Verifying Instructional Theory through Analysis of Patterns in Time

Theodore W. Frick Charles M. Reigeluth

Department of Instructional Systems Technology
School of Education, Indiana University, Bloomington

Paper for Presentation at the Annual Conference of the
American Educational Research Association
April 22, 1992

Abstract

Confusion and controversy exist among educational research methodologists, apparently because ***methods*** *of research have been conflated with* ***outcomes****. The spirit of educational inquiry is in danger if the "baby is thrown out with the bath water." We attempt to "keep the baby" by first discussing methods of fixation of belief, including the general method of inquiry, drawing on theory from Charles Sanders Peirce. Next is a discussion of kinds of knowledge that can result from inquiry, which builds upon theories from Gilbert Ryle, Elizabeth Steiner and George Maccia. Examples of qualitative, quantitative and praxiological knowledge are then provided. It becomes clear that development of instructional theory and its verification is an attempt to create praxiological knowledge of education. A conceptually simple method of verification of praxiological theory, called analysis of patterns in time (APT), is proposed and illustrated by example. Finally, implications of praxiological theory and expert systems are discussed in relation to design of instructional systems.*

Introduction

Over the last two decades considerable controversy has emerged in the field of educational research over methodology, as well as its purpose. One only has to survey issues of the *Educational Researcher* in the past decade to gain a flavor of the debate. Words such as 'qualitative,' 'naturalistic,' 'paradigm shift,' 'logical positivism,' 'multiple perspectives,' and 'phenomenology,' tend to stand out in the numerous discussions. This is the polite way to talk about the problem. To be more blunt about matters, the field of educational research is, in our view, a mess right now. It appears that we are returning to the Dark Ages in an earlier millennium, that we are losing the spirit of the renaissance in science, the fruits of which we now enjoy and take for granted.

This paper is an attempt to bring some order to the chaos that looms on the horizon. We then discuss where instructional theory and methods of verification fit into the larger picture of educational research as we see it.

The Spirit of Disciplined Inquiry: "It's the Right One, Baby — Uh-Huh"

To set the stage, imagine a researcher in the Middle Ages. The socially constructed reality of that time was that the world was flat, and moreover the earth was the center of the universe over which the sun and stars moved. This was common knowledge, as any naturalistic inquirer who merged into the discourse of the time would discover. To believe otherwise was heretical, as Copernicus and Galileo would discover later (cf., Bronowski, 1973).

The renaissance in science was that theory must be confirmed, refuted and revised by legitimate evidence. To simply hold a theory as true or best because the majority believes so is not sufficient evidence. Prejudice and tenacity of beliefs abounded in the Middle Ages — and still do now.

The theory that the world is flat was falsified, for example, by evidence from Magellan's voyage around the world. How could he sail around a flat world? The belief and the evidence were inconsistent. If the world were flat and finite, then we would expect to find some evidence of the edges or ends (e.g., from eye witnesses). To persist in the belief and ignore the evidence, is to use the *method of tenacity*, as Charles Sanders Peirce (1958) reminded us half a millennia later.

Even in the legal system today we do not convict someone of a crime on the basis of hearsay evidence. All of the legitimate pieces of direct evidence must fit together in a coherent and consistent manner. If they do not, then we say there is reasonable doubt and the defendant is set free. And there are clear rules about what constitutes legitimate evidence.

Peirce (1958) discussed four ways of fixating belief: Tenacity, authority, agreeableness to reason, and inquiry.

1. The *method of tenacity* is simply to hold onto one's beliefs (or theories) despite legitimate evidence to the contrary. An example in this century would be one who believes in behaviorism to the exclusion of mind or will, for example, despite evidence that Socrates *chose* to drink the hemlock rather than change his ways some 2400 years ago. A further example is those who believe completely in constructivism. When this particular theory is taken to its logical conclusion, however, it contradicts considerable historical evidence that guided learning is usually more efficient and effective than totally self-directed learning[1].

2. The *method of authority* is when we hold beliefs based on what we are told is true or right by someone in a position of power or prestige—i.e., a perceived authority. People who followed and upheld the dictates of Hitler or Saddam Hussein are good examples. They believe what they do because their leader says so, even when there is legitimate evidence to the contrary. For example, Hitler preached that a particular line of the German race was superior. Those who continue to believe this, despite the fact that there is considerable evidence that falsifies it, are following the method of authority. As another example, if you believe that constructivism is untenable because David Merrill said so, and only on that basis, then you are following the method of

[1] If self-directed learning *alone* (i.e., discovery learning or trial-and-error learning) were as effective as guided learning, then there would be no third-world countries. Think about it.

authority[2]. As a final example, people who believe that birth control is wrong and only because their religious leader says so are following the method of authority, despite evidence of starvation and disease due to overpopulation where they live.

3. The *method of agreeableness to reason* is when we hold beliefs because they should be true according to rational thought. For example, it is reasonable to believe that humans with larger brains would be more intelligent. It stands to reason that if there is more cortical matter, then one could remember more and do more with those memories. Therefore, people with larger brains should be smarter. To continue to hold this belief, despite considerable evidence to the contrary, is to use the method of agreeableness to reason.

4. The *method of inquiry* is the scientific way to resolve uncertainty and form belief. We refer to science in its broadest sense, not in the narrow sense of the logical positivists or of the more recent phenomenologists, but in its original Greek sense (i.e., knowing). According to Peirce, and as implied in the title of this section, the method of inquiry is "the right one, baby — uh-huh," to borrow from a 1991 soft drink commercial. Now, if you believe this because Peirce says so, or if you believe that Pepsi is the best soft drink because Ray Charles endorses it, then you are using the method of authority.

Instead, consider the evidence that supports the inquiry method. Consider the Renaissance and the advances in human knowledge since the Middle Ages, for example. Consider the automobiles you now can drive, planes you fly in, television you watch, the electric light you read by, the books you read, the antibiotics you take for infections, immunizations against diseases, and on and on. The spirit of the scientific method — the method of inquiry — has proven itself to be central to advances in knowledge. That research is *sine qua non* in higher education is no historical accident.

Knowledge should not be confused with uses to which it is put. In addition to uses which benefit humankind, knowledge can be ill-used for the development of more powerful and sophisticated ways to maim and kill each other, and our environment.

The Essence of Disciplined Inquiry

The epitome of disciplined inquiry is development and verification of theory. When verified, theory becomes knowledge. The central purpose of inquiry is to create knowledge. Knowledge in turn can be used to explain phenomena, to predict, or to achieve ends. Knowledge also can be created simply for its own sake — i.e., for better understanding of ourselves and our world.

Knowledge is neither exclusively subjective nor solely objective, but instead intersubjective. Knowledge is a set of public signs[3] shared among humans which represent states of affairs (cf. Steiner, 1978; Peirce, 1977; Bruner, 1990).

Moreover, knowledge is not absolute. Knowledge changes, mostly as a function of further inquiry. Karl Popper (1963) summarizes it succinctly: Knowledge evolves

---

[2] See *Educational Technology Magazine*, May and September 1991 issues for more on this debate.

[3] Signs are the "stuff" of communication. Signs can include spoken and written language (linguistic symbols), icons, gestures, pictorial representations, non-linguistic sound patterns, smells, touch, facial expressions, expression of emotion, and demonstration by enactment. Signs take on meaning for humans through use during situated action (see Bruner, 1990, for further discussion).

through conjectures and refutations. Thus, theories are never finished products but subject to extension, revision or discard through methods of verification (confirmation/refutation).

Kinds of inference. Three kinds of reasoning are used in disciplined inquiry according to Peirce (1958): 1) retroduction, 2) deduction, and 3) induction. *Retroduction* is reasoning by analogy or metaphor to develop a new theory or to modify a falsified theory. *Deduction* is reasoning according to logic, to derive specific implications from axioms or general principles in a theory. *Induction* is the making of observations in order to verify a theory. If the observations do not fit the theory, we normally modify the theory further; or we make additional observations if we suspect some interference or error in our judgments or observational methods; or we simply discard the theory[4]. These three kinds of inference occur repeatedly during a particular inquiry, and their order may vary. For example, inquiry could begin with induction, when one observes something which refutes a known theory.

