Tools, Researchable Issues & Conjectures for Investigating
What it Means to Understand Statistics (or Other Topics) Meaningfully

Richard Lesh
Indiana University
ralesh@indiana.edu

Abstract
This paper describes a variety of sharable multi-purpose research tools which have evolved from recent studies in which models & modeling perspectives (Lesh & Doerr, 2003) were used to investigate what it means for students to develop meaningful understandings of foundation-level concepts and abilities in an introductory statistics course. Such a course typically is intended to “cover” topics ranging from basic measures of probability, centrality, and spread, to more advanced topics such as analysis of variance, hypothesis testing, or regression and correlation. Although MMP studies have involved students from grade school through graduate school, as well as professionals in fields ranging from education to engineering, the students referred to in this paper were sophomore elementary education majors at Indiana University. This group is identified simply to give readers a frame of reference for observations made - not because it served as a treatment group or a control group of any kind… These students had sufficiently strong academic records to be admitted to a university program with high academic standards; but, compared with peers in other fields, none considered themselves to be outstanding in mathematics. The tools described here were designed mainly to investigate the following questions. (i) What are the most important “big ideas” that should be emphasized in a given mathematics topic area (e.g. statistics)? (ii) What does it mean to “understand” these ideas? (iii) How do these understandings develop? In fields like engineering, or in other “design sciences” which are more mature than mathematics education, researchers in leading research communities often devote significant portions of their time and efforts toward the development of tools and artifacts for their own use. Whereas, in mathematics education, our research community has spent relatively little time developing such tools and resources. One result of this neglect is that mathematics educators are unable to reliably observe, document, or measure either students’ or teachers’ levels of development for nearly any of the deeper or higher-order achievements that current theories hypothesize to be important. For this reason, this paper focuses on theory-based tools and tool development; and, to begin, main elements of MMP theoretical foundations are described in enough detail so that the tools will be useful to other researchers. Throughout this paper, bulleted questions or statements indicate issues about which a great deal is known – but that still should be investigated more thoroughly in order to more thoroughly understand their meanings and implications.

1. Distinctive Goals of MMP Research

MMP research begins with the recognition that, in virtually every field where researchers have investigated differences between high ability people and those with lesser abilities, results consistently show that exceptionally effective people not only DO things differently, but they also SEE (or interpret) things differently. So, MMP research generally begins with studies designed to clarify the nature of the most important interpretation systems (i.e. models) that underlie competence in a given course or topic area (e.g., algebra, calculus, geometry, statistics). For example:

- What makes statistical interpretation systems different than geometric interpretation systems, or algebraic interpretation systems, or other types of interpretation systems?
- What changes are taking place in the kinds of problem solving and decision making situations where some type of “statistical reasoning” is needed in a technology-based age of information?
- What concepts, abilities, and underlying interpretation systems are most useful in the preceding kinds of problem solving situations?
- Which among the preceding concepts, abilities, and interpretation systems are most fundamental - because of their power, accessibility, and generativity?
- What does it mean to “understand” the underlying interpretation systems associated with these concepts and abilities?
- How and in what ways do these understandings develop?
- (Last of all, and only after answers to the preceding questions become more clear), how can we facilitate document, measure, and assess development?

Notice that, in MMP research, models are expected to be among the most important kinds of knowledge that students, teachers, and researchers or developers develop. And, in the case of researchers, new models usually are expected to be accompanied by revised research questions, new conjectures, or new tools and artifacts for investigating the nature of students’ evolving knowledge and abilities.

2. Theoretical Foundations for MMP Research:

Regardless whether the “subjects” of MMP research are students, teachers, developers, researchers themselves, or other education decision makers, the models that any of these “subjects” are expected to develop are of two basic types. The first type of model is used to interpret, describe, or explain systems that already exist. For example, for curriculum developers, such models may be needed to assess new kinds of mathematical thinking that are becoming important beyond school in the 21st century. The second type of model is used to design new systems (such as new assessment systems) or to shape the development of dynamically evolving systems (such as students’ or teachers’ thinking when new technology-based tools are available). In either case, a distinctive characteristic of MMP research methodologies is that all levels of participants – from students, to teachers, to developers, to researchers - are considered to be model developers (Lesh in English, et. al. 2005). For example, in a typical MMP study: (i) Students often develop models to describe or design mathematical tools or artifacts in responses to carefully designed problem solving activities that MMP researchers refer to as model-eliciting activities (MEAs). (ii) Teachers often develop artifacts (e.g., assessment activities focusing on higher-order understandings) or tools (e.g., observation tools for observing students’ thinking during MEAs. (iii) Researchers develop models to facilitate and make sense of interactions between students and teachers within carefully designed learning environments. … For any of these model developers, important underlying assumptions of MMP research include the following.

- Model development is expected to involve a series of parallel and interacting sequences of iterative cycles in which current ways of thinking are repeatedly expressed, tested, and revised. Furthermore, for situations that involve learning or problem solving, first draft descriptions, explanations, or designs are not expected to be adequate for the situations, purposes, or problems at hand. Second-iteration drafts and higher-iteration drafts tend to be needed. And, Nth-iteration models tend to inherit characteristics from more than a single evolutionary strand – in somewhat the same way that grandchildren inherit characteristics from more than a single grandparent.

- Every model is recognized to be (no more than) a useful oversimplification of the situations, artifacts, or tools that the model is intended to describe, design, or develop. So, every current model (and every model-based artifact or tool) is expected to be the Nth in a continuing series of extensions, revisions, or rejections; designing for future adaptations is an important attribute of successful model development; and, understanding causes of failure tends to be as important as explaining success.

- Early models are expected to be relatively barren and distorted compared with later models. For example, using the language of Piaget (Lesh, 2005): (i) primitive models are expected to center on only a subset of relevant characteristics, (ii) primitive models are expected to involve a small collection of relatively fuzzy, undifferentiated, and unintegrated partial interpretations of the situations, and (iii) primitive models are expected to alternate among fuzzy and loosely connected forest-level and tree-level interpretations, descriptions, or designs. In other words, primitive models are expected to be relatively unstable compared to later-iteration models.

Beyond the preceding assumptions about the nature of mathematical thinking and learning, MMP research also emphasizes the fact that one of the most important characteristics which distinguishes mathematical thinking from other types of thinking is that mathematics is the study of
structure. So, unlike physics (which studies the physical properties of relevant systems), or chemistry or biology (which study the chemical or biological properties of relevant systems), mathematics studies the structural properties of systems. Consequently, because the term “model” is the simplest and most straightforward name that is used to refer to interpretation systems that scientists develop in mathematics and the natural sciences, “model” is the term that MMP uses to refer to the interpretation systems that students develop. However, compared with formal and analytic models that professional scientists develop, students’ models are models in-the-making. That is, they are at beginning or intermediate stages of development. They are.

What is a model? In MMP research, an entry-level conception of a model is: A model is a system for describing or designing some other system for some specific purpose. (According to this definition, notice that it is possible for a system to be modeled at one moment and to serve as a model for some other system at a later moment – and that it is possible to use a model without much awareness of its existence.)

Assumptions about the Systemic Nature of Mathematical Models. Another significant assumption underlying MMP research should be attributed to Piaget – even though this assumption is easier to state using the modern mathematical language associated with complex adaptive systems (Lesh, 2005). One of the most important characteristics of complex adaptive systems is that many of their most important properties cannot be described using only a single input-output rule. Such properties are called emergent properties of the system-as-a-whole; and, they arise because of feedback loops, second-order effects, or patterns of the system which arise because of interactions among elements of the system – but which cannot be deduced from properties of the elements themselves.

Assumptions associated with complex adaptive systems lie at the heart of Piagetian psychology (Lesh & Carmona, 2002). This is why many of Piaget’s most famous tasks were designed to: (i) demonstrate the systemic nature of many of the most important foundation-level concepts in mathematics, (ii) demonstrate the nature of pre-systemic thinking (i.e. pre-operational thinking), and (iii) describe the ways that thinking develops from non-systemic to systemic thinking in tasks which focused on issues such as: invariance (with respect to a system), transitivity (with respect to a system), or patterns and regularities (of systems) that might involve the maximization, minimization, or stabilization of relevant systems.

In retrospect, after seeing the results of Piaget’s ingenious studies, the claim that mathematical ideas tend to have meanings which depend on emergent properties of conceptual systems-as-a-whole should not seem surprising. To see why, consider the fact that the formal construction of every mathematical system begins with “undefined terms” which make it possible to state axioms whose meanings are assumed to be understood. But, in the thinking of real life learners and problem solvers, there is no such thing as beginning with “undefined” terms; and, even in mathematical formalizations, what it really means to be and “undefined term” is that all of the mathematical meanings of these terms come from the system-as-a-whole in which they reside. In other words, they are emergent properties of relevant systems-as-a-whole. Similarly, in mathematical formalizations, even though relevant axioms are stated sequentially, they are assumed to function simultaneously and synchronously. Yet, in the thinking of real life learners and problem solvers, it is entirely possible for conceptual systems to function in “uncoordinated” ways (i.e., in ways so that relevant operational/relational systems do not function as systemic wholes). Consequently, when the functioning of a system is relatively “uncoordinated” (i.e., unsystemic) students tend to “see the ‘forest’ and loose cognizance of the ‘trees’” – or vice versa. And, when they focus on one type of detail, they often loose cognizance of other details.

Stated in the language of complex adaptive systems, Piaget’s research demonstrated that many of the most powerful elementary-but-deep concepts in mathematics are (or depend on) emergent properties of conceptual systems-as-a-whole. One important implication of this claim is that “understanding” such properties and concepts is not reducible to learning a condition-action rule – or even a list of condition-action rules. This is because emergent properties cannot take on their intended meanings until students (or other MMP research subjects) interpret situations using well coordinated conceptual systems-as-a-whole.
The preceding claims about the systemic nature of mathematical thinking are especially important in countries such as the USA, whose curriculum standards consistently fragment knowledge into lists of low-level rules (i.e., basic facts and skills) or high-level rules (e.g., problem solving rules or processes which operate on lower-level facts and skills). Consequently, regardless whether attention focuses on low-level or high-level rules, the development of powerful, sharable, and reusable interpretation systems tends to be ignored.

