A Proposal for the Collaborative Development of Authentic Performance Technology

Richard E. Clark and Fred Estes[1]

Executive Summary

In a companion article (Clark & Estes, 1999 Proposal for Collaborative Development), we presented evidence that many popular performance interventions do not work and suggested that the reason for this was an over-reliance on craft methods, rather than use of a technology of performance improvement firmly rooted in scientific knowledge. After briefly summarizing the bad news about the success of many performance solutions, we describe a template for the design, development and assessment of authentic, science-based performance technology. Our template is presented as the answer to four key questions: 1) What performance problem are we solving?; 2, What body of scientific research addresses this problem?; 3, How do we tease out the "active ingredient(s)" of successful research interventions and produce a "generic" (T1) performance technology' and 4, How do we translate effective, generic T1 interventions into culturally and organizationally acceptable (T2) performance technologies. Since authentic technologies have to be focused on achieving an organizational goal or solving a problem, we propose a preliminary taxonomy of three general performance problem types: 1) knowledge; 2) motivation to perform and,3) organizational design (structure, policy and procedures) at work. Finally, we briefly describe the type of collaboration needed to make new and authentic performance technology a reality and describe the role the International Society for Performance Improvement (ISPI) can play in the collaborative effort.

Introduction: The Need for Science-based Performance Technologies

In recent years, a number of blue-ribbon scientific groups including the National Academy of Sciences and the National Research Council, have reviewed the quality and results of evaluation and research on some of the most important human performance solutions used and promoted by many ISPI members and others. The reviews identify a huge gap between what we think we accomplish, on one hand, and the picture painted by scientific analyses of our successes, on the other hand. What follows is a brief sample of conclusions reached by diverse scientific groups about a number of commonly used performance solutions:

---

[1]Portions of this article have been taken or summarized from a number of past articles and presentations including Clark & Estes, 1998; 1999; 2000 and Estes & Clark, 2000. Full citations for these articles are in the list of references at the end. Richard Clark is Professor of Educational Psychology and Technology, Mentor of the Human Performance at Work doctoral program at the University of Southern California and CEO of Atlantic Training Inc.; Fred Estes is an Adjunct Assistant Professor in Human Performance at Work at the University of Southern California and head of training design for a Fortune 100 company. Clark can be reached at clark@usc.edu , Estes at FDEstes@aol.com .

▸   Scientific studies estimate that many of our training programs may actually leave people significantly more confused, less able to remember important information and unable to use their work-related knowledge as effectively after training, when compared with their ability before training began (Clark, 1988; Druckman and Bjork, 1994; Druckman et al, 1997;) .

    ▸   More than 60% of organizational change strategies undertaken in companies are quickly abandoned (Druckman et al, 1997).

    ▸   Kirkpatrick (1994) Level 1 (reaction) assessment, currently the most popular method of evaluating performance products, gives the wrong information as often as it gives correct information. Sometimes this wrong information indicates that a product worked very well when it actually made things worse (Clark, 1988).

    ▸   Studies of employee empowerment motivation strategies suggest that while they may have a limited success rate in some organizations, they can have very negative results in others. More distressing is the finding that when more rigorous evaluation designs are used, the positive results of these strategies tend to disappear (Golembiewski and Sun; 1990).

    ▸   After hundreds of studies in every imaginable context, there is strong scientific evidence to support the claim that multimedia, Internet and intranet training provides absolutely no learning benefits beyond that available from older technologies such as manuals and human trainers. Additionally, there is evidence that new instructional delivery technologies are much less cost-effective than older technologies in many, but not all circumstances (Clark, 1994).

    ▸   Studies of research on our most popular performance feedback strategies indicates that about one-third of the time, feedback makes performance worse not better. In another one-third of the cases, performance does not improve with feedback (Kluger and DeNisi, 1998).