Choosing one theory over another. The problem with 'multiple perspectives' is that it is often the case that all are not equally valid. Some theories are more adequate than others. Parsimony in the evolution of knowledge requires that the best theories are retained in human cultures. Steiner (cited in Maccia, 1974) offers these criteria for choosing one theory (*X*) over another (*Y*):

1) Theory *X* is more complete than *Y* if and only if *Y* is derivable from *X*, and *X* describes relations which are not described by *Y*.
2) Theory *X* is inclusive of *Y* if and only if the conjunction of *X* and *R* includes *Y*, where *R* is a set of translation rules matching expressions in *X* and *Y*.
3) Theory *X* is a strong alternative to *Y* if and only if *X* includes all those data, phenomena, events or event relations that *Y* includes; *X* and *Y* are empirically inconsistent; and *X* has higher empirical content than *Y*.[5]

Thus, one theory can be chosen over another in disciplined inquiry on the basis of the following criteria: completeness, inclusiveness, or strength (coherence and consistency). These criteria apply not only to comparisons of qualitative theories with each other, but also to comparisons of quantitative and praxiological theories, respectively, which are discussed next.

Epistemology: "What Do You Know? Ya Know?"

The next area we address is kinds of knowledge. Not all knowledge is of the same kind. It would appear that much of the current debate and confusion in educational research methodology might be clarified by making first a distinction among kinds of knowledge, and second a distinction among various methods that can be used to create

---

[4]Unless of course we are tenacious -- i.e., keep the theory and ignore the data!

[5]See Steiner (1988) for a more precise and elaborated discussion of criteria for evaluating theories, especially chapters 4 and 5.

such different kinds of knowledge. This would help separate issues concerning *outcomes* of research versus *methods* to achieve them.

1. Again, let us start simply. We may believe that our earth has seven continents surrounded by large oceans and has two polar ice caps. It is like no other planet in the universe we know. This is a belief in a *specific* entity. The object of the belief is unique.

Knowledge of the specific is qualitative; that is, it distinguishes what is unique about some particular state of affairs, *what qualities set it apart from all else.*

2. We may further believe that most heavenly bodies are spherical, such as our earth, sun and other stars. This is a belief not in a statement about a unique entity but instead about a commonality among multiple entities. We typically refer to such a belief as a *generalization.* The belief concerns a *class* of objects or states of affairs. Gilbert Ryle (1959) refers to such cognition as 'knowing that.'

George Maccia (1987; 1988) however makes the distinction between the general and the specific. In Maccia's lexicon 'knowing that' pertains to the general and 'knowing that-one' pertains to the unique. In David Merrill's (1983) component display theory, he distinguishes among knowing facts from knowing concepts and principles. Facts are specific; concepts and principles are generalizations[6].

Knowledge of the general is quantitative; that is, it concerns what is generalizable. Such knowledge is not bound to particular persons, places, events, things, or their interrelationships. Quantification implies extension relative to class. We view this as an instance and that as a non-instance of some category, pattern or relation. We want to make statements that are *generally* true. Or at least we try to make probabilistic statements about the occurrence of states of affairs[7].

Quantitative knowledge includes the non-empirical as well. Philosophy, which deals with matters of value, is quantitative. For example, a philosophic proposition in education is that teachers should respect all students as human beings first and foremost. Philosophic knowledge concerns what is worthwhile *in general.*

Quantitative knowledge is concerned not only with generalizations about 'what is' but also about 'what ought to be.'

3. Israel Scheffler (1965) points out that our language does not have an expression for 'believe how.' We do not normally say we *believe* how to read printed text, believe how to build a skyscraper, believe how to perform open heart surgery, believe how to solve quadratic equations, etc. Instead we say we 'know how' to do these sorts of things. Such knowing is not mindless action, but mindful. Know-how requires intelligent cognition of particular circumstances, making judgments based on the conditions, and choosing appropriate courses of action when warranted — all in order to achieve some desired outcome. Know-how is purposeful.

---

[6]Instantiation of a generalization should not be confused with knowing that-one. To know our specific earth in all its beauty and ugliness, its unique continents, lakes, rivers and oceans is 'knowing that-one.' On the other hand, to view Earth as an instance of a planet is 'knowing that.' The generalization is the concept of 'planet.' An instance is Earth, but other instances include Mercury, Venus, etc.

[7]An example of a probabilistic statement is: The likelihood of student non-engagement with academic tasks when no direct instruction is occurring in elementary schools is 13 times greater than when direct instruction is occurring (cf., Frick, 1990).

Moreover, such know-how can be shared. Hence it can become knowledge in a human social system. It can become intersubjective. It can be represented by a public set of signs[8] that are shared among humans[9].

Knowledge of how-to-do is praxiological; that is, it concerns means to achieve ends.

In education, we often refer to such knowledge as pedagogy — i.e., knowing how to facilitate learning. Instructional theory and its verification is the subject of the latter half of this paper. Such knowledge is distinct from knowing that and knowing that-one.

Summary. Disciplined inquiry can result in three distinct kinds of knowledge: of the specific, of the general, and of how-to-do. These are referred to, respectively, as qualitative knowledge, quantitative knowledge, and praxiological knowledge[10].

It is also patent that, while these three kinds of knowledge are distinct, they can certainly be related — indeed it is often desirable that they are. Also note that these distinctions are *not* based on how knowledge is used, such as conclusion- versus decision-oriented inquiry (Cronbach & Suppes, 1969), prescriptive versus descriptive, or basic versus applied research.

Examples of Qualitative Knowledge

To conduct inquiry to create qualitative knowledge is to develop and verify a theory about some *unique* state of affairs. The purpose is not to generalize but to adequately describe some specific state of affairs (event, person, place, setting, etc.).

Example 1. A familiar example of such inquiry outside of education was the investigation and report on the assassination of John F. Kennedy in 1963. An official group known as the Warren Commission attempted to determine who killed the U.S. president and how. At least two theories were seriously considered: 1) that Lee Harvey Oswald acted alone and shot the president from the 6th floor of the Texas Book Depository as the motorcade passed below on the street, and 2) that others were involved in addition to Oswald, and shots were fired from several other locations (the conspiracy theory). The Warren Commission concluded in its report that the sum total of legitimate evidence supported the first theory more than the second. Numerous attempts have been made to refute this theory over the past three decades, including a recent movie, *JFK*. Possibly some day enough additional evidence may become available which clearly refutes the first theory and supports the second. And possibly not.

It would be interesting to apply Steiner's criteria for choosing one theory over another to the J.F.K. assassination.

---

[8]See footnote 3.

[9]Unfortunately, in a capitalistic economy, private enterprises tend not to share their know-how with outsiders in order to protect their profits. Patents and copyrights abound.

[10]It should be also evident that methods of inquiry are instances of praxiological knowledge (i.e., ways of doing research).

Example 2. An example of inquiry to create qualitative knowledge of education would be to document the historical evolution of the Montessori education movement in Cincinnati, Ohio. Again, theory is developed and it is tested by collecting evidence to confirm, refute or refine it[11]. In this case, verificational methods could include interviewing persons who were and are involved; obtaining authentic written documents such as minutes of school board meetings, letters of correspondence among teachers, parents and others involved; finding authentic photographs, video or film footage, etc. There is no attempt to generalize beyond the object of inquiry. The inductive inference pertains to the specific, in order to establish coherence between the theory and the legitimate evidence for the unique state of affairs.

Example 3. A further example of inquiry to produce qualitative knowledge would be to evaluate a specific educational program. For example, an accreditation agency such as NCATE will visit a school of education and attempt to develop and verify a theory about the worth of that specific educational program. Such an evaluation is not to determine the worth of schools of education in general but to make a disciplined judgment about an individual, unique educational program.

Example 4. A final example, not in education, would be an Olympic diving contest. Contestants are given opportunities to complete a fixed number of dives. Their performances are observed by a panel of judges who know criteria for good dives and who are competent observers. Extreme ratings from judges are discarded, and the best diver is determined to be the one with the highest aggregate score after all dives are completed. *This outcome is qualitative knowledge, even though quantitative inquiry methods are used.* The outcome should not be confused with the method to achieve it. In this case the outcome is determination of the best three divers among the competitors on this particular day by ranking them according to their aggregate scores.