One reason for neglecting models as goals of learning is because it is not widely understood how to document or assess system-dependent concepts or abilities. But, another reason is because the notion of emergent properties of systems involves many assumptions which appear to contradict a variety “common sense” notions about teaching, learning, and assessment. For example, if Piagetian perspectives are taken seriously, then issues soon arise which are similar to the riddle: What comes first, chickens or eggs? A Piagetian answer to this question is: “Both!” Both develop in parallel and interdependently. So: Which comes first, the system or its emergent properties? Piaget’s answer to this question was to study both developmentally, and to notice that (i) systems evolve out of systems, and, (ii) for young children, the earliest of these mathematical systems tend to involve systems of interactions with concrete materials or with other people.

Assumptions about the Situated Nature of Mathematical Models. For the purposes of mathematics educators, one of the foremost weaknesses of Piaget’s theory results from the fact that he was most interested in describing the systemic nature of thinking. Consequently, he tended to deemphasize the situated nature of thinking. For example, in his efforts to demonstrate the importance of systemic thinking, he focused special attention on major conceptual reorganizations that tend to occur in the thinking of children. For example: (i) When children are approximately 5-7 years of age, a wide range of concepts begin to emerge which presuppose thinking based on “concrete operational systems” (i.e., systems in which the relevant relationships and operations act on point-to-able “objects”). (ii) When children are approximately 12-16 years of age, a wide range of concepts begin to emerge which presuppose thinking based on “formal operational systems” (i.e., systems in which the relationships and operations act on concrete operational “objects”). … But, whereas Piaget was most interested in major conceptual reorganizations which occur at the onset of concrete operational thinking or formal operational reasoning, mathematics and science educators tend to be more interested in conceptual developments which occur before, between, and beyond these periods when major conceptual reorganizations occur. So:

− During intermediate periods of development, most mathematics concepts (and underlying conceptual systems) tend to be at intermediate stages of development – not pre-operational, but also not fully operational; and, especially during these intermediate periods, knowledge tends to be organized around experience at least as much as around abstractions.

For example, in nearly every one of Piaget’s books which provide evidence about the systemic nature of children’s thinking, the results also show that; (i) The same tasks that Piaget used to demonstrate the importance of systemic thinking also clearly show that two tasks which appear to involve the same kind of systemic thinking often are significantly different in difficulty. (ii) A child whose thinking is at stage N for one task (or concept) often is at stage N±n on another task which seems to involve the same structure. In fact, Piaget used the term decalage to refer to this fact. (iii) In mathematics education research, a series of projects known as “The Rational Number Project” clearly demonstrated that small changes in Piaget-like tasks often significantly change their level of difficulty. (iv) Vygotsky used the term zones of proximal development to describe how the difficulty of a given tasks can be changed significantly depending on interactions with a teacher or conceptual guide. (v) Research based on Models & Modeling Perspectives (MMP) has shown that, during the solution of specially designed model-eliciting activities (MEAs), the thinking of learners or problem solvers often evolves from a Piaget-like stage N to stage N+n during a single 60-minute model development activity (Lesh & Harel, 2003). That is, significant adaptations often occur during brief periods of time; and these adaptations are molded and shaped by the situations in which they occur.

− One of the most important characteristics that distinguishes MMP from Piagetian perspectives is that models are assumed to be highly situated, continually adapting, richly distributed, and socially
shaped human constructs. Furthermore, models tend to integrate ideas and procedures from a variety of textbook topic areas.

Note: It is important to recognize that claims about the situated nature of mathematical knowledge do not mean that relevant understandings are only context specific – with little or no transfer to other situations. For example, imagine a graphics-oriented spreadsheet which students might develop for some purpose (such helping another person to buy a car, or figuring out how much to pay their governments in taxes); then, imagine how some spreadsheets are more powerful (in the given situation), more sharable (with other people), or more reusable (in other situations) than others. … Similarly, even though interpretation systems are conceptual tools that are strongly shaped by the situations and purposes for which they were developed, they generally are not worthwhile to develop unless they are intended to be not only powerful (in one specific situation), but also sharable and reusable (for other people in somewhat different situations). And, powerful, sharable, and reusable knowledge is transferable knowledge. So, whereas lack of transfer has been the “Achilles heel” of projects which have attempted to teach higher-order processes (or “metacognitive processes”), overgeneralization tends to be more of a difficulty for learning related to interpretation systems (or models). To see why, consider the similarity between unstable conceptual systems and uncoordinated physical activities such as hitting baseballs, tennis balls, or golf balls. When a baseball player first begins to play golf or tennis, s/he is likely to engage her baseball-hitting schemes to hit golf balls or tennis balls. Yet, in both golf and tennis, baseball-hitting schemes are not the best ways to swing a club or a racquet. Similarly, if a person develops a metaphor for thinking about some kind of situation, then they are likely to engage this metaphor in other situations that they consider to be similar – even when the metaphor is only loosely appropriate. ... For these reasons, overgeneralization tends a greater difficulty than lack of transfer when developing interpretation systems.

Assumptions about the Distributed and Infrastructural Nature of Mathematical Models. People in modern societies are continually off-loading functions which once needed to be executed “in their heads” (with no support from conceptual tools); and, such people also are continually projecting their interpretation systems into the world in the form of designed tools and artifacts – which have the characteristic that, as soon as people come to better understand them, they tend to change them. … In a technology-based age of information, many of the most important “things” that impact the lives of ordinary people are systems that are designed or developed by humans. So, the worlds in which we live are increasingly designed worlds; and, the systems that impact lives range in size from global economic systems or ecological systems to the kind of systems that are embodied in gadgets that are ubiquitous in modern societies.

Assumptions about the Connected and Cross-Disciplinary Nature of Mathematical Models. Models which are useful in realistically complex “real life” situations often require problem solvers to integrate concepts and procedures drawn from a variety of disciplines, theories, and textbook topic areas. One reason this is true is because, in fields like engineering, knowledge development is problem-driven as much as it is theory-driven. In “pure” sciences, relevant theories dictate what problems are most important to solve, and the theories also provide criteria for judging when problems are solved; whereas, in applied sciences, problems tend to arise from outside of any theory, and the criteria for being done also reside outside of any theory. The following facts provide some other reasons why solutions to realistically complex problems usually need to draw on a variety of practical and theoretical perspectives. (i) Problem solvers usually do not have as much time, money, or other resources as they would want under ideal conditions. (ii) Solutions often need to satisfy conflicting purposes (such as low costs and high quality) - when “customers” do not agree about the relative importance of different criteria for success (such as low risks but the possibility of high gain). (iii) In “real life” problem solving situations beyond school, problem solvers often are not isolated individuals. Instead, “problem solvers” often are teams of diverse specialists – each interpreting the situation in different ways, and each using different kinds of technology-based tools. So, for all of the preceding reasons, models tend to be chunks of knowledge which integrate ideas and procedures drawn from a variety of textbook topic areas – as well as a variety of disciplines or theories.

Assumptions about the Undifferentiated and Unintegrated Nature of Primitive Mathematical Models. Learners’ and problem solvers’ primitive interpretations of situations tend to involve several
undifferentiated and un-integrated partial interpretations. In particular, when people interpret situations, they often engage more than a single model (Lakoff & Nunes), metaphorical story (Schank’s “scripts”), conceptualization/interpretation/explanatory system (Novak), representational system (deSessa), or conceptual framework. So, in a given situation which is being interpreted, there tend to be a variety of levels and types of interpretation systems which are partly accessible and engaged; and, each tends to clarify some parts or aspects of the system being described while at the same time ignoring or distorting other parts or aspects of the system.

− The development of a model tends to be less like cycles in which a single interpretation system is expressed, tested, and revised, and more like the evolution of an ecological system which includes a variety of interacting organisms – where the processes that contribute to evolution include: (i) diversification, (ii) selection, (iii) communication, and (iv) accumulation.

Assumptions about the Continually Adapting Nature of Mathematical Models. Interpretation systems develop. They don’t go from unknown to learned in a single step. Furthermore, development tends to occur along a variety of interacting dimensions – such as: concrete-abstract, specific-general, preoperational-operational, intuitions-formalizations, global-analytic, external-internal, and unstable-stable. Yet, in a given situation, the most useful interpretation (or model) is not necessarily the one that is most abstract, most formal, most decontextualize, and so on. Yet, primitive interpretation systems often function similarly to Piaget’s preoperational stages of thinking. That is, they tend to be relatively barren (Piaget calls this “centering”), distorted (Piaget calls this “egocentrism”), and unstable (e.g., when attention focuses on the “forest”, pre-operational thinkers tend to lose cognizance of the “trees” - and vice versa). Yet, significant conceptual adaptations may occur (locally) within brief periods of time.

Assumptions about Beyond-Mathematical Aspects of Mathematical Interpretations. When learners or problem solvers interpret situations mathematically, they don’t simply engage logical mathematical systems, they also engage feelings, values, beliefs, dispositions, and a variety of metacognitive functions; and conversely, in a given situation, whether or not a feeling, value, belief, disposition, or metacognitive process is engaged depends on how the situation is interpreted. So, these latter attributes and processes develop as part of the models that learners or problem solvers’ develop - instead of being learned as isolated abstractions, and then connected to substantive mathematical concepts, dispositions, beliefs, attitudes, values, and feelings. Furthermore, instead of being fixed characteristics which act on students, an important part of competence is expected to involves learning to manipulate profiles of feelings, values, and dispositions to fit changing circumstances.

− All of the assumptions throughout this section could have been bulleted as being issues deserving further research; and, this is especially true when attention shifts from student-level understandings to teacher-level understandings. Also, if references were made to statistics-related thinking, the MMP research base has included many other topic areas. So, the statements that were made are not restricted to statistics.

3. Methodological Assumptions Underlying MMP Research

MMP considers mathematics education to be a “design science” (Lesh & Shriraman, 200x) in which most of the systems that researchers are trying to understand and explain also tend to be systems that they themselves are actively engaged in designing or developing. These systems include those that underlie curriculum materials or learning programs, but they also include the conceptual systems that students, teachers, developers, and others used to make sense of learning or problem solving situations. So, in the context of such systems, new ways of thinking lead to change, and change creates the need to develop new ways of thinking. As new understandings develop, new opportunities become apparent; new challenges arise; and, new changes need to be made. So, researchers tend to be inseparable parts of the systems they seek to understand and explain; and, methodological difficulties arise which are reminiscent of Heizenberg’s famous indeterminancy principle in physics. Not only is it true that researchers and developers tend to be integral parts of the systems they are seeking to understand, but, to assess the development of powerful interpretation systems, the kind of tasks that are most useful are those in which “subjects” need to interpret things in
ways so that important aspects of their interpretations can be observed directly. But, such thought-revealing activities tend to result in conceptual adaptations that enhance understandings of the interpretation systems being assessed. So again, documenting and assessing these understandings tends to change them in significant ways.