    ▸   Experiments that check the "generalizability" or transfer of performance solutions indicate that while some of them work once, they seldom transfer between organizational contexts. Since we seldom evaluate solutions that may have worked for someone else in another organization, most of us are largely unaware of this "failure to generalize" phenomenon (Druckman and Bjork, 1994, Druckman et al, 1997).

    ▸   While scientifically-tested and successful strategies for performance improvement do exist, they are seldom incorporated into many of our most popular performance solutions (Clark & Estes, 1998, 1999; Estes & Clark, in press).

The results reported by scientific reviews is not all negative. We do have successes. Positive examples of authentic human performance technologies were described in Clark & Estes (1998; 1999). They included familiar programs such as the famous children's instructional television

series Sesame Street; new training systems for teaching complex knowledge such as the award-winning system proposed recently by van Merrienboer (1997); and a model for drug abuse treatment and personal behavior change proposed by Prochaska et al (1997). Our purpose for focusing here on what is not working is because it is our experience that most performance technologists are not aware of this data. Our field faces a major problem similar to the performance gaps our client organizations face. Like our clients, our own performance problems require careful analysis and systematic solutions.

Thus, the bad news in performance technology is that many of the interventions we now use are unreliable solutions that are not transferrable between organizational settings. The good news is that a reliable and transferrable authentic technology of human performance is available and can produce stunning results. Scientific knowledge, coupled with the vast experience of organizational consultants, are the only reliable basis for building a technology of human performance and solving the crucial performance problems facing us.

Two recommendations for ISPI
There are at least two ways in which ISPI as a professional organization can now take advantage of recent developments in performance science to enhance the development of authentic performance technology:

First, ISPI should strongly encourage all professional members to systematically evaluate the results of their work by using at least results (level 2) and transfer (level 3) assessment and, where possible, bottom-line (level 4) assessment. Our suggested approach is derived from Kirkpatrick's (1994) four-level model applied to human performance technology solutions. Phillips (1997) spells out a clear procedure for doing this that is both practical and business-results driven. We must be willing to subject our work to assessment or we will stop growing and slip backwards. How can we know whether we have made a difference unless we consistently and scientifically measure our results, as Tom Gilbert (1996) insisted? Only when assessment is firmly entrenched as a foundation of our field will we will have a firm basis to hold a dialogue about the value we add to organizations. A commitment to rigorous, science-based assessment is a necessary foundation for the development of an effective and authentic human performance technology.

Second, ISPI should support the development of a process model of how an authentic performance technology is created – and encourage the use of the model to generate authentic performance technology solutions. The solutions must be systems based and focused on gaps in organizational goals due to knowledge, motivation and organizational design problems. ISPI should not support the generation of basic performance research. Instead, we can draw on the basic behavioral and organizational research being conducted in our universities and by more research-oriented professional associations representing the behavioral and social sciences. ISPI should serve as a common meeting ground for scientists, performance technologists, innovators, artisans, and practitioners who agree to participate in the type of collaboration achieved by the team that developed Sesame Street. The purpose of our collaboration should be the development and systematic validation of new, systemic and authentic performance technologies. Our views about how we can collaborate and a high-level conceptual model of authentic technology follow.

Overview of the Four-Stage, Science-to-Technology Model

The four-stage model for developing an authentic performance technology begins with an attempt to answer four key questions. Each question characterizes one stage in the science-to-technology model we propose:

    1) What problem is being solved by an authentic technology?
    2) What body of research and theory addresses the problem most effectively?
    3) How do we translate research into effective technologies? and,
    4) How do we determine that a technology has "solved" an educational problem?

Stage 1: What problem is being solved by performance technology?
We propose that science-based performance technology design requires first a clear analysis and validation of the problem being solved and a determination to avoid the all-too-common error of "advocating solutions and looking for problems they will solve." Too many performance consultants specialize in performance solutions (for example training, sales strategies, employee empowerment, organizational change) and not in performance problem solving. Although the necessity to analyze and validate the performance problems we solve seems obvious, Rossett and Czech (1998) found distressing evidence that not much analysis occurs. The reasons most often given in their study were "not enough time" and "lack of management support." There is no need to recount here the very important reasons for problem analysis or the hit-or-miss nature (and possible negative effects) of the solutions implemented without careful analysis. Many others have very eloquently articulated this issue (see –add some good references here) and Rossett's new book (1999) lays out how to do analysis and problem definition in "Internet time."
.