Examples of Quantitative Knowledge

To conduct inquiry to create quantitative knowledge is to develop and verify theory about generalizations regarding states of affairs. The goal is not to deal with uniqueness as in qualitative knowledge, but to deal with commonality or similarity among states of affairs. Quantitative knowledge includes what has historically been considered scientific knowledge — about 'what is' in general. Quantitative knowledge also includes philosophy — about 'what ought to be' in general (Steiner, 1978; 1981; 1988).

Examples of quantitative knowledge abound. Most people are more familiar with such knowledge, since it has been emphasized heavily in our prior educational experiences. We will discuss just a few examples, to help clarify and elaborate the concept of quantitative knowledge.

Example 1. In physics Isaac Newton developed a theory of gravitation. Concepts of force, mass, velocity, acceleration, etc. were developed. He discovered mathematical relationships such as: force equals mass times acceleration. These relationships were found to be generally verifiable when dealing with large objects

---

[11] History (his story) is after all human belief about the past. It could be called herstory or ourstory as well. It is nonetheless theory about 'that-one' state of affairs.

moving through physical space. Planetary motion could be predicted. Trajectories of cannon balls could be forecast, but not those of falling leaves. Newton's theory did not work well when large numbers of bodies were involved, velocities approached the speed of light, or with entities at the atomic level such as protons and electrons. Einstein later extended significantly Newton's theory by developing a theory of relativity. Einstein's theory is more adequate than Newton's on the basis of completeness, inclusiveness and strength.[12]

Example 2. Another well-known example of quantitative knowledge is Mendelian genetics. In contrast to Newtonian physics, one could not predict a characteristic of a *single* off-spring, such as whether a pea plant would have wrinkled or smooth pods. But the probability of such characteristics could be predicted and confirmed by experiments in genetics (e.g., 25 percent of the plants might be expected to have wrinkled pods, given knowledge of the configuration of their parent's genes for this characteristic).

Example 3. An educational example of quantitative knowledge was Rieth and Frick's (1983) observational study of student and teacher activities in different special education settings in elementary schools. One finding was that, when direct instruction occurred, student engagement in academic tasks was very high (97 percent). When no direct instruction occurred, students were engaged about 57 percent of the time. While there was some variation from individual to individual, these patterns were observed to hold regardless of the type of setting (resource room, special class, regular class)[13].

Example 4. In the field of education, philosophic inquiry to produce quantitative knowledge would address questions such as, "What *should* all students learn? or "What criteria should be used to evaluate student learning outcomes?"

To conflate the empirical ('what is') with the philosophical ('what ought to be') is to commit the naturalistic fallacy. Philosophic inquiry does not deal with empirical questions such as, "What *do* students learn?" — either specifically or in general. Nor does it treat praxiological questions such as, "How *can* we best facilitate student learning?" Nonetheless, philosophic inquiry is quantitative since it pertains to what is worthwhile or of value in general.

Discussion. Quantitative knowledge applies to both deterministic and stochastic systems. For example, Newton's laws of motion pertain to a deterministic system, whereas Mendel's pea plants and elementary classrooms are stochastic systems. Yet generalizations can be made about both kinds. The certainty of the outcomes is not the critical attribute, nor the methods of verification. Rather, a theory must be extensible to some *domain* of states of affairs.

---

[12]See the discussion at the end of an earlier section on criteria for choice of one theory over another.

[13]See Frick (1983, 1990) for a discussion of the research methodology, analysis of patterns in time, and of the issue of stochastic versus deterministic systems. See also Gerald Weinberg's (1975) lucid discussion of deterministic and stochastic systems.

Methods of empirical verification can include naturalistic observation (e.g., of planetary motion, of classroom activities) or experimental manipulation (e.g., crossing breeds of pea plants, shooting cannon balls at different angles)[14].

Two further issues are noteworthy for inquiry to create quantitative knowledge: 1) Development of theory involves retroductive and deductive inference. For example, in Newton's case the falling apple was apparently the metaphor that retroductively implied the idea that planetary bodies might also be "falling" around the sun, and the results of deduction were his three laws and numerous mathematical relationships. 2) Inductive inference requires observations of states of affairs which are *representative* of the domain of interest, in order to make generalizations about it.

Induction for creating quantitative knowledge is essentially this kind of reasoning:

1) A is true of $b_1, b_2, ..., b_n$.
2) $b_1, b_2, ..., b_n$ are members of some class B.
3) Hence, A is true of all members of class B. (Steiner, 1978, p. 9).

When empirical matters are of concern in quantitative theory (i.e., about 'what is'), the form of inductive inference is statistical (cf. Salmon, 1966; 1971). Therefore, issues of how an inquirer *samples* from some domain (class of states of affairs) are paramount. Moreover, the number of samples (observations) and how they are taken is related to the certainty of the inductive inference. These topics have been treated extensively in traditional research methodology. One thing is very clear: It is often hazardous to make a generalization based on an extremely small sample — the margin of error can be rather large in this situation.

The first author has observed that undisciplined inquirers are prone to make generalizations from insufficient evidence. When they do so, I ask them to get one of two covered coffee cans from the bookshelf in my office. Each contains one hundred cubes, one of which has 40 red and 60 white, the other 60 red and 40 white. I require them to determine which container they have selected, by randomly sampling one cube at a time and observing its color. After each draw, I ask them how much money they are willing to bet on correctly identifying which container they are holding.

The irony is that most students are not willing to bet any money at all until a sample of 10 to 20 cubes is taken. Yet they were perfectly willing to make some generalization based on only a few instances earlier in the conversation — somehow it is o.k. to do so if they are doing qualitative or naturalistic research, but it is definitely not o.k. when they have to put money on the line and make a quantitative inference! This, by the way, is an example of the confusion that obtains in the present educational research milieu.

When philosophical matters are of concern in quantitative inquiry (about 'what ought to be'), inductive inference of the kind discussed above is not appropriate. Induction instead occurs through the dialectic, through criterial analysis and philosophic discussion. Socratic dialogue in *The Republic of Plato* is a prime example, where the question, "What is justice?" was central.

---

[14]Moreover, we can do experimental manipulation to produce qualitative knowledge. Consider for example an auto mechanic conducting a series of tests in order to understand your particular malfunctioning car.

Examples of Praxiological Knowledge

To conduct inquiry to create praxiological knowledge is to develop and verify theory about means to achieve ends. We can design procedures, methods, inventions, etc. and observe how well they work. We often refer to this as knowledge of practice — i.e., knowledge of how-to-do. It is important that the practices are replicable — i.e., not just a one-time solution to some particular problem. Praxiological knowledge is technology in its broadest sense, not just equipment or instruments.

As with qualitative and quantitative knowledge, inquiry to create praxiological knowledge involves: 1) development of theory by means of retroductive and deductive inference, and 2) testing of theory for confirmation or refutation (inductive inference).

Qualitative and quantitative knowledge are non-instrumental, represented by public signs that characterize 'what is unique' versus 'what is general' and 'what ought to be in general.' Praxiological knowledge, on the other hand, is instrumental and represented by public signs that characterize 'what works' — means to achieve ends.

Criteria for comparing praxiological theories include: effectiveness, efficiency and compatibility. For example, one treatment for cancer may be effective (i.e., it kills the cancer cells), but it is not efficient (i.e., requires many expensive injections), and it is not compatible (i.e., it has undesirable side effects such as complications in the kidneys). Another treatment for cancer might be less effective, but more efficient and compatible (i.e., though it does not kill as many cancer cells, it is less expensive and does not harm internal organs). The second theory would appear to be more adequate than the first.

Example 1. For thousands of years humankind tried, and failed, to fly like birds, as did the mythic Icarus. Most of these inquiries developed and attempted to verify variations on a theory of flapping wings. The Wright brothers created and tested a different kind of praxiological theory, that of a plane slicing through the air which was propelled by a source of power. Their theory of flying by means of a propelled airplane underwent many trials until they made their famous inaugural flight at Kitty Hawk.

Even then the Wright brothers' design (theory) required many further revisions and tests before stable and replicable flight was achieved. It was not too many years before passengers were able to enjoy a new form of transportation. Now, nearly a century later, the theory has been extended to air and airless flight as exemplified by NASA's space shuttle program.