The preceding observations do not imply that systemic thinking is impossible to document or measure. In fact, in most modern sciences outside of education, systems with such properties are both documented and measured routinely. MMP research often investigates systemic thinking using multi-tier design studies of a type that has been described in detail in several recent books (Lesh & Kelly, 2000; Lesh, Kelly & Baec, 2007; English et. al., 2008). In such studies: (i) student development, teacher development, and the development knowledge by researchers or developers are recognized to interact, and (ii) model developers at all levels express their current ways of thinking in the form of purposeful and thought-revealing artifacts and tools whose development involves sequences of iterative and interacting cycles of testing and revising – and these processes automatically generate auditable trails of documentation in which important aspects of the students ways of thinking can be observed.

During student-level thought-revealing activities of the type used in MMP research, problem solvers typically make significant conceptual adaptations during sufficiently brief periods of time (e.g., 60-90 minutes) so that researchers can go beyond observing successive states of knowledge to directly observe processes that lead from one state to another. In other words, thought-revealing activities are similar to Petrie dishes in high school chemistry or biology laboratories – where significant changes occur during sufficiently brief periods of time so that processes of change can be observed directly.

The specific kind of thought-revealing activities that have been emphasized in MMP research are called model-eliciting activities (Lesh & Doerr, 2003); and again, a number of past publications have given detailed examples, design principles, and research methodologies for creating MEAs (Lesh, et. al. 2000; Lesh, Kelly & Yoon, 2007; Hjalmarson & Lesh, 2009). In general, MEAs are simulations of “real life” problem solving situations in which students clearly recognize the need to develop some specific type of mathematical model. The products that students produce usually are thought-revealing tools and artifacts in which important aspects of the problem solvers underlying ways of thinking are apparent; and, because the artifacts and tools need to be powerful (for some specific purpose), sharable (with other people), and reusable (in other situations), problem solvers themselves are able to assess the strengths and weaknesses of alternative ways of thinking. So, solution processes tend to involve sequences of iterative cycles in which students repeatedly express > test > revise their current ways of thinking. Students are able to proceed in directions that are progressively “better” without being guided toward a preconceived notion of “best.” This is possible because the problem statements are similar to design “specs” which are given to engineers or architects when they develop artifacts such as spacecraft to skyscrapers. These design “specs” typically include: (i) information about the resources that are to be use, (ii) issues that need to be addressed, and (iii) criteria that should be used to assess strengths and weaknesses of alternative products and ways of thinking. Therefore, such “specs” give direction to the products that engineers produce; and, they optimize the chances that desired developments will occur; but, they do not dictate the exact nature of what develops, nor do they dictate how development should be thought about or achieved. So, in many ways, these “specs” functions like photographic negatives of the products that are needed.

For the purposes of this paper, one main point to emphasize about model-eliciting activities (MEAs) is that they were NOT developed to be instructional “treatments” whose worth depends on their ability to induce significant learning gains in students. Instead, they were developed primarily to be research tools or assessment tools which were designed to help MMP researchers investigate the nature of the interpretation systems that students, teachers, developers, or researchers develop to make sense of situations that involve important kinds of mathematical thinking. Nonetheless, MEAs do in fact provide powerful tools to promote important types of learning - because their value as research tools depends heavily on optimizing the chances that significant kinds of conceptual adaptations (i.e., learning) will occur during brief periods of time (so that researchers can go be observing states of knowledge toward directly observing what enables students to evolve from one state to another). In fact:
Average ability students often make significant conceptual adaptations related to powerful mathematical concepts during 60-90 minute MEAs. In fact, they often develop impressively deep understandings of powerful concepts that their instructors were skeptical that they could be taught—even with large amounts of teacher guidance.

- The students whose model development work is most impressive in the preceding situations often do not have outstanding academic records on problem sets of the type emphasize in traditional textbooks and tests—and often come from highly underprivileged backgrounds.

- The understandings and abilities that develop during MEAs often exhibit extraordinary levels of transferability, adaptability, and durability.

- The students who perform surprisingly well often correlate closely with students that job interviewers most like to hire (Zawojewski, Diefes-Dux & Bowman, 2008, Lesh, Hamilton & Kaput, 2007).

4. What is an Example of a Model-Eliciting Activity?

The Paper Airplane Problem described below is an example of a model-eliciting activity (MEA). And, to emphasize the future-oriented marketability of the concepts and abilities that MEAs tend to emphasize, it is similar to a “case study” that MMP researchers first saw being used in Purdue University’s graduate program for aeronautical engineering—where a group of graduate students were in a “wind tunnel” where they were trying to develop more powerful ways of thinking (and quantifying) the concept of “drag” for various shapes of planes and wings.

To prepare students to work on the Paper Airplane Problem, homework assignments include reading a newspaper article which describe how to make a variety of different kinds of paper airplanes using one sheet of 8.6”x11” typewriter paper. The rules for making planes stipulate that planes cannot have any cuts, and cannot be made using tape, or paper clips, or any other such materials to weight the plane or hold its shape. Then, students are asked to bring data to class that could be useful for assessing how accurately their plane can fly. In particular, they are asked to measure how close they come to hitting a target that is at least 30 feet away. … One of the key things that students are expected to learn from this homework assignment is that many of their plane’s flight characteristics depend on how the plane is tossed—and by whom (because different people toss planes differently).

The Paper Airplane Problem

This photograph shows nine of the best paper airplanes from last year’s Paper Airplane Contest. Prizes were given for characteristics such as: most accurate, best floater, fanciest flier, and most creative design. Each of the planes must be made using a single sheet of 8.5”x11” piece of papers-with no tape, no paper clips, and now cuts.

Last year at the paper airplane contest, arguments arose about which planes really should have won several of the contests. These arguments arose because of three main reasons. (1) The differences often were not large between planes or pilots who where ranked 1st, 2nd, and 3rd. (2) Planes often flew quite differently when different "pilots" tossed them. Yet, nothing was done to control for which judges tossed the planes. (3) The judges weren’t explicit about the criteria they were using to rate planes. So, their decisions often didn’t match the intuitive judgments of most people who watched the contests. … For these reasons, next year, the judges want to have better and more explicit rules for judging planes for each award; and, as much as possible, they want to avoid having their judgements depend on vague opinions—or differences among judges who threw the planes. In fact, they would like to have a “formula” or some which uses objective measurements to calculate which plane is (for example) the most “accurate” flyer.

Your Task: Please write a letter to the judges for the Paper Airplane Contest. In your letter, describe some kind of a formula or rule that they can use to give awards for the paper airplane contest which will be held next week. Use the sample of data shown in the table below to show exactly how your formula works. Notice that each plane will be thrown three times by three “pilots”. Also notice that, even though this table of data only shows information about four planes that were in the contest last year, more than one hundred
planes are expected for the contest next week. The judges would like for your formula to help them judge: (a) which plane is most accurate, and (b) which pilot is most accurate. ... Your formula should have the characteristic that, when it is used, the planes that emerge as being ranked highest are generally consistent with intuitive judgements about the best planes. If formal rules don’t correspond with intuitive judgements, then people complain. But, they also complain if judgements aren’t made based on “objective” criteria.

As part of the problem statement for the Paper Airplane Problem this table and two graphs were given to each three-person team of students. This table and graphs show information from last year’s paper airplane contest. The table shows data about four different planes that were each thrown three times by three different pilots. The planes were thrown in a gymnasium - in a large square area that measured approximately 80’ on each side. The target was at the point marked (40,40) on the graph; and the planes were thrown from the lower left corner (0,0)

**Figure 2: Landing Points for Four Paper Airplanes**

Note: (i) The first graph above shows where each of the four planes landed for each of their nine flights. The second shows which pilot threw the plane that landed at each point. (ii) The graphs that the students were given measured 8 inches on each side. So, 1 inch on the students’ rulers corresponded to ten feet in the gymnasium where the planes were thrown.
5. A Research Topic where Problem Reformulation & Tool Development have been Important

Nowhere has tool development and the formulation of falsifiable conjectures been more clearly missing than in research on mathematical problem solving. For example, in the most recent Handbook of Research in Mathematics Education (Lester, 2007) which is published by the USA’s National Council of Teachers of Mathematics, Lesh’s and Zawojewski’s review of the literature on mathematical problem solving sited numerous examples where unproductive conjectures have been recycled again and again with new language (such as “habits of mind”) replacing old conjectures (such as “metacognitive processes”) without any significant changes in the underlying constructs. So, for these and other closely related issues, this review of the literature concluded that little progress has been made since Begel’s comprehensive review of the literature in 1979 concluded: “(T)here are enough indications that problem-solving strategies are both problem- and student-specific often enough to suggest that hopes of finding one (or a few strategies) which should be taught to all (or most students) are far too simplistic. (N)o clear cut directions for mathematics education are provided by the findings of these studies.” (p. 145).

Earlier, literature reviews by Silver (1985), Stanic & Kilpatrick (1989), Schoenfeld (1992), and Lester and Kehle (2003) support similar conclusions. The basic difficulty is clear.

− When researchers or developers focus on a small number of general processes, these processes tend to be more like names for large categories of processes rather than being well defined processes in themselves. They are too vague to be useful. (Schoenfeld, 1992).

− When researchers focus on longer lists of prescriptive processes, these list tend to become so lengthy that knowing when to use specific rules becomes the heart of understanding. (Lester & Kehle, 2003).

− Introducing the notion of metacognitive strategies (i.e., higher-level rules which operate on lower-order rules) simply moves this same basic problem to a higher level. And, the relationship between concept development and the development of problem solving processes is not clear.

− Even in those few studies in which some successful learning of problem solving strategies or metacognitive processes has been reported, transfer of learning has been undocumented or unimpressive (Lesh & Zawojewski, 2007); and, the causes of improved student performance may have little to do with the learning of general problem solving processes (Charles & Silver, 1989). For example, improvement could have occurred due to new mathematical concepts were learned – not because problem processes were learned.

