry Wood (twood@purdue.edu), Frank Lester (lester@indiana.edu), Beatrice D'Ambrosio (bdambro@topaz.iupui.edu), Erna Yackel (yackeleb@calumet.purdue.edu).
11. For information about this course, contact Jim Middleton (jmiddleton@)asu.edu) or Marilyn Carlson (carlson@math.la.asu.edu), Doug Clements (clements@acsu.buffalo.edu), Lyn English (l.english@qut.edu.au), or Carolyn Kieren (kieran.carolyn@uqam.ca).
12. For information, contact Helen Doerr (hmdoerr@sued.syr.edu) or Richard Lesh (rlesh@purdue.edu).
13. For information about how to participate in this DDRCME conference, contact Richard Lesh (rlesh@purdue.edu), Judi Zawojewski (judiz@purdue.edu), or Kay McClain (Kay.McClain@vanderbilt.edu).

References

Doerr, H. & Lesh, R. (in press). *Beyond constructivism: A models & modeling perspective on mathematics problem solving, learning, & teaching*. Hillsdale, NJ: Lawrence Erlbaum.

Kelly, A. & Lesh, R. (2000). *Handbook of research design in mathematics & science education*. Hillsdale, NJ: Lawrence Erlbaum.

Lesh, R., & Lamon, S. (Eds.). (1993). *Assessment of authentic performance in school mathematics*. Hillsdale, NJ: Lawrence Erlbaum.

Steffe, L. (1974). *The Georgia Center for Research in Mathematics Education*. Funded by a small grant from the National Science Foundation.

Grows, D.A. (1992). *Handbook of research on mathematics teaching and learning*. New York: Macmillan.

Sierpinska, A. and Kilpatrick, J. (1998). *Mathematics education as a research domain: A search for identity.* Dordrecht, The Netherlands: Kluwer Academic.

Steen, L. (1999). Theories that gyre and gimble in the wabe. *Journal for Research in Mathematics Education 30(2),* 235-241.

NOTICE

Reproduction Basis

- [x] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

- [ ] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").