Praxiological knowledge should not be confused with how it might be used. The actual invention, e.g., a particular airplane, is created to empirically test the theory — to see if the *theory* works. If it does work consistently[15], then the theory becomes knowledge. The knowledge in turn can be shared with others who can apply it — e.g., to use a blueprint, a model, and a set of directions to build a replicate[16].

---

[15] Consistency does not imply certainty of outcomes. See the below discussion of stochastic systems.

[16] In this case the public signs are drawings in the blueprint, symbols such as words and numbers on the blueprint and in the directions, and a model (a picture of what it is supposed to look like when finished).

Example 2. The first author and several of his doctoral students have invented an expert systems approach to computerized adaptive mastery testing. The goal of a computerized adaptive test is to determine a student's mastery of a given educational objective with the fewest questions possible, while ensuring that the decision agrees with that reached if a large number of test questions were administered to that student.

We began by first investigating a sequential decision procedure developed by Abraham Wald (1947) known as the sequential probability ratio test (SPRT). Wald's praxiological theory worked reasonably well as long as a fairly large number of test items was available to sample from, and the decision error rates were kept small. However, the theory was potentially suspect because easy and difficult items on a test were weighted equally. The praxiological theory was revised by incorporation of theory about an expert system rule base that would contain item difficulty and discrimination information on past masters and non-masters of a test. This new theory (EXSPRT) was then empirically tested and compared to other praxiological theories of adaptive testing such as item response theory. The EXSPRT theory appears to work better than the others for computerized adaptive testing to determine mastery of a single educational objective.

The praxiological knowledge resulting from this research is signified in papers presented and published (Frick, 1992; Frick, 1990; Frick, Plew & Luk, 1989). Additional public signs include decision algorithms, computer software listings, demonstration programs that illustrate the praxiology, and a textbook (under preparation).

Example 3. Someone devised a theory of selling merchandise such as food and soft drinks automated by a machine. We do not know to whom to give credit, but such knowledge is nonetheless praxiological. The theory allows for customer input of money and selection of the wanted product. Based on the amount of money inserted and which selection is made, a corresponding product is dispensed; and change is returned if necessary. The vending machines which are commonplace today are applications of automated vendor theory.

This example was chosen because it, like the above examples, is a theory about a *deterministic* system. If the antecedent conditions are known then the outcomes are predictable — at least most of the time.

Occasionally, however, airplanes crash, computerized tests make incorrect mastery decisions, and vending machines take our money and provide nothing in return.

If the desired outcomes do not happen regularly enough, praxiological theory may be flawed; or one of its applications has "broken" and no longer corresponds with the originally verified theory. Deterministic (or mechanistic) systems are very predictable. A given set of input conditions normally results in a single output condition.

We can also develop praxiological knowledge about stochastic systems, as discussed in the next section.

Where Does Instructional Theory Fit into the Picture?

The development of instructional theory and its verification, historically referred to as pedagogy, is one kind of praxiological knowledge of education. Such theory basically addresses the question of how to best facilitate learning. Learning can occur by trial-and-error, although we would not consider this to be education, for to educate is to guide or lead. Educational systems are guided learning systems (cf., Steiner, 1988;

Frick, 1991). Guidance does not imply that all instruction must be direct. An example with minimal direct face-to-face instruction from teachers is the Montessori system of education.

A problem in pedagogical inquiry is that educational systems are stochastic. Unlike the automated vendor theory, we cannot predict with near certainty an outcome, given a known set of antecedent conditions and methods of instruction.

Imagine for a moment a vending machine that has gone somewhat but not totally awry. We'll call it the Un-Machine for now. For example, when you insert money and push the Pepsi button, sometimes you get a Pepsi (45 percent), sometimes you get a Coke (50 percent), and occasionally you get an empty cup (5 percent). On the other hand if you push the Coke button, you get a Coke (20 percent), a Pepsi (70 percent), and an empty cup (10 percent of the time). Suppose you have tested this bizarre vending machine hundreds of times, and the above probabilities are found to be relatively consistent[17].

Well, it's not the greatest of worlds, but if that is all we have, we can nonetheless use this stochastic knowledge to make choices. For example, if we want a Pepsi we should push the Coke button. Why? The probability of a Pepsi is .7 when we choose Coke, whereas it is .45 when we push the Pepsi button. On the other hand, if we want a Coke, then we should press the Pepsi button. Why? The probability of getting a Coke is .5 here, but only .2 if we press the Coke button. Of course, we'll still get a Pepsi more often than a Coke if we press the Coke button (odds are 7 to 2).

If we don't care if we get a Coke or Pepsi but just want some cola to drink, then we should press the Pepsi button, since it results in half as many empty cups as does the Coke button.

And if you're about ready to kick the Un-Machine, just be glad that the probabilities are stable over time. If the probabilities start shifting, then you might think about getting a drink of water instead — and then start collecting new data on the Un-Machine.

Human activity systems, which include educational systems, appear to be more stochastic than deterministic (cf., Frick, 1983). There appear to be at least two reasons for this:

1) People have wills, feelings and cognitive limitations. For example, sometimes students really want to learn, and other times they don't. And sometimes they are impaired by fatigue, frustration, or even too much excitement. And sometimes they can't learn because they are intellectually incapable at the time.

2) Human activity systems are highly *complex*, unlike our relatively simple but bizarre stochastic vending machine. There are many elements in educational systems, and the number of interactions among elements is often incredibly large, making such systems difficult to understand.

Thus, the development and verification of instructional theory is even more formidable than the task faced by the Wright brothers. At least they were dealing with a relatively deterministic system[18]. Nonetheless, pedagogical knowledge can be created through disciplined inquiry.

In the next section we discuss a formative methodology for instructional theory *development*. In the section following that we propose a means of keeping track of *verification* attempts through analysis of patterns in time (APT). In the final section we

---

[17]Remember that Mendel's reproducing pea plants were consistent stochastic systems as well.

[18]Their airplane didn't complain about having to get up so early, or daydream about lying on the beach ...

consider how APT probabilities can become part of an expert system for instructional design.

The Formative Approach to Instructional Theory Development

The purpose of formative research is to improve our praxiological knowledge base, where the focus is on how useful the theory is rather than how truthful it is (Snow, 1973). Formative research differs from experimental research in that the latter is concerned with comparison rather than with improvement. This distinction is parallel to that between formative and summative evaluation. Furthermore, formative research is concerned with prescribing means to ends, rather than with describing what is. When we ask, "How can we improve this method?", we need formative research. Neither experimental studies nor standard descriptive qualitative methodologies are very helpful for answering the question.

So what is the formative research methodology? We have found the following to yield useful results (see e.g., Farmer, 1989; Roma, 1990; Simmons, 1991), but much work remains to develop helpful details of this methodology. First, pick a method for facilitating learning (e.g. an instructional theory or model) and design and develop an instructional system solely on the basis of it (much like one designs a treatment in an experimental study). Try not to use any other prescriptions, not even your own intuition, in designing that system. This should make the system a true instance of the method in the same way that treatments in an experimental study are instances of the parameters of the independent variable.

Second, conduct a series of one-on-one and small-group formative evaluations of that instructional system. As in an experimental study, be sure that your students are typical of the population to which you wish to generalize. Your initial one-on-one evaluations should probably be fairly obtrusive, with constant interaction between you and the student to identify any problems the student is having and to get suggestions from the student as to how the method might be improved. You should correct any gross problems with the system right away, as long as you have reason to believe the corrections will be beneficial for all students, or at least not detrimental to any.

You should compare these results (problems, suggestions, and actual improvements) across a representative range of students until you get consistent findings. The less consistent your results, the more students you need to run through this process. If you keep getting inconsistent results, try to identify any conditions (factors) that appear to be interacting with the methods used in the system.

Then you should conduct some unobtrusive one-on-one and small-group evaluations to corroborate those results — i.e., to assess their validity under more naturalistic conditions. These might be done by videotaping the instruction and then, immediately after the instruction, reviewing the videotape with the student and asking questions related to the earlier evaluation results. You should correct all problems with the system.

Several rounds of obtrusive followed by unobtrusive formative evaluations are advisable, especially if major revisions are made in the instructional system each round. Naturally, as in an experimental study, replications with different content, different aged learners, and different settings and media would also promote external validity. This is addressed in more detail in the next section.