In contrast to most traditional research on mathematical problem solving, MMP does not begin with the assumption that MMP researchers already know what it means to “understand” the most important “big ideas” for a given mathematics course or topic area - nor that they already understand what it means to understand related problem solving processes. In fact, MMP research has found that it must significantly revise many of the most basic assumptions that mathematics educators traditionally have associated with research on problem solving. For example:

− Whereas traditional research has characterized problem solving as a process of getting from givens to goals when the path is blocked (or where next steps are not apparent), MMP research has found it more productive to assume that: Problem Solving is a goal-directed activity which requires the problem solver to make significant adaptations to beginning interpretations of givens, goals, and possible solution processes. (Lesh & Zawojewski, 2007)

− Similarly, because MMP focuses on problems in which model-development, explanation-development, or design-development are important, MMP researchers expect solution processes to involve several parallel and iterative progressions of iterative cycles in which development from first-drafts to final-drafts generally seldom evolve along ladder-like linear sequences - nor like points moving linear or circular paths in space (Lesh & Harel, 2002). In particular, solutions should not be expected to evolve through sequences of modeling cycles of the type that we and others commonly used to describe model development.
− When MMP researchers look back at students’ 60-minute model development activities, students’ thinking often can be describe as evolving through a series of modeling-cycles that may look remarkably similar to Piaget-like accounts of the ways relevant concepts develop. On the other hand, when model-eliciting activities are used in which significant conceptual adaptations occur during 60-minute problem solving episodes, the tend to look nothing like steps up a ladder. In general, for the kind of problem solving that is emphasized in MMP research, organic models of solution processes (which envision the development of models to be similar to the development of an organisms in a complex ecosystem) appear to be far more appropriate than ballistic models (which envision the development of models to be similar to points moving along paths in space). During model development episodes, several poorly differentiated partial interpretations tend to gradually evolve along a variety of dimensions (concrete-abstract, situated-decontextualized, intuitions-generalizations, etc.). So, evolution often involves gradually integrating, differentiating, revising, or rejecting several competing partial interpretations as problem solvers go through a series of parallel and interacting cycles of expressing, testing, and revising their current ways of thinking.

− Whereas traditional research has characterized productive problem solving strategies and heuristics as answers to the question: What can I do when I’m stuck – with no ideas that appear to be useful? MMP has found that, in “real life” situations where mathematical thinking beyond school, problem solvers tend to be aware of a variety of relevant concepts and procedures – each focusing on different aspects of the situations, and all being at intermediate stages of development. So, MMP research recognizes that problem solving heuristics and strategies which provide answers to the question “What should I do when I’m stuck?” (i.e., when no ideas or procedures are apparent to guide next steps) may not be useful to provide answers to the question “How can I go beyond my current interpretation of givens, goals, and relevant processes? (i.e., when understandings of the relevant concepts and processes are at intermediate stages of development – not missing). In fact, MMP research has shown that drawing diagrams, clearly identifying givens and goals, or looking for similar problems lead to negative as well as positive outcomes. For example, they can lock students into current ways of thinking – instead of helping them develop beyond current ways of thinking. So, understanding when and why to use such strategies appears to be the heart of what it means to understanding them.

− MMP also recognizes that heuristics and strategies which are useful for describing the past behaviors of highly effective problem solvers may not be useful for prescribing next steps for inexperienced problem solvers. In MMP research, this has become apparent partly because participants in our studies often involve students, teachers, developers, or researchers who have expertise in fields ranging from dancing, to football or basketball, to tennis or to teaching – as well
as in academic fields ranging from education to engineering. And, in many such fields, videotape analyses of past performances are carefully analyzed to help performers develop powerful language to describe past decision-making activities. Yet, in general, no assumption is made that the goal of these reflection activities is to reduce decision-making to a rule-governed activity. One reason why this is true is that, during on-going performances in contexts like basketball, decision-making must occur far too rapidly and must involve far too many factors to be handled formally and analytically. So, one of the primary goals of reflection activities may be to develop more powerful ways to interpret relevant experiences.

Whereas traditional problem solving research has characterized solution processes as putting together concepts and processes that have been learned separately, MMP focuses on problem solving that involves differentiating, integrating, reorganizing, adapting, or extending interpretation systems that are at intermediate stages of development. In fact, in MMP research which has involved simulations of decision-making situations that reasonably might arise in the everyday lives of the relevant subjects, it has become clear that relevant mathematical knowledge is organized around experiences at least as much as around abstractions — and that working knowledge also is far more piecemeal, fragmented, fuzzy, situated and unstable than traditional theories have suggested. Even though the entire K-16 mathematics curriculum is largely restricted to situations that can be modeled using a single, solvable, and differentiable function, “real life” situations often involve feedback loops and second-order effects whose significance often far outweighs the significance of first-order effects. So, the situations cannot be modeled using single one-direction functions. In fact, useful models (descriptions or designs) often involve chunks of knowledge which integrate ways of thinking drawn from a variety of textbook topic areas. So, after relevant models have been developed, ideas that have been used usually need to be unpacked, named, and stored so that they will be retrievable when they are needed later.

6. Why Focus on Statistics or Data Modeling?

Future-oriented statements of mathematics curriculum goals often identify data modeling as a topic that should receive priority attention; and, for a wide range of students from elementary school through graduate school, many data modeling concepts are both accessible and empowering. One reason why this is true is because many of these concepts and abilities involve straightforward extensions of basic ideas in elementary mathematics, and because both the power and the accessibility of these concepts and abilities have been strongly impacted by the development of new technologies for conceptualization, design, and communication. However, as we enter the 21st century, these same new technologies also are inducing significant changes in: (i) the kind of problem solving situations in which some type of “mathematical thinking” is needed for success beyond school (Lesh, Caylor, & Gupta, 2007; Lesh & Caylor), (ii) the levels and types of “mathematical thinking” that are needed (Lesh, Hamilton, & Kaput, 2007), and (iii) the ways that these concepts and abilities can be thought about, learned, documented, and assessed (Lesh & Lamon, 1992). Yet, when we look at statistics courses such as those that are taught to graduate students in education (or the cognitive sciences, learning sciences, life sciences, social sciences, or environmental sciences), then it is clear that there is perhaps no other course in the entire K-20 mathematics curriculum where such an overwhelming majority of bright students emerge with only the most superficial “cook book” conceptions of relevant concepts and procedures.

7. What Kind of Tool Development is Emphasized in MMP Research?

Even though the tools that are described in this paper focus on statistics, most of these tools can be adapted in straightforward ways for research about other topic areas. Such tools include the following.

− Model-Eliciting Activities (MEAs) are aimed at clarifying what it means to “understand” the most important concepts and procedures that should be emphasized in an introductory statistics course for beginning doctoral students in education,
Model-Exploration Activities (MXAs) are aimed at investigating “What kind of development still remains be achieved after students have participated in MEAs,

Concept-Analysis Wheels are used to help document students’ reflections about important concepts that they believe that they have used in their responses to MEAs. And, they also are used to help document how teachers are thinking about students’ thinking: (i) while they are working on MEAs, (ii) when they are making oral reports during “poster sessions” reporting the results of their work on MEAs, or (iii) when feedback needs to be given about strengths and weaknesses of students written responses to MEAs,

Reflection Tools are used to help document students’ development concerning: (i) levels of engagement during MEAs, (ii) group functioning during MEAs, (iii) individual roles played during MEAs, or (iv) metacognitive beliefs, dispositions, or processes that influence students’ effectiveness during MEAs.

Model Development Sequences are designed to be completed in a small number of 90-minute class periods; and, they are modularized to be easily modifiable to support a variety of research purposes.

8. What is an Example of a Model Development Sequence?

In MMP research, many types of model-development sequences and model-development units have been used for different research purposes. But, the following schematic diagram shows one typical example. … Notice that an important characteristic of model-development units is that they are highly modular – so that individual components can be added, subtracted, modified, or rearranged for a variety of purposes. For example, one recent study demonstrated that:

- The models that students develop during MEAs tend to be significantly different depending whether: (i) they begin by engaging in an MEA before working on the model-exploration activity (MXA) which served to extend or empower the models that they themselves had developed, or (ii) they begin by working on the MXA activities before working on the MEA which served as an application of ideas and procedures introduced in the MXA. In general, MEA-to-MXA understandings are far more richly connected to topics associated with a variety of textbook topic areas; they are far better grounded in students’ personal experiences; and, they were far more resistant to forgetting.
Two additional points are important to emphasize about the preceding kinds of learning and problem solving units. First, just like the MEAs and other MMP model-development activities, model development units are NOT designed to be ideal curriculum materials to optimize learning. Instead, they are designed to provide useful contexts for researchers to use to investigate the development of important aspects of students’ mathematical thinking. Therefore, like MEAs, they provide direction to students without depending on large amounts of teacher guidance; and, they do this in much the same way that, in engineering, good design specs ensure that clients get what they need without telling the engineer how to think and what to do. Instead of guiding students in desired directions, they get students to recognize the need to develop in desired directions. Second, even though MEAs are the most widely known components of MMP research, students’ model development activities do not stop with MEAs. So, other components of model development sequences are designed to investigate important types of model develop which tend to occur after students have completed MEAs.

Notice that the preceding model development units involve only three 90-minute classroom activities. One reason this is true is because MMP research is intended to observe significant conceptual adaptations over periods of time which are sufficiently brief so that most of what happens takes place in contexts that cannot be observed. But, another reason is to debunk the notion that inquiry-oriented learning is necessarily much more time consuming than traditional teacher directed instruction. For example, in one recent study, MMP researchers compared what was taught and learned in traditionally taught statistics classes versus the understandings that developed in a 15-week course consisting of ten model-development units like the one depicted in Figure 4 (Lesh, Carmona & Moore, 2010). Results:

− The modeling group not only “covered” far more topics, but they also outscored the traditionally-taught students on the half of the final examination that was designed explicitly for the traditionally-taught class.

− The traditionally-taught students scored close to zero on the half of the final examination which was designed to show how well students understood that every computational procedure in statistics presupposes a model which makes assumptions that may or may not be appropriate for the situation being investigated.

− In interviews that were conducted one year after the preceding course ended, the MEA-to-MXA students often were able to give impressively detailed, accurate, and insightful descriptions of concepts and activities that had been emphasized in their course. Whereas, among the traditionally-taught students, recollections about what they presumably had learned were nearly non-existent. Also, the traditionally taught students made impressively more computational errors and conceptual misstates on a brief review of basic skills from the course.