Select the Appropriate Body of Research and Theory
We suggest three criteria for choosing research to be translated and applied in an authentic technology. First, choose the most comprehensive, experimentally verified theory that predicts the outcomes (problems to be solved) of interest to the new technology. Second, the independent variables and interventions described in the theory must be morally, ethically and practically acceptable in the target application environments. Third, the theory and related experimental studies that provide evidence for the theory, must permit a description of the "active ingredient" that causes or alleviates the problems being solved in a way that permits "no plausible alternative explanation". Let's examine each of these points in turn.

<u>Choose the Best Theories</u> In the first place, technologists must look for comprehensive theories. In choosing theory, we must go beyond the limited and narrowly focused theories that are often available. For example, if our problem involves adults, research with children may often be relevant. If the problem exists in a school setting, research conducted in the military or business may be relevant. The key to deciding which research applies is the ability to identify common structural features of the problems, even if the surface features differ. In a classic study of problem-solving in physics (Chi, Feltovich & Glaser, 1981), novices classified problems according to surface features, such as grouping problems that involved incline planes in one group and problems with pulleys in another. Experts, on the other hand, classified the problems

according to the basic principles involved in solving them, such as Newton's first law. And if you think this example is only relevant to physics problems, you missed the point.

Thus, we must conduct the widest search possible for the most inclusive and robust theories. We must also avoid the temptation to choose a theory because it fits our bias or reflects our current experiences or beliefs. We must be willing to change our minds if well-designed experiments fail to confirm our prejudices or our experience.

For example, Ford (1992) counts over thirty-two, research-based theories of motivation to learn. He notes that many of these theories overlap. Pintrich and Schunk (1996) describe four different theories of motivational learning goal orientation developed by researchers who have given different construct names to almost exactly the same constructs. Because of this duplication and lack of communication among researchers, we need to be very cautious in picking among large body of theories. In fact, since only one or two motivation theories attempt to incorporate the others (e.g. Ford, 1992) the choice in the motivational area is made easier.

The problem is made more difficult as the number and diversity of comprehensive theories increases. This is certainly the case when one considers research on learning and instruction. The field is ripe with many small and large theoretical efforts (for example, see a discussion by Clark, 1988a). Yet most researchers who have recently attempted to produce design technologies have settled on the comprehensive body of research produced by John R. Anderson (1993). Other bodies of research-based theory might eventually produce authentic technologies that have different strengths because they focus on different learning problems.

Prochaska made the decision to accept all major theories but focus the technology based on those theories on the type of problem that reflected one of the six stages in the cycle of change in addictive behavior. This way he and his colleagues did not have to discard many powerful theories; they merely used them appropriately. The design technology of van Merrënboer is focused on the most effective theories on the learning of complex knowledge.

Moral, Ethical and Cultural Issues. As an integral part of our initial analysis, we must ask whether the problems and solutions addressed by a specific technology will be acceptable to our clients. Robert Heinich (1984) has addressed similar acceptability problems in the educational technology literature. The second issue, our search for practical, moral, cultural and ethical acceptability, requires a personal knowledge of the application context for a technology. One of us had a personal epiphany on this issue in India when we made a remark to an Indian colleague about India's "population problem". We were politely but firmly told that "India does not have a population problem. It has a food supply problem". At that point, we learned at the deepest level that the way a problem is defined in the application setting for a technology determines, in large part, the receptiveness of people in the setting for the solution being developed.

Prochaska wisely did not discard theories that were philosophically acceptable to their advocates and prospective clients. He selected the set of theories that could be used practically to solve at least one of the problems addressed