On one level, the results of these formative evaluations identify weaknesses in the instructional system and suggestions for improving it; but, since the system is an instance of the theory, the results also reflect similar weaknesses in the theory and provide suggestions for correcting those weaknesses. This kind of study yields much more data about a broad range of features of the theory, and these data are far more relevant for improving the theory than any experimental study could be.

An Example of Verification of Instructional Theory by APT

Verification of instructional theory, when compared to its development, is analogous to the difference between summative evaluation and formative evaluation. *Verification* of an instructional theory requires additional considerations in order to warrant its replicability in practice. Traditional experimental methods are not well-suited for this primarily because they are designed for inquiry into deterministic systems (cf., Frick, 1983). Instructional systems are clearly stochastic. Analysis of patterns in time (APT) was invented for empirical investigation of stochastic systems (Frick, 1990; 1983). The relevance and fruitfulness of using APT to verify instructional theory and add to praxiological knowledge of education is discussed below.

C x M x O. Reigeluth (1983) has made general distinctions among conditions (C), methods (M) and outcomes (O) when developing instructional theory. Conditions are circumstances that are expected to make some difference in outcomes attained by certain instructional methods.

Conditions. One kind of condition is the type of learning objective. For example, in Component Display Theory Merrill (1983) considers a matrix resulting from crossing student behaviors (remembering, using and finding) with certain content types (facts, concepts, rules and procedures). In George Maccia's epistemology learning outcomes are characterized as knowing that-one (recognitive, acquaintive, appreciative), knowing that (instantial, theoretical, criterial), knowing how-to-do (procedural, performative) and knowing what-to-do (innovative, creative) (see Frick, 1992, for further details). While Maccia's taxonomy is more comprehensive than Merrill's, we will use Merrill's for purposes of illustration and because it is more well-known at this time.

In analysis of patterns in time (APT) it is necessary to define a set of classifications which are relevant to the theory being investigated. One of Merrill's condition classifications is type of content. A classification must consist of a set of mutually exclusive and exhaustive categories. If we use Merrill's schema, we would have the following:

Condition classification name:	CONTENT TYPE (CT)
Categories within CT:	fact, concept, procedure, principle, null[19].
Condition classification name:	STUDENT BEHAVIOR TYPE (SBT)
Categories within SBT:	remember, use, find, null.

---

[19] Null is an operational way of indicating that there is nothing occurring at some point in time which is relevant to the classification. The use of the null category will become apparent in the APT scores (observational record) in the appendix.

Further condition classifications that would appear relevant are student motivation to learn and relevant prior knowledge. For purposes of illustration, let us add these two condition classifications:

Condition classification name:	STUDENT MOTIVATION (SM)
Categories within SM:	Positive, neutral, negative, null.
Condition classification name:	RELEVANT STUDENT PRIOR KNOWLEDGE/EXPERIENCE (RSPKE)
Categories within RSPKE:	Adequate, inadequate, null.

Additional condition classifications could be considered as an instructional theory is further developed, such as teacher knowledge of the subject matter and teacher competence in carrying out various instructional strategies and tactics. One might also consider whether teaching is face-to-face with students or mediated through technology (i.e., indirect). Subject matter could be further characterized using, for example, Howard Gardner's (1983) schema (mathematical/logical, linguistic, kinesthetic, musical, visual/spatial, interpersonal and intrapersonal knowing). Or one could cross Gardner's types of intelligence with Maccia's epistemology.

Methods. Let us move on to method classifications. Again, for purposes of illustration and simplicity, we will continue with Merrill's characterizations of instructional tactics. Merrill (1983) discusses four kinds of presentation forms, two kinds of presentation types, and nine kinds of secondary presentation elaborations[20] :

Method classification name:	PRIMARY PRESENTATION FORMS (PPF)
Categories within PPF:	expository (expos), inquisitory (inqui), process displays (procd), procedural displays (prced), null.
Method classification name:	PRIMARY PRESENTATION TYPE (PPT)
Categories within PPT:	generality (gener), instance (insta), null.
Method classification name:	SECONDARY PRESENTATION ELABORATIONS (SPE)
Categories within SPE:	context (cntxt), prerequisite (prerq), mnemonic (mnem), mathemagenic help (mathp), representation (repr), correct answer feedback (cafb), help feedback (hlpfb), use feedback (usefb), null.

We can also characterize methods at a broader level, i.e., instructional strategies. Strategies are at a macro level, whereas tactics are at a micro level. Strategies can be viewed as a combination of tactics in order to facilitate wanted learning outcomes.

[20] If Merrill had known about APT he would not have needed to invent such awkward ways of identifying combinations of things such as Eeg, Ieg, IG, etc. Numerous classifications can be considered temporally in APT. Mnemonic abbreviations in parentheses here are those used in the APT scores in the appendix.

Again, for purposes of illustration we will keep it simple and define a few strategies that could be used to facilitate concept learning:

Method classification name:	INSTRUCTIONAL STRATEGIES (IS)
Categories within IS:	Strategy 1, Strategy 2, ... Strategy *n*, null.

Strategy 1 is defined as follows: Present a generality, then present examples and non-examples, then provide practice with feedback. Strategy 2 is defined: Present best examples and non-examples, then provide practice with feedback. Strategy 3 is defined: Present examples and non-examples, then present a generality.

Outcomes. Finally, we attempt to characterize learning outcomes. Following the same pattern as above, we will look at tactical outcomes and strategic outcomes. Tactical outcomes are intermediate, ones that occur during instruction as indicated by learner attention, verbal responses, and task performance. Strategic outcomes are the ones wanted when instruction is completed. We typically determine those by testing or observing students.

Outcome classification name:	STUDENT TACTICAL PERFORMANCE (STP)
Categories within STP:	Pays attention (pattn), correct answer (cans), incorrect answer (ians), asks question (askq), makes relevant comment (comm), performs task with high success (thi), performs task with medium success (tmed), performs task with low success (tlo), inattention/off-task (offt), absent (abst), null.
Outcome classification name:	STUDENT LEARNING ACHIEVEMENT (SLA)
Categories within SLA:	Mastery, partial mastery, nonmastery, null.
Outcome classification name:	STUDENT AFFECTIVE ACHIEVEMENT (SAA)
Categories within SAA:	Positive, neutral, negative, null.

Additional outcome classifications could be considered such as development costs for an instructional strategy in terms of time, money and necessary resources. Efficiency can be determined as well, since APT scores permit easy calculation of temporal duration.

Building case histories with APT scores. Six case histories are presented in the appendix. These are not real data but constructed to illustrate some basic principles of analysis of patterns in time in a brief amount of space. Basically the task of an observer is to construct a case history of a student in an instructional context using the classifications and categories relevant to the theory under investigation. Whenever there is a change in state in one or more classifications, the observer records the time and the new categories which describe the changed situation[21].

[21]This task can be facilitated greatly if entries are made directly into a portable computer system which automatically records the time whenever a category change is observed and entered.

An APT score is a set of parallel timelines. When represented on paper as in the appendix each horizontal timeline increases from left to right, and the various classifications correspond to the rows of timelines.

For example, the observer began the case history of the first student on December 14, 1991 at 10:30 a.m. by testing and interviewing the student. At that time the observer determined that student motivation was positive, relevant prior knowledge/experience was adequate, and the student had not mastered the instructional objective. The 'null' code was used to characterize states in classifications that were either not relevant at the time or unobservable.

Two days later at 8:25 a.m. the observer began characterizing a concept lesson taught to a class containing this student. The goal was for a student to remember a concept and the instructor was using strategy 1.

The observer recorded the particular instructional tactics used and this particular student's tactical performance (below the dotted line on the APT score). For example, at 8:25:15 the concept was presented expositorily as a generality. There was no secondary presentation elaboration at that time and the student was paying attention (watching/listening). At 08:27:05 the teacher began to present examples and non-examples (instances). At 08:34:52 the teacher began to ask students to classify further examples and non-examples. It can be seen that this student participated actively by giving two correct and one incorrect answers during the practice session, and also paid attention to questions, other student answers, and to instructor feedback during the lesson.