An advocate of direct instruction might ask: How could this be? If you know what you want students to learn, why don’t you just tell them the correct ways to think? Then test them to make sure they get it. … An MMP-based response might be. How successful do you think you’d be if you tried to teach a child how to ride a bicycle by telling them the correct way to do it? Is it possible to teach something correctly without teaching it meaningfully? … Other responses should include the following observations.

− When students are engaged in the kind of model-development units described here, far more students than usual tend to be intensely engaged for far longer periods of time; and, by working in
teams in which many alternative ways of thinking are expressed, tested, and revised, students tend to make far more connections among ideas in different textbook topic areas, and they also tend to make far more connections to their own real life experiences? And, they also weed out misconceptions at the same time that they are developing productive ways of thinking.

Model-Eliciting Activities – To Change Students’ Thinking, You Must First Engage Student’s Thinking.

This section briefly describes responses that students in one class gave to the Paper Airplane Problem. The students referred to were thirty elementary education majors who were enrolled in a mathematics course for students majoring in elementary education during their sophomore year at Indiana University. For all MEAs, the students were randomly divided into groups of three students each; and, for the students whose work is described here, the responses were given during the fourth week of the course – after they had participated in two previous model-development units. The first unit was about quantifying qualitative information and about transforming data into forms suitable for comparing or aggregating data of different types. The second was about developing useful ways to “operationally define” constructs which cannot be measured directly. Examples of these operationally defined constructs included what it means to be a “productive” worker (using information about dollars earned and time spent across several distinct periods of time) and what it means to be a student who deserves an A, B, C, D, or F” (using information such as performance on quizzes, tests, projects, and homework). … In these activities, the students also became familiar with data analysis software such as Excel, Fathom, or Tinkerplots to create graphs and to calculate weighted sums or weighted averages involving several kinds of quantities.

Because of prior experiences during the first two weeks of their course, the students naturally thought of the Paper Airplane Problem as being about developing an operational definition about what it means to be an “accurate” paper airplane. So, the list below briefly describes several alternative first-iteration ways that different teams thought about the problem. When describing these ways of thinking, however, it is important to reiterate that, in situations like the Paper Airplane Problem, early interpretations tend to involve several loosely related, unstable, partial interpretations – not one single and internally consistent interpretation. So, the generalizations that follow are too simplistic; and, transitions from first-draft, to second-draft, third-draft, and Nth-draft ways of thinking would have been unlikely to occur unless several alternative ways of thinking co-existed and were developing interactively and in parallel.

− Groups G1 and G2 began by using colored marking pens to mark the landing points for each plane on the paper graphs, and they made intuitive judgments about which plane they thought should win – before doing any detailed measurements or calculations.

− Group G3 did essentially the same thing as groups G1 and G2 but they did it using Tinkerplots versions of the paper graphs that were given.

− Groups G4, G5, and G6 began their assessment of each plane’s accuracy by drawing concentric rings around the target (like a dart board as shown in Figure 15 later in this paper) on the two paper graphs that were given in the problem statement.

− Groups G7 and G8 began by using a ruler to measure distances from the target to the landing points on the paper graphs. Then, these distances were entered into the table that was given in the statement of the problem.

− Groups G9 and G10 focused mainly on the table of data and used the Pythagorean Theorem \( A^2 + B^2 = C^2 \) to calculate the distance from the target (0,0) to each landing point.
One main difference among 1st interpretations of the *Paper Airplane Problem* involved the extent to which students focused on details about individual flights rather than focusing on patterns of landings for different planes or pilots. The distinction is easy to see if graphs are plotted for each of the four planes – as shown in Figure 5. Students themselves did not create such graphs during the first fifteen minutes when they started working on the *Paper Airplane Problem*. But, most groups created such graphs during later interpretations.

For their first interpretations of the *Paper Airplane Problem*, groups G1-G6 focused mostly on landing patterns for each plane; and, G7-G10 focused mostly on details for each flight. Whereas, during second and third interpretations, all groups tended to shift attention in the opposite directions than in their first interpretations. So, groups that began by focusing on details about individual flights gradually shifted their attention toward aggregating this information for each plane; and, groups that began by focusing on general landing patterns for each plane gradually shifted attention toward details about each flight. Furthermore, in both cases, all of the groups gradually become aware of mismatches between results of calculations and intuitive judgments about which planes were “most accurate” flyers. For example, for groups that began by focusing on the sums of the distances from the target, planes with the lowest sums often were not the planes that the students’ intuitive judgments considered to be the “most accurate” fliers. So, for similar reasons, all of the groups gradually became aware of two distinct kinds of factors that where involved in their intuitive judgments about “most accurate” planes. The first factor had to do with the “spread” of the landing points for each plane; and, the second had to do with the distance from the target to the “center” of the landing points for each plane. So, for all ten groups, final operational definitions of what it means to be an “accurate” plane involved four issues. (i) finding a “center” of the landing patterns for each of the planes. (ii) finding a way to measure the distance from this “center” to the target for each plan. (iii) finding a way to combine the preceding two measures into a single measure of accuracy. (note: Steps “i”, “ii” and “iii” usually interacted.)

Five of the ten groups ended up using mean values to operationally define the “centers” of the landing points for each plane. But, five groups used a “center” that minimized the sum of the straight-line distances to all of the landing points for each plane (rather than the sum of the squares of the distances). (note: Mean values minimize the sum of squares of distances – not the sum of straight-line distances.) ... Because of students’ experiences during the first two weeks of the course, they knew how to use software such as Excel, Fathom, or Tinkerplots to easily find either a least distance point or a least square point for a collections of data. So, after finding “center” for each plane’s landing pattern,
all ten groups used some type of weighted sum to combine: (i) their measure of spread, and (ii) their measure of the distance from the target to the “center” of the landing points for each plane. Intermediate interpretations involved alternative ways of resolving (or ignoring) the preceding three issues. And, all groups went through at least 3-4 distinct interpretation cycles during the process of developing final responses to the problem. Different interpretations focused on distinct types of mathematical characteristics or quantities, and/or different quantities, relationships, or operations.

**Poster Sessions – An Example where Model-Development Sequences treat Classroom Time as Precious.**

In model-development units, the main criteria for deciding whether an activity should be done in class or out if class is whether the activity provides a rich site for researchers to directly observe students’ thinking; and, past experiences in MMP research suggest that the richest activities for this purpose tend to be those that emphasize teacher-student interactions or student-student interactions. So, activities such as oral reports from individual groups, where only small percentages of students tend to be active participants, often are relegated to outside-of-class activities. On the other hand, poster sessions similar to those that occur at educational research conferences, have proven to be far less time-consuming and far more rich and intense versions of the kind of one-at-a-time oral presentations of students’ work that many teachers like to use at the end of MEAs. Even though oral presentations can be both thought-revealing and thought-revising activities, poster sessions usually can be conducted in a way so that most of the beneficial characteristics of oral presentations are not lost – and less time is required. But, even if poster sessions are used, it is important for them to be recognized having a productive purpose. For example, when poster sessions are used in model-development sequences, MMP researchers often randomly assign one student in each group to be the “presenter” while the remaining two students serve as “rovers” whose jobs are to gather as much information as possible from other groups. Then, each group is allowed to produce one more iteration in their solutions to problems – in which each group can adopt new ideas that they learn from other groups, and both N-1st and Nth responses (as well as poster presentations) count toward grades in the course.

**Self-Assessment Activities - Require Students to Reflect about Important Aspects of their Past Work.**

The ability to self-assess is important during the solution of MEAs. Otherwise, students wouldn’t recognize the need to go beyond first-draft or intermediate-draft responses; and, they also wouldn’t be able to assess the strengths and weaknesses of alternative ways of thinking – or to make judgments about when their current ways of thinking need to be modified, rejected, or accepted as being good enough to the client’s purposes. But, even beyond these on-going self-assessments, it sometimes is both useful and illuminating for students, and/or teachers, and/or researchers to play the role of “clients” during in-class activities in which results of MEAs are evaluated. But, in these “outside assessments” of students’ work, it is important to avoid introducing scoring rubrics which “change the rules of the game” by assessing students’ work using criteria that were not given (explicitly or implicitly) in the statement of the MEA. Otherwise, students soon conclude that stated problems are not the “real” problems – and will not take future problem solving statements seriously.

One of the most important design principles for MEAs is that solutions which are “most practical” also need to be solutions which involve the kind of mathematical thinking that MMP researchers are trying to investigate. And, perceptions of “political correctness” generally are not what MMP researchers investigate. On the other hand, sharability and reusability of tools and artifacts that students produce during MEAs nearly always are among the most important criteria which are emphasized implicitly or explicitly in MEA problem statements. Therefore, the Quality Assessment Guide that is shown in Figure 6 usually works for all MEAs – because it simply returns to criteria that were given in the statement of the problem. Consequently, to the greatest extent possible, all criteria for assessing the results of MEAs should be clear in the problem statements.

This Quality Assurance Guide is designed to help teachers (and students) evaluate the products that are developed in response to model-eliciting activities with the following characteristics: (a) the goal is to develop a conceptual tool, (b) the client who needs the tool is clearly identified, (c) the client’s purposes are known, and (d) the tool must be sharable with other people and must be useful in situations where the data
are different than those specified in the problem.

<table>
<thead>
<tr>
<th>Performance Level</th>
<th>How useful is the product?</th>
<th>What questions should be asked?</th>
<th>What might the client say?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requires Redirection</td>
<td>The product is on the wrong track. Working longer or harder won’t work. The students may require some additional feedback from the teacher.</td>
<td>To assess students’ work, put yourself in the role of the client. To do this, it’s necessary to be clear about answers to the following questions. 1. Who is the client? 2. What conceptual tool does the client need? 3. What does the client need to be able to do with the tool?</td>
<td>“This is impossible to understand. Or: “Start over. This won’t work. Think about it differently. Use different ideas or procedures.”</td>
</tr>
<tr>
<td>Requires Major Extensions or Refinements</td>
<td>The product is a good start toward meeting the client’s needs, but a lot more work is needed to respond to all of the issues.</td>
<td>Then, the quality of students’ work can be determined by focusing on the question – How useful is the tool for the purposes of the client? To assess usefulness, and to identify strengths and weaknesses of different results that students produce, it would be helpful to consider the following questions. 1. What information, relationships, and patterns does the tool take into account? 2. Were appropriate ideas and procedures chosen for dealing with this information? 3. Were any technical errors made in using the preceding ideas and procedures? But, the central question is - Does the product meet the client’s needs?</td>
<td>“This is close to what I need. You just need to add or change a few small things for it to be useful.”</td>
</tr>
<tr>
<td>Requires Only Minor Editing</td>
<td>The product is nearly ready to be used. It still needs a few small modifications, additions, or refinements.</td>
<td>The product should make it clear that: • The students went beyond producing a tool that they themselves could use to also produce a tool that others could use – by including needed explanations, and by making it as simple, clear, and well organized as possible. • The students went beyond thinking with the tool to also think about it – by identifying underlying assumptions (so that others know when the tool might need to be modified for use in similar situations). • The students went beyond blind thinking to also think about their thinking – by recognizing strength and weaknesses of their approach compared with other possible alternatives.</td>
<td>“Excellent, this tool will be easy for me to use again in future situations – and it will be easy to modify for slightly different situations.”</td>
</tr>
<tr>
<td>Useful for this Specific Data Given</td>
<td>No changes will be needed to meets the immediate needs of the client.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharable and Reusable</td>
<td>The tool not only works for the immediate situation, but it also would be easy for others to modify and use it in similar situations.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 6:** An Example of an MEA Quality Assessment Guide

MMP research suggests that the ability to self-assess are especially important when the goal of research is to identify students whose capabilities seldom are assessed using traditional school tests and tasks. Bright-but-disenchantment students tend to be especially turned off by: (i) tasks which tell students that one thing is wanted but then score responses based on other things that students were not told, and (ii) tasks in which “school answers” and not “smart answer” outside of school. Conversely, such students often are clearly empowered by activities in which they themselves are able to make decisions that are valued.

*Concept Analysis Wheels (or Ways of Thinking Sheets)* – Meta-Level MEAs for Students & Teachers: *Concept analysis wheels*, like the one shown in Figure 7, have proven to provide useful sites for investigating students’ and/or teachers’ meta-aware awareness of the concepts and procedures that students use during solutions to MEAs. Furthermore, when such tools are used multiple times throughout a course, MMP research has shown that both students and teachers often become significantly more insightful about changes in thinking during solutions to MEAs.
The general idea of the concept analysis wheel was developed by Carmona (200x); but, the specific version shown in Figure 7 was developed jointly by the students and teachers who participated in the study described here. Sometimes these concept analysis sheets look more like matrices than wheels; and sometimes, they look more like networks or “concept maps” of the type described by Novak and others. Carmona borrowed the notion of a concept analysis wheel from the kind of “flavor wheels” that are used to teach people to become wine tasting connoisseurs. Consequently, it may not be surprising that they also have proven to be useful for helping researchers, research assistants, teachers, and students to become connoisseurs of students’ thinking. Such reflection tools often contribute to learning at the same time that they are documenting or assessing learning. That is, they are self-fulfilling assessments. So, in MMP research, students and teachers often are cast in the roles of co-developers of “concept analysis wheels” that can be used for a variety of purposes – some of which as similar in function to Japanese lesson plan activities – except they don’t take teachers away from teaching by treating them as curriculum developers. Instead, in MMP reflection activities, teacher-level tasks are emphasized which focus is on thinking about student thinking. So, especially projective tasks often involve having teachers work in teams of three to co-develop sharable and reusable “concept analysis wheels” or “ways of thinking sheets” which can be used for a variety of purposes – such as making insightful observations about students’ thinking: (i) while students are working on thought-revealing activities such as MEAs, (ii) while students are making oral reports about their responses to MEAs, (iii) when feedback is given to students about strengths and weaknesses of their responses to MEAs. … A variety of different kinds of teacher-level reflection tools and teacher-level MEAs have been described by Zawojewski, Chamberlain, Hjalmarsen & Lewis, (2007) and by Lesh & Zawojewski (2007).
In MMP research, the development of *concept analysis wheels* has proven to be a very effective example of a tool whose development often treats students and/or teachers as co-developers. So, even if MMP researchers believe that they themselves already understand the nature of a high quality *concept analysis wheel* for a given course, engaging students and/or teachers in their development often is a powerful *thought-revealing task* – as long as development has a clear purpose (so that it is possible for developers to self-assess the power, sharability, and reusability of products that they produce). For example, in the mathematics course for elementary education majors which has been referred to throughout this paper, the *concept analysis wheel* that was used at the start of the course highlighted lofty-sounding “big ideas” that sounded like chapter headings in traditional statistics textbooks. But gradually:

- As teachers, research assistants, and researchers repeatedly tried to make sense of students work (and plan for what teachers should do next based on observations of students’ work) the *concept analysis wheel* shifted away from highlighting “capstones” of the course – and shifted toward highlighting “cornerstones” of the course.

The preceding cornerstones came to be thought of as reoccurring themes that were woven throughout virtually every MEA and model-development sequence in the course; and, this view in turn reflected dramatic shifts in how the teachers and researchers thought about prerequisites – and the extent and ways in which some ideas and abilities were prerequisites to others. Concepts or topics which, at the start of the course, had been thought of as being capstones increasingly came to be thought of as cornerstones whose meanings evolved dramatically throughout the course – as they were integrated in different ways into the results of nearly every MEA in the course.

Past researchers and developers have used the term “spiral curriculum” to refer to something similar to this issue of reoccurring themes. But, because the results of MEAs tend to be models which integrate concepts and procedures associated with a variety of “cornerstones” (which then need to be unpacked), the notion of revisiting a single idea repeatedly takes on meanings that gives new meaning to the notion of a spiral curriculum.

**Reflection Activities – Focusing on Beliefs, Dispositions, Attitudes & Meta-Level Thinking:**

When used by students, the kind of *concept analysis wheels* shown in Figure 7 also can be used to set the stage for a variety of other kinds of reflection activities – such as those that focus on changes in levels of engagement, group functioning, roles of individuals within groups, or a variety of metacognitive processes, beliefs, attitudes, dispositions, values, or feelings. This is because one distinctive characteristics of MMP research is that all of these beyond-mathematical aspects of thinking are important parts of the models that students develop to make sense of their experiences.

- When students interpret situations mathematically, MMP research expects that they don’t simply engage interpretation systems that are purely logical or mathematical in nature. Interpretations also involve attitudes, values, beliefs, dispositions, or metacognitive processes.

- MMP does not treat these attributes as if they were fixed characteristics which determine students’ behaviors. Instead, students are expected to develop profiles of variable attributes which students themselves can learn to manipulate to fit circumstances which are continually changing.

How can *concept analysis wheels* be used to investigate the development of students’ or teachers’ beliefs, dispositions, attributes, or problem solving processes? For the mathematics course described here, one way that *concept analysis wheels* can be useful is to begin by having students “colored in” (or in some other way reflected about) concepts and processes that they used during MEAs that they had completed (Lesh, Carmona & Moore, 2010; Zawojewski, et.al., 2007) For example, after students complete work on MEAs, three-person teams can be asked to draw a “step diagram” (see Figure 5) to show “changes in thinking” that their group went through during the solution of a given MEA. For our purposes, it generally was not necessary for these “step diagrams” that students draw to closely fit the interpretation cycles that researchers later developed to describe stages in thinking that they believe students’ actually went through - based on detailed analysis of videotapes and/or transcripts of the sessions. To promote productive reflections and discussions about problem solving sessions, it was only necessary for the students in each three-person team to reach a consensus about the location of several “benchmarks” which they could then use to ground their reflections about shifts in their thinking during sessions.
Richard Lesh

The second part of such reflection activities asks each student to mark a half dozen “significant points” on each step diagram to show where they believe that “something significant happened” concerning whatever attribute or process is being investigated - such as their level of engagement, the quality of their group functioning, or individual problem solving behaviors, beliefs, attitudes, dispositions, or problem solving processes. For example, to investigate students’ perceptions of “levels of engagement” (i.e., Csikszentmihaly’s concept of FLOW, 1990), immediately after finishing an MEA, and after reaching a consensus about solution steps and “significant points” on these steps, students can be asked to quickly responded to the questionnaire shown below for each of the “significant points” that they identified. Then, this information about locations and levels of engagement can be used as reference points to stimulate reflections and discussions about the problem solving session that they had just completed.

When I am in a high state of FLOW …

- Time seems to be passing rapidly because I’m enjoying what I’m doing.
- Thing seem to be going in slow motion because I see patterns. So, everything fits.
- I’m really on top of things. So, I’m am not feeling overwhelmed by details.
- I’m so totally engaged that I’m not distracted by other things.
- I’m not at all frustrated because I can see exactly what’s needed to reach my goals.
- I’m really cruising along, and everything is going smoothly.
- I feel like I’m really making progress and accomplishing something that’s not easy - but worthwhile.
- I have a good sense of how everybody in my group is thinking – and how everything is fitting together.

Figure 8: Indicators for High Levels of FLOW

1. Very Low Low High Very High
2. Very Low Low High Very High
3. Very Low Low High Very High
4. Very Low Low High Very High
5. Very Low Low High Very High
6. Very Low Low High Very High
7. Very Low Low High Very High
8. Very Low Low High Very High

Figure 9: My Level of FLOW at Different Points During a Problem Solving Session

For the mathematics class referred to throughout this paper, Figure 10a shows the average height of students’ self-assessed first step, second step, next to last step, and last step interpretations of steps that they took during solutions to MEAs. Then, Figure 10b shows a simplified version of how students’ self assessments of “levels of engagement” were recorded at “significant points” during the first two steps and the last two steps that the students identified during their solutions to all MEAs where this reflection tool was used during in the entire course. Then, Figures 10c and 10d show how these data were aggregated across several MEA activities. … Looking across these graphs, several observations are apparent. For example, in Figure 10c, students’ assessments of FLOW generally are positive and remain approximately the same from the beginning to the end of each step; and, in fact, changes are small even when students move from one step to the next. Whereas, in Figure 10d, assessments of FLOW generally go from positive at the start of each step to negative at the end of each step. So, feelings of FLOW are highest immediately after students shift to new ways of thinking, and they are lowest as students recognize the need to shift to a new way of thinking – or as they notice shortcomings of current ways of thinking.
The preceding observations also were supported by anecdotal information. That is, at the start of the course, students tended to report liking or disliking activities as a whole – with little change in these assessments throughout solution attempts. But, one reason for this lack of change was that students were not very aware of changes in their own thinking during solution processes. So, after completing 3-6 MEAs, and after using concept analysis wheels and other reflection tools to think back about progress that was made during each MEA, students tended to become much more aware of changes in their own thinking during solutions to MEAs. And, accompanying this growing awareness of changes in thinking came corresponding growing awareness of changes in feelings of FLOW (or other attributes) throughout MEA session. For example, the students began to think of feelings of FLOW as a continually changing characteristic. Furthermore, rather than thinking of FLOW as an attribute that impacted them, they started to think of FLOW as something that they themselves could manipulate to their own advantage. In particular, they became aware that feelings of discomfort often can be expected during solutions to problems – and that they often occur precisely at points in time when changes in thinking are needed. So, recognizing this fact tended to make feelings of discomfort more tolerable – so that some students and groups actually welcomed feelings of discomfort because
they recognized that these feeling tended to be short lived and often led to new ways of thinking. Similarly, as students’ thinking shifted to new levels of thinking, feelings of FLOW tended to accompany newfound abilities to see patterns where they used to see only disjointed pieces of information. Yet, because the interpretation systems that students engaged were tied to past experiences more than to abstractions, their abilities to manipulate feelings of FLOW also depended on the power of past experiences to facilitate productive interpretations.