On the next morning, this student took a mastery test on the previous day's instructional objective. She also indicated positive affect towards the experience on a brief rating scale. Cases 2 and 3 were other students on whom case histories were kept at the same time. It can be seen, for example, that the second student was not motivated to learn, had inadequate prior knowledge, was off-task much of the time during the instruction, failed the test the next day, and was ambivalent about the instruction.

The last three cases were begun about a month later, when the teacher tried instructional strategy 3, which is expository only and does not provide practice with feedback as does strategy 1. There were no observations of strategy 2 in these 6 cases.

The reader should note that the case histories in the appendix are not completely visible, due to the limitation of the width of each page. Imagine that the page is like looking through a window down onto a set of parallel conveyor belts running at a constant rate, which move information from right to left. The movement of the conveyor belts is analogous to the passage of time, and the objects riding along at various points represent the category changes in each classification as they were recorded by the observer(s) who contributed to the construction of the case. Thus, the APT score represents the temporal configuration of observed changes in various classifications relevant to the theory under investigation.

Each APT score is a characterization of an *individual* student's history with respect to his/her changing states of knowledge, motivation, feelings and experiences of instructional methods and tactics. As such, an APT score is qualitative knowledge about a specific student's learning experiences.

An APT score is constructed through whatever observation techniques are appropriate. Sometimes systematic observation is needed, for example, to verify that an instructional strategy is being used by a teacher and to see how students are responding and participating. At other times, students may need to be observed or tested under appropriate conditions to see what they currently know or if they have learned something. Individual interviews, questionnaires, rating forms, etc. may also be needed to get at other information, which is then characterized at the appropriate temporal locations on the APT score for a specific student.

Individual APT scores can be valuable during development and revision of instructional theory. While generalizations are not appropriate unless a sufficient number of relevant APT scores is accumulated, an inquirer can nonetheless see patterns emerging even from a few APT scores. As an instructional theory is modified, new classifications might be added which are thought to be relevant. Within existing classifications categories might be added or modified. Definitions of meanings of categories and classifications may need clarification if there is ambiguity when observers are constructing APT scores[22].

Finally, an APT score is a description of a student's learning experiences in a naturally occurring educational system. While APT could be used to characterize contrived experiments, this is not really necessary most of the time. APT is not intended for causal explanations[23] but for description of temporal patterns. In praxiological theory development, causal explanations are not needed; rather the concern is that the theory works in practice. The Wright brothers did not need to explain *why* their invention of an airplane succeeded in flying, but only demonstrate that it did so repeatedly and consistently.

Generalization from APT scores. It is hazardous to make generalizations if only a few APT scores are constructed from observations of educational systems in order to verify an instructional theory under investigation. The Wright brothers had to show that their plane flew more than once or twice. It took numerous replications of flying before they could feel confident that their design was verified — and they were working with deterministic systems.

While the problem is more complicated when dealing with stochastic systems, this does not obviate the need for repeated observations in order to make generalizations. Recall the stochastic vending machine in the previous section. It took a sizable number of observations before we could be confident that the probabilities of various temporal patterns were relatively stable (e.g., if you push the Coke button, the probability of getting a Coke is .20, a Pepsi is .70, and an empty cup is .10; whereas if you push the Pepsi button, the probabilities of getting a Coke, Pepsi or empty cup are .50, .45 and .05 respectively).

Making queries of APT scores. In verifying an instructional theory we need to look for recurrence of condition-method-outcome patterns. Two queries are illustrated here and results are based on the data in the six APT scores in the appendix.

---

[22]See Frick and Semmel (1978) for issues regarding observer agreement and reliabilities of classroom observational measures.

[23]See Frick (1990) for further discussion of APT and causality. Causal inferences are not warranted in APT unless special conditions exist.

QUERY 1		CASE RESULTS					
		1	2	3	4	5	6
	Phrase 1:						
If							
	STUDENT LEARNING ACHIEVEMENT is **nonmastery**	**hit**	hit	**hit**	**hit**	**hit**	**hit**
and	CONTENT TYPE is **concept**	**hit**	**hit**	**hit**	**hit**	**hit**	**hit**
and	STUDENT BEHAVIOR TYPE is **remember**	**hit**	**hit**	**hit**	**hit**	**hit**	**hit**
and	STUDENT MOTIVATION is **positive or neutral,**	**hit**	miss	**hit**	**hit**	**hit**	miss
	Phrase 2:						
then	**if**						
	INSTRUCTIONAL STRATEGY is **strategy 1,**	**hit**	NA	**hit**	miss	miss	NA
	Phrase 3:						
then							
	STUDENT LEARNING ACHIEVEMENT is **mastery**	**hit**	NA	**hit**	NA	NA	NA
and	STUDENT AFFECTIVE ACHIEVEMENT IS **positive?**	**hit**	NA	miss	NA	NA	NA

QUERY 2		CASE RESULTS					
		1	2	3	4	5	6
	Phrase 1:						
If							
	STUDENT LEARNING ACHIEVEMENT is **nonmastery**	**hit**	hit	**hit**	**hit**	**hit**	**hit**
and	CONTENT TYPE is **concept**	**hit**	**hit**	**hit**	**hit**	**hit**	**hit**
and	STUDENT BEHAVIOR TYPE is **remember**	**hit**	**hit**	**hit**	**hit**	**hit**	**hit**
and	STUDENT MOTIVATION is **positive or neutral,**	**hit**	miss	**hit**	**hit**	**hit**	miss
	Phrase 2:						
then	**if**						
	INSTRUCTIONAL STRATEGY is **strategy 3,**	miss	NA	miss	**hit**	**hit**	NA
	Phrase 3:						
then							
	STUDENT LEARNING ACHIEVEMENT is **mastery**	NA	NA	NA	miss	**hit**	NA
and	STUDENT AFFECTIVE ACHIEVEMENT IS **positive?**	NA	NA	NA	NA	miss	NA

A query specifies a temporal pattern to search for in the database of APT scores. Queries consist of phrases. Each phrase is evaluated against a particular APT score (case). If a match is found for a phrase, then it is counted as a 'hit.' Otherwise, it is counted as a 'miss.' All *prior* phrases must be true in order to evaluate a *subsequent* phrase as a hit or miss. And for an entire phrase to be a hit, all its phrase segments must be hits. When scanning Case 1, for example, in query 1 the first phrase is a hit because all four segments are hits. However, the first phrase is a miss when scanning Case 2 because the fourth phrase segment does not match the recorded category (student motivation was negative in Case 2). Once a miss is encountered, a case is considered 'not applicable' to the remainder of a query.

Query 1 was verified in the APT scores in 1 out of 2 relevant cases ($p = 0.5$). Cases 2 and 6 became irrelevant because the first phrase was a miss, and Cases 3 - 5 were rejected because phrase 2 was a miss (different instructional strategy). Query 2 was verified in the APT scores in 0 out of 1 relevant case ($p = 0.0$).

Queries 1 and 2 are identical except for the second phrase — strategy 1 versus strategy 3. *If our sample of relevant cases in the APT scores were sufficiently large*, then we would conclude that strategy 1 was more effective than strategy 2 when conditions are that students do not previously know a concept, their motivation to learn is positive or neutral, and the goal is to remember concepts. Strategy 1 resulted student mastery of the learning objectives and positive affect more often than did strategy 2. However, strategy 2 took 33 percent less time than did the first strategy (10 minutes versus 15 minutes).

Sampling and generalizability from APT scores. Normal considerations of representativeness in survey research apply also to APT when generalizations are of concern. Random sampling of cases from the relevant universe of possible cases is desirable. For example, if the instructional theory is to apply to elementary students, then students should be selected to represent corresponding age and grade levels. If the theory is to apply only to learning of mathematics, then APT scores should be obtained accordingly.

As in survey research, confidence bandwidths can be determined for probability estimates. For example, a Gallup Poll might find that 80 percent of the American people do not favor a tax increase, with a margin of error of 3 percent. This would normally mean that we would be 95 percent confident that the actual percentage is somewhere between 77 and 83 percent. The same notion can be applied to results from queries of APT scores. In the above example the margin of error is too great to conclude that one strategy is more effective than another. We would need to increase the sample size until the confidence band for a probability estimate no longer overlaps another estimate. How large the sample needs to be depends on the size of the difference in probabilities of the temporal patterns.