For more details about studies of this type, see Lesh, Carmona & Moore (2010). But, for the purposes of this paper, the main points to emphasize are that: (i) similar studies are straightforward to conduct which focus on many different kinds of beliefs, dispositions, attributes, or processes, and (ii) so far in MMP research, the meanings of these attributes have tended to be quite different than they have been portrayed in tradition research on learning and problem solving. For example, in the case of Csikszentmihalyi’s feelings related to FLOW (1990):

- Students can manipulate FLOW as much as feelings of FLOW impact them.
- How feelings of FLOW are engaged depends on how situations are interpreted.
- Exceptionally proficient problem solvers do not have fixed profiles of attributes such as FLOW.
- As students increased their problem solving proficiency, they developed sophisticated profiles of attributes which they manipulated in ways that were useful in continually changing situations.
- In particular, students increasingly manipulated their own profiles of attributes from one stage of model development to another during MEAs.
- Developing productive attributes related to FLOW appears to be closely related to the development of powerful models for interpreting experiences. And, MMP research hypothesizes that similar claims will prove to be true for other kinds of beliefs, dispositions, and problem solving processes.

Homework – Model Development Sequences treat Time in Class as Precious.

In model-development sequences, any types of learning that are not thought-revealing tend to be off-loaded (if possible) to homework activities where it is not important to be able to directly observe students’ thinking. For example, in the case of the Paper Airplane Problem, three types of homework assignments were used. First, students were given a transcript from past MMP research which gave details about how “another group of students” had solved the Paper Airplane Problem. Then, students used their current concept analysis wheel to reflect about their own work and to also identify “shifts in thinking” that were apparent in the transcript. Second, brief explanations and practice exercises were given which focused on skills that the students “might have found to be useful”) in their solutions to the Paper Airplane Problem. The most important characteristics of these skills are: (i) Students should clearly recognize them as being useful in future situations that they judged to be similar to the just-completed MEA – or to similar situations in their lives beyond school. (ii) Students should become familiar with ways to visualize processes and products of using these skills. For example, following the Paper Airplane Problem, the homework assignment included a few quick exercises (presented in TinkerPlots or Fathom) to remind students about important issues related to alternative ways of measuring centrality and spread.

First, answer this question using MEANS as “Centers” and then Calculate the Sum of Squares of Distances to the Other Points.

Second, answer this question using LEAST DISTANCE POINTS as “Centers” and then Calculate the Sums of the Distances to the Other Points.

Figure 11: Which Collection of Points below has the Smallest Spread?
Note: The same question that is posed in Figure 11 also were posed for points on a line (where MEDIANS coincide with Least Distance Points). One reason why this is important is because, for points on a line, it is more obvious how to solve the problem that a unique MEDIAN sometimes does not exist – such as in cases where there is an even number of points. For example, if additional restrictions also are included (such as minimizing the difference between points on the left and on the right) then unique values can be determined.

This second version of the Paper Airplane Problem also was different because the data were generated using the kind of TinkerPlots samplers shown in Figure 13. Then, the goal of the task was: Take as many samples as you like to determine whether you believe these four paper airplanes are really different from one another – or if apparent differences are simply due to luck (or unexplained sources of variation that are not recorded in the data). (Note: When this problem was given, the
graphs in the samplers were hidden. So essentially, the students’ goal for the task was to take samples to try to figure out whether the samplers (i.e., the graphs) for the four planes were significantly different from one another.

Figure 13: A Repeatable (& Easily Constructable) TinkerPlots Simulation for the Paper Airplane Problem

In Figure 13, the sampler that is given on the left side of the simulation chooses one of three “pilots” to throw a paper plane. Then: (i) For each of the three samplers in the second column of the simulation, bar graph samplers are used to give each of the three pilots a different profile of characteristics (e.g., tendencies to throw to the left or right, or short or long. (ii) The fourth column assigns paper airplanes to each pilot for each flight; and again, columns 4 and 5 give each plane a profile of characteristic such as tendencies to fly short or long, or left or right. (iii) The seventh column assigns a flight number to each plane and pilot.

Notice that this second version of the Paper Airplane Problem is different than the first in another very significant way. In the first version, statistics are used to describe past events (Which plane should be declared the winner?); whereas, in the second version, statistics are used to make inferences about future events (What is likely to happen if we continue to toss the planes a larger number of times?) In particular, information drawn from a sample is used to make inferences about an infinite population of possibilities.

Each of the three samplers below could be used to act out the same experiment: A sample of twelve items is selected (with or without replacement) and the probability of selecting an “a” is twice the probability of selecting “b” or “c”.

The sampler shown on the right would generate the following list: (a,b,c,a,b,c,a,b,c,a,b,c)
For the *concept analysis wheel* that is shown in Figure 7, notice that both descriptive statistics and inferential statistics are given equal attention. This balance between descriptive statistics and inferential statistics is not usual in traditional textbooks – where little attention tends to be given to aspects of “statistics understanding” that are related to: (i) formulating sensible operational definitions of attributes associated with collections of data, or (ii) alternative ways to collect or mathematize data – e.g., by organizing it, dimensionalizing it, coordinatizing it, or quantifying it. Yet, in MMP research, these latter kinds of understandings have emerged as being primary sources of sources of bias in research that uses statistical reasoning. Furthermore, such issues appear to be primary sources of misconceptions when students who complete introductory statistics class emerge with only exceedingly impoverished understandings about the fact that every computational procedure in statistics presupposes assumptions about the situation being described. So, according to MMP research, one of the primary questions that students should learn to answer is: *Do assumptions about distances, centrality, spread – or about proportions, patterns, relations, and interactions – fit the situation that is being described?* Or, when attention shifts from descriptive to inferential uses of statistics: *What is the relationship between statistics about samples and statistics about the populations from which samples are collected?*

Model-Exploration Activities (MXAs) – Investigating Development after MEAs are Completed

For MMP research purposes, one of the main goals of MXAs is to investigate the nature of levels and types of understandings students still need to develop after they have completed MEAs. This is important because, even though students responses to MEAs often are impressive, many of the strengths of MEA-stimulated models also have down sides. For example:

- One important strength of MEA-influenced models that students develop is that their meanings tend to be highly situated and enriched by a variety of “real life” experiences. But, even though many aspect of thinking probably should continue to be organized around experience as much as around abstractions, MMP research also expects that a certain amount of decontextualization may be important for students to develop; and, if names are not given to important constructs that students develop, then they tend to be quickly forgotten. … So, MXAs often begin with brief teacher-led summaries comparing strengths and weaknesses of alternative ways that different groups of students solved (or could have solved) the relevant MEA (e.g., the *Paper Airplane Problem*).

- Another strength of MEA-influenced models is that they meanings tend to integrate ideas and procedures drawn from a variety of textbook topic areas. But, even though the connected nature of these chunks of knowledge is important, MMP research has found that it is be important for students to unpack their models - so that connections among ideas are apparent which advance formal understandings and abilities. This is another reason why MXA’s often begin by using a *concept analysis wheel* to give a class-level summary of concepts and processes that the class as a whole used to solve the *Paper Airplane Problem*.

- Another characteristic of MEA-influenced models is that students often use a variety of expressive media (e.g., spoken language, written notations, diagrams, pictures, graphs, concrete models or experience-based metaphors) - each of which clarifies some aspects of the problem solving
situations, but ignores or distorts other aspects. For this reason, MMP research anticipates that it may be important for students to gain access to powerful ways of visualizing and expressing the interpretation systems that they develop. For example, the kinds of graphs given in Figures XX, XX, XX, and XX, represent a graphic method of analyzing sources of variation in the Paper Airplane Problem. In other words, they represent a graphic form of analysis of variance. So, MXAs often begin with a few brief quiz questions or teacher-led exercises designed to focus on a ways to visualize a variety of ways of thinking that would have been useful in the MEA being discussed.

Another characteristic of MEA-influenced models is that they often enable students to “see” patterns and relationships which they previously had not recognized. But, students’ ways of thinking often function more like transparent “windows” through which they view experiences – rather than being “objects” which they also look at, and examine, and analyze. For this reason, MMP research investigates levels and types of understanding that are involved when students explicitly analyze and compare alternative models to identify strength and weaknesses. ... So, MXAs often introduce elegant graphs or other tools which clarify important distinctions among various ways of thinking about the relevant MEA. For example, in the case of the Paper Airplane Problem, a series of brief activities were presented to help students clarify the following kinds of issues.

- When alternative solutions to the Paper Plane Problem involve different ways of measuring “centrality” and/or “spread”, different planes tend to end up being ranked differently (e.g., concerning “accuracy” or paper planes).
- When alternative solutions to the Paper Plane Problem involve different ways of weighting relevant attributes, different planes tend to end up being ranked differently (e.g., concerning “accuracy” or paper planes).
- When alternative solutions to the Paper Plane Problem involve drawing concentric circles (like a “dart board) to measure how far landing points were from the target, different kinds of circles transform the data in ways that invariably end up ranking planes differently (e.g., concerning “accuracy” or paper planes).