APT Scores and Expert Systems for Instructional Design

The reader may have noticed that the *if-then* syntax of APT queries is very similar to that used in formation of rule-based expert systems. Expert systems in general are decision aids used in problem solving. Expert systems have been developed, for example, to help physicians identify types of bacterial infections, to assist in the ordering of components for computer systems, for making decisions about where to drill for oil, for assisting underwriters in making insurance policies, and for diagnosing causes of equipment failures (cf., Winston & Prendergast, 1984; Neapolitan, 1990).

The use of APT results is likewise possible in an expert system to aid instructional designers and teachers. Such a data-based expert system could make recommendations on which instructional strategies and tactics would be most likely to succeed, given that a user specifies the *conditions* that are applicable to the problem at hand. In other words, the expert system would be the vehicle to translate the stochastic instructional theory and provide the decision maker with relevant probabilities.

As an analogy, recall the stochastic vending machine described above. If an expert system were devised for this, based on APTs of its observed performance, then the expert system would tell you which button to press if you first tell it what you want. If you want a Pepsi, then the expert system would advise you to press the Coke button because the odds of getting a Pepsi are 7 to 3, whereas the odds are disfavorable (9 to 11) if you press the Pepsi button.

With today's computer technology it would be possible to construct a large, centralized database of individual APT scores, such as those illustrated in the appendix, if instructional researchers were to coordinate their efforts. If consistent terminology and measurement procedures were used in creation of the APT scores, then it would be possible to query the database as it grows over time. The queries and resultant probabilities would become input as rules used by the expert system for instructional design.

While this is still a dream, it is nonetheless a feasible way to create praxiological knowledge of instructional strategies and tactics. It would require the collective effort of numerous instructional researchers and the cooperation of teachers and students in many different places.

Likewise, those who develop computer-assisted instruction could contribute to the cause. They would need to build into their courseware information which describes conditions, methods and outcomes as the courseware is used by students. If properly designed, these characterizations could be used to automatically construct APT scores as output. Those scores in turn could be sent by electronic mail to the centralized database.

This may all sound fantastic. It is. And it can be done with today's computer technology and electronic networks. All we need to do is work together, and we can advance the development of pedagogical knowledge — simply by using the method of inquiry. It's the right one, baby — uh-huh.

References

Bronowski, J. (1973). *The ascent of man.* Boston: Little, Brown and Co.

Bruner, J. (1990). *Acts of meaning.* Cambridge, MA: Harvard University Press.

Cronbach, L. & Suppes, P. (1969). *Research for tomorrow's schools: Disciplined inquiry for education.* Toronto: Macmillan.

Farmer, T. (1989). *A refinement of the ARCS motivational design procedure using a formative evaluation methodology.* Bloomington, IN: Doctoral dissertation, Indiana University Graduate School.

Frick, T. (1983). *Non-metric temporal path analysis: An alternative to the linear models approach for verification of stochastic educational relations.* Bloomington, IN: Doctoral dissertation, Indiana University Graduate School.

Frick, T. (1990). Analysis of patterns in time: A method of recording and quantifying temporal relations in education. *American Educational Research Journal, 27*(1), 180-204.

Frick, T. (1991). *Restructuring education through technology.* Bloomington, IN: Phi Delta Kappa Educational Foundation.

Frick, T. (1992). Computerized Adaptive Mastery Tests as Expert Systems. *Journal of Educational Computing Research, 8*(2), 187-213.

Frick, T., Plew, G. and Luk, H.-K. (1989, March). EXSPRT: An Expert Systems Approach to Computer-Based Adaptive Testing. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Frick, T. & Semmel, M. (1978). Observer Agreement and Reliabilities of Classroom Observational Measures. *Review of Educational Research, 48*(1), 157-184.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

Maccia, G. (1974). Epistemological considerations of educational objectives. *Proceedings of the XVth World Congress of Philosophy,* Varna, Bulgaria.

Maccia, G. (1987). Genetic epistemology of intelligent natural systems. *Systems Research, 4*(1), 213-218.

Maccia, G. (1988). Genetic epistemology of intelligent systems: Propositional, procedural and performative intelligence. Paper presented at the Fourth Annual Conference on Systems Research, Cybernetics and Informatics, Baden-Baden, West Germany.

Merrill, D. (1983). Component Display Theory. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: An overview of their current status*. Hillsdale, NJ: Lawrence Erlbaum.

Neapolitan, R. (1990). *Probabilistic reasoning in expert systems*. New York: Wiley.

Peirce, C. S. (1958). *Selected writings: Values in a universe of chance*. New York: Dover Publications.

Peirce, C. S. (1977). *Complete published works*. Greenwich, CT: Johnson Associates.

Popper, K. (1963). *Conjectures and refutations: The growth of scientific knowledge*. New York: Harper and Row.

Reigeluth, C. (1983). *Instructional-design theories and models: An overview of their current status*. Hillsdale, NJ: Lawrence Erlbaum.

Reigeluth, C. (1989). Educational technology at the crossroads: New mind sets and new directions. *Educational Technology Research and Development, 37*(1), 67-80.

Rieth, H. & Frick, T. (1983). *An analysis of the impact of different service delivery systems on the achievement of mildly handicapped children*. (Final Report, Project No. 443CH00168). Washington, DC: Department of Education.

Roma, C. (1990). *Formative evaluation research on an instructional theory for understanding*. Bloomington, IN: Doctoral Dissertation, Indiana University Graduate School.

Ryle, G. (1959). *The concept of mind*. New York: Barnes & Noble.

Salmon, W. (1966). *The foundations of scientific inference*. Pittsburgh, PA: University of Pittsburgh Press.

Salmon, W. (1971). *Statistical explanation and statistical relevance*. With contributions from R. Jeffrey and J. Greeno. Pittsburgh, PA: University of Pittsburgh Press.

Scheffler, I. (1965). *Conditions of knowledge*. Glenview, IL: Scott-Foresman.

Simmons, J. (1991). *Formative evaluation research on an instructional theory for teaching cause principles*. Bloomington, IN: Doctoral dissertation, Indiana University Graduate School.

Snow, R. (1973). Theory construction for research on teaching. In R. M. Travers (Ed.), *Second handbook of research on teaching*. Chicago, IL: Rand McNally.

Steiner, E. (1978). *Logical and conceptual analytic techniques for educational researchers*. Washington, DC: University Press of America.

Steiner, E. (1981). *Educology of the free*. New York: Philosophical Library.

Steiner, E. (1988). *Methodology of theory building*. Sydney: Educology Research Associates.

The Republic of Plato. Translated with introduction and notes by. F. Cornford. London: Oxford University Press, 1941.

Wald, A. (1947). *Sequential analysis*. New York: Wiley.

Weinberg, G. (1975). *An introduction to systems thinking*. New York: Wiley.

Winston, P. & Prendergast, K. (Eds.) (9184). *The AI business: The commercial uses of artificial intelligence*. Cambridge, MA: The MIT Press.

(filed as aera92.apt, November 21, 1992)

APT SCORE FOR AN INDIVIDUAL CASE

Classifications with categories that do not change during the APT score:

Case ID: **1** Student Name: **Susanne S.** Instructor: **J. Shrock**

Classifications with categories that DO change during the APT score: **1991**

	1	2	3	4
DATE is	**12/14**	**12/16**		**12/17**
HOUR:MINUTE is	**10:30**	**08:25**	**08:40**	**11:05**
Condition observations:				
CONTENT TYPE is	**null**	√**concept**	**null**	
STUDENT BEHAVIOR TYPE is	**null**	√**remember**	**null**	
STUDENT MOTIVATION is	√**positive**			
RELEVANT STUDENT PRIOR KNOWLEDGE/EXPERIENCE is	**adequate**			
Strategic Method Observations:				
INSTRUCTIONAL STRATEGY is	**null**	**strategy 1***	**null**	
Strategic Outcome Observations:				
STUDENT LEARNING ACHIEVEMENT is	√ **nonmastery**			**mastery**
STUDENT AFFECTIVE ACHIEVEMENT is	**null**			**positive**

. . . .