If a “dart board” diagram is used to measure distances from the target to each of the landing points for planes, then all of the distances in the outside ring get collapsed to a single number – and similarly for each of the other rings. So, these shifts in values often make significant differences when a collection of such scores are combined. Similarly, significant differences often result when different numbers of concentric circles. Landing points that used to be treated as being equivalent now may be treated as being different – and vice versa.

After finishing the model-eliciting activities for the Paper Airplane Problem, another important kind of model-exploration activity often engages students in designing their own simulations to generate data for an MEA that is similar to the Paper Airplane Problem. For example, teachers can use a simple-to-construct TinkerPlots simulation (similar to the one shown in Figure 13) in which the samplers are hidden for a “darts game” in which four players throw three darts – trying to be as accurate as possible. Then, after each student has created their own simulation, they can hide the samplers in their simulations and can trade their simulation with another student in the class. Then,
the other student’s goal is to determine whether the darts players are really different from one another concerning their abilities to throw darts accurately.

Model-Adaptation Activities – Investigating Students Abilities to Modify & Adapt Existing Models:

One topic that tends to evolve from the Paper Airplane MEA (and surrounding activities) involves sorting out sources of variation in a collection of data – which leads directly to the formal topic of analysis of variance. Therefore, to further emphasize simulations related to analysis of variance, two useful model adaptation activities (MAAs) are described below. They are called the School Friendliness Problem and the Virtual Golf Tournament. Both require students to modify basic ways of thinking that they developed in the context of the Paper Airplane Problem.

This table shows results from a study about "friendliness" at three schools (Gotham City South, Central, and North). At each school, data were collected from twenty-four boys and twenty-four girls at four grade levels (freshmen, sophomores, juniors, seniors). So altogether, data were collected from 576 students.

The scores for friendliness were based on information from a questionnaire - and from on-site observations that were made at each school midway through the school year.

Your Task: Write up a report, which will be published in the local newspaper, in which you answer the following questions. Were the scores significantly different for the three schools? For different classes? For boys versus girls? What are the chances that apparent differences were only the result unexplained variations (that is, factors not related to school, grade level, or gender)?

As the graphs in Figure 17 show, the differences were not large between schools, between grade levels, or between genders. Nonetheless, because the data in Figure 16 were generated using a TinkerPlots simulation, the student could repeat the experiment as many times as they wished; and, when they did so, they noticed that most of the relationships among schools and grade levels remained largely unchanged. So, by the time teams of students finished this MEA, they usually made specialized versions of the School Graph – one for each grade level (with boys and girls indicated separately using different colors) – and then they calculated (i) differences between the “centers” of the graphs, and (ii)
the amount of spread within each cell of these graphs. … In other words, these students invented their own methods for analysis of variance. And, regardless what kind of procedures they developed, it was not difficult to use visual methods to introduce them to traditional data analysis procedures.

The Computer Golf Problem
(Calculating "Adjusted Handicaps" for Players in a Computer Golf Tournament)

Real golf courses usually give "par scores" for the course as a whole and for each hole on the golf course. The "par score" is intended to be the score that a typical "good player" would be expected to get on each hole and for the course as a whole. Then, in tournaments, so that all levels of players will have more equal chances to win, individual players often are given "handicaps" - which are points that are added to their scores. For example, if a player has a handicap of 3, then, in order to win, this person needs to score three points better than a player with a zero handicap. (note: Players can have negative handicaps.)

In computer games, it is now possible to purchase computer golf games where players actually swing plastic "golf clubs" which determine how an avatar on a computer screen hits balls on virtual golf courses. And, just like in real golf, "par scores" and "handicaps" can be calculated - even after tournaments have ended - so that scores can be adjusted to give everybody a more equal chance of winning.

Next week, as a fund raising event, students at Center High School are planning to hold a computer golf tournament in which the players will be boys or girls from 10 to 17 years of age. Each player will find a list of sponsors - who are people who pledge money for points that their players score in the tournament.

Players who qualify for the tournament will be able to choose among sixteen different virtual "golf courses" which appear to be set in virtual locations around the world. So, the landscapes and weather conditions vary significantly from one course to another. Some are in windy and rocky seaside locations. Some are located near ski resorts high the mountains; and, others are located in deserts or wetlands. This means that some golf courses are likely to be more difficult than others - and neither the players nor the organizers of the tournament will know about these difficulties before the tournament starts. In fact, just like on real golf courses, each course can be played at more that one level of difficulty. For example, on real golf courses, it often is possible to "tee off" at more than a single location. And, on virtual golf courses, it is easy for each hole to be designed so that players can play each virtual golf course at one of four levels of difficulty - - which range from low, to low-middle, to high-middle, to high.

Based on past experience, the organizers of the tournament have concluded that the tournament is more fun, and that it gives more people a chance to win, if each player gets to choose their own golf course and level of difficulty. But, such choices also raise problems for the judges who must decide who are the winners for the tournament. ... To solve this problem, the judges have decided to give each player an "after-the-game handicap" which will be calculated after the tournament has ended. For example, if boys-as-a-whole score differently than girls, then the judges want to adjust boys handicaps to take these average differences into account.

And, they want the size of these adjustments to depend on: (a) how much difference there is between players of different ages, (b) how much difference there is between the various golf courses, and (c) how much difference there is between difficulty levels that are chosen at each golf course. ... So, after the tournament ends, and after it is possible to examine data to show how girls performances compare to boys, and how much difference there is among players at different grade levels, whether ninth grade students are better than seventh grade students, the judges want you to develop a formula that they can use to calculate after-the-game handicaps for each player. After the games have ended, the judges want the size of adjustments to somehow be related to differences that emerge during the tournament - for factors such as age level, gender, and course and difficulty level chosen.

YOUR TASK: Write a letter to the tournament's judges using the sample of data shown in the table below to show how they can use results of the tournament to calculate after-the-game handicaps that will be fair for all players who participate in the tournament that will be held next week. ... Your answers to these questions should include graphs, tables, and ways to measure the factors that you believe are important. ... Remember: There is no single "correct" formula to operationally define something like a "handicap" (or an "adjusted score"). You need to use basic operations (+, -, x, /, ^, sqrt()) to combine information in ways that makes sense in the given situation.
Whereas the *Friendliness Problem* involved inferential statistics, the *Computer Golf Problem* involved descriptive statistics. And, whereas the *Friendliness Problem* only involved assessing whether differences exist between the schools, grade levels, and genders, the *Computer Golf Problem* requires students to quantify similar types of differences. But, especially when students are engaged in creating their own Tinker plots simulations of these and similar problems, students themselves tend to create simulations where covariance is important to take into account. For example, consider a simpler version of the *Friendliness Problem* in which the grade level factor is omitted, and the number of schools is reduced to two. Figure 18 shows two different kinds of simulations that students often create to fit this story. The first assumes that boys and girls are impacted the same ways at both schools; whereas, the second assumes the boys and girls might be impacted differently at the two schools. In other words, the second assumes that there is an interaction between schools and gender; whereas, the first assumes that there is no such interaction.

For both of the simulation shown in Figure 18, final friendliness scores for each student is assigned using the formula:

\[
\text{Overall Friendliness} = \text{Individual Friendliness} + \text{School Friendliness} + \text{Gender Friendliness}.
\]

But, in other simulations, it might make more sense to combine factors multiplicatively, or using quotients or differences. ... Choices about all such issues involve choices of different models to describe the situations.

Starting from the left in the following two simulations, the first sampler assigns numbers 1-48 to each of the forty-eight students who are selected to be in the study. The second gives each students an certain amount of “friendliness” as individual people. The third assigns each individual to a school (North or South). The fourth column includes two separate samplers, and each uses a different sampler to assign friendliness scores to students at each school. The fifth column is where the two simulations are different.

The first simulation assigns “friendliness scores” in the same way for boys and girls at each school. But, the second simulation uses different samplers for boys (or girls) at one school and boys (or girls) at the other schools.
9. Conclusions

Partly as a result of three different mathematics education research design books which have been published during the past ten years (Kelly & Lesh, 2000; Kelly, Lesh & Baec, 2010, English, Bartolini, Bussi, Jones, Lesh, Srirman & Tirosh, 2008)), model-eliciting activities (MEAs) have come to be used widely in many different countries, at many different grade levels (K-16), and for different issues that involve student development (Lesh, 2003) or teacher development (Lesh, Haines, Galbraith & Hurford, A., 2010). But, the development of students’ and teachers’ thinking does not end with MEAs. So, this paper has described a variety of activities that have proven to be especially useful for investigating post-MEA development – or for investigating the development of thinking that doesn’t fit nicely within disciplinary topic areas – as well as knowledge that involves beliefs, feelings, values, dispositions, and a variety of metacognitive processes (Lesh, Carmona & Moore, 2020).

This paper has focused, in particular, on research tools and artifacts that can be used to investigate the development of mathematical thinking in an introductory course about statistics. But, the tools described here also are easy to adapt for use in courses focusing on many other mathematical topic areas. And, in any of these topic areas, an important point to emphasize is that the value of these tools does NOT depend on their ability to lead to significant learning gains in students. Nonetheless,
because *model-development units* do indeed tend to lead to significant types learning gains. This is because they are intended to create situations in which researchers will be able to reliably observe significant types of conceptual adaptations during sufficiently brief periods of time so that researchers can directly observe processes that lead to change.

Model-development units also tend to focus on deeper and higher-order understandings than the kind of knowledge or procedures that are emphasized in the kind of word problems in traditional textbooks and tests. One reason why this is true is because, whereas traditional tests tend assess only students’ first iteration answer to questions, *modeling activities* focus explicitly on students abilities to describe and design “things” mathematically – and to do so in the form of *thought-revealing artifacts and tools* that are sharable and reusable. For example, in statistics classrooms, students seldom are asked to “operationally define” constructs such as centrality, variation, error, or distance. Consequently, in traditional statistics classrooms, the decisions that students make about which routines to use usually are based on issues of correctness – rather than on the “goodness of fit” between assumptions about the situation and assumption which underlie the statistical procedures being used.

This distinction between “choosing correct computational procedures” and “developing sensible mathematical descriptions” is perhaps the most essential difference between (a) the kind of understandings and abilities emphasized in this paper, and (b) statistics as it has been portrayed in traditional statistics courses. Furthermore, it is a distinction that becomes especially important when the availability of powerful technology-based tools is recognized.

References