Tactical Method Observations:* HOUR: **8

MINUTES:SECONDS is	**25:15**	**27:05**	**34:52**	**34:56**	**35:02**	**35:22**	**36:07**	**36:10**	**36:20**	**36:25**	**36:30**	**36:38**	**36:42**
PRIMARY PRESENTATION FORM is	**expos**		**inqui**		**expos**	**inqui**	**expos**	**inqui**		**expos**	**inqui**		**expos**
PRIMARY PRESENTATION TYPE is	**gener**	**insta**					**gener**	**insta**					
SECONDARY PRESENTATION ELABORATION is	**null**				**cafb**		**hlpfb**			**cafb**			**hlpfb**
STUDENT TACTICAL PERFORMANCE is	**pattn**			**cans**	**pattn**				**cans**			**ians**	**pattn**

APT SCORE FOR AN INDIVIDUAL CASE

Classifications with categories that do not change during the APT score:

Case ID: **2** Student Name: **Bradley G.** Instructor: **J. Shrock**

Classifications with categories that DO change during the APT score: **1991**

DATE is	**12/14**	**12/16**		**12/17**
HOUR:MINUTE is	**10:30**	**08:25**	**08:40**	**11:05**
Condition observations:				
CONTENT TYPE is	**null**	**concept**	**null**	
STUDENT BEHAVIOR TYPE is	**null**	**remember**	**null**	
STUDENT MOTIVATION is	**negative**			
RELEVANT STUDENT PRIOR KNOWLEDGE/EXPERIENCE is	**inadequate**			
Strategic Method Observations:				
INSTRUCTIONAL STRATEGY is	**null**	**strategy 1***	**null**	
Strategic Outcome Observations:				
STUDENT LEARNING ACHIEVEMENT is	**nonmastery**			**nonmastery**
STUDENT AFFECTIVE ACHIEVEMENT is	**null**			**neutral**

....

Tactical Method Observations:* HOUR: **8

MINUTES:SECONDS is	**25:15**	**27:05**	**34:52**	**34:56**	**35:02**	**35:22**	**36:07**	**36:10**	**36:20**	**36:25**	**36:30**	**36:38**	**36:42**
PRIMARY PRESENTATION FORM is	**expos**		**inqui**		**expos**	**inqui**	**expos**	**inqui**		**expos**	**inqui**		**expos**
PRIMARY PRESENTATION TYPE is	**gener**	**insta**					**gener**	**insta**					
SECONDARY PRESENTATION ELABORATION is	**null**				**cafb**		**hlpfb**			**cafb**			**hlpfb**
STUDENT TACTICAL PERFORMANCE is	**offt**		**pattn**					**offt**					

APT SCORE FOR AN INDIVIDUAL CASE

Classifications with categories that do not change during the APT score:

Case ID: **3** Student Name: **Michael R.** Instructor: **J. Shrock**

Classifications with categories that DO change during the APT score: **1991**

DATE is			
12/14	**12/16**		**12/17**
HOUR:MINUTE is			
10:30	**08:25**	**08:40**	**11:05**

Condition observations:

CONTENT TYPE is			
null	**concept**	**null**	
STUDENT BEHAVIOR TYPE is			
null	**remember**	**null**	
STUDENT MOTIVATION is			
neutral			
RELEVANT STUDENT PRIOR KNOWLEDGE/EXPERIENCE is			
adequate			

Strategic Method Observations:

INSTRUCTIONAL STRATEGY is			
null	**strategy 1***	**null**	

Strategic Outcome Observations:

STUDENT LEARNING ACHIEVEMENT is			
nonmastery			**mastery**
STUDENT AFFECTIVE ACHIEVEMENT is			
null			**negative**

....

Tactical Method Observations:* HOUR: **8

MINUTES:SECONDS is												
25:15	**27:05**	**34:52**	**34:56**	**35:02**	**35:22**	**36:07**	**36:10**	**36:20**	**36:25**	**36:30**	**36:38**	**36:42**
PRIMARY PRESENTATION FORM is												
expos		**inqui**		**expos**	**inqui**	**expos**	**inqui**		**expos**	**inqui**		**expos**
PRIMARY PRESENTATION TYPE is												
gener	**insta**					**gener**	**insta**					
SECONDARY PRESENTATION ELABORATION is												
null				**cafb**		**hlpfb**			**cafb**			**hlpfb**
STUDENT TACTICAL PERFORMANCE is												
pattn												

APT SCORE FOR AN INDIVIDUAL CASE

Classifications with categories that do not change during the APT score:

Case ID: **4** Student Name: **Jennifer R.** Instructor: **J. Shrock**

Classifications with categories that DO change during the APT score: **1992**

DATE is			
01/10	**01/11**		**01/12**
HOUR:MINUTE is			
10:30	**13:05**	**13:15**	**08:30**

Condition observations:

CONTENT TYPE is

null	**concept**	**null**

STUDENT BEHAVIOR TYPE is

null	**remember**	**null**

STUDENT MOTIVATION is

positive

RELEVANT STUDENT PRIOR KNOWLEDGE/EXPERIENCE is

adequate

Strategic Method Observations:

INSTRUCTIONAL STRATEGY is

null	**strategy 3***	**null**

Strategic Outcome Observations:

STUDENT LEARNING ACHIEVEMENT is

nonmastery			**partial mastery**

STUDENT AFFECTIVE ACHIEVEMENT is

null			**neutral**

....

Tactical Method Observations:* HOUR: **13

MINUTES:SECONDS is

05:21 08:29 13:47

PRIMARY PRESENTATION FORM is

expos

PRIMARY PRESENTATION TYPE is

insta gener

SECONDARY PRESENTATION ELABORATION is

null

STUDENT TACTICAL PERFORMANCE is

pattn offt pattn

APT SCORE FOR AN INDIVIDUAL CASE

Classifications with categories that do not change during the APT score:

Case ID: **5** Student Name: **Benjamin F.** Instructor: **J. Shrock**

Classifications with categories that DO change during the APT score: **1992**

DATE is			
01/10	**01/11**		**01/12**
HOUR:MINUTE is			
10:30	**13:05**	**13:15**	**08:30**

Condition observations:

CONTENT TYPE is

null	**concept**	**null**

STUDENT BEHAVIOR TYPE is

null	**remember**	**null**

STUDENT MOTIVATION is

positive

RELEVANT STUDENT PRIOR KNOWLEDGE/EXPERIENCE is

adequate

Strategic Method Observations:

INSTRUCTIONAL STRATEGY is

null	**strategy 3***	**null**

Strategic Outcome Observations:

STUDENT LEARNING ACHIEVEMENT is

nonmastery	**mastery**

STUDENT AFFECTIVE ACHIEVEMENT is

null	**negative**

....

Tactical Method Observations:* HOUR: **13

MINUTES:SECONDS is

05:21 13:47

PRIMARY PRESENTATION FORM is

expos

PRIMARY PRESENTATION TYPE is

insta gener

SECONDARY PRESENTATION ELABORATION is

null

STUDENT TACTICAL PERFORMANCE is

pattn

APT SCORE FOR AN INDIVIDUAL CASE

Classifications with categories that do not change during the APT score:

Case ID: **6** Student Name: **Kevin R.** Instructor: **J. Shrock**

Classifications with categories that DO change during the APT score: **1992**

DATE is

01/10	**01/11**		**01/12**

HOUR:MINUTE is

10:30	**13:05**	**13:15**	**08:30**

Condition observations:

CONTENT TYPE is

null	**concept**	**null**

STUDENT BEHAVIOR TYPE is

null	**remember**	**null**

STUDENT MOTIVATION is

negative

RELEVANT STUDENT PRIOR KNOWLEDGE/EXPERIENCE is

adequate

Strategic Method Observations:

INSTRUCTIONAL STRATEGY is

null	**strategy 3***	**null**

Strategic Outcome Observations:

STUDENT LEARNING ACHIEVEMENT is

nonmastery	**nonmastery**

STUDENT AFFECTIVE ACHIEVEMENT is

null	**neutral**

....

Tactical Method Observations:* HOUR: **13

MINUTES:SECONDS is

05:21 06:43 08:19 09:24 12:01 13:47

PRIMARY PRESENTATION FORM is

expos

PRIMARY PRESENTATION TYPE is

insta	**gener**

SECONDARY PRESENTATION ELABORATION is

null

STUDENT TACTICAL PERFORMANCE is

pattn offt pattn offt pattn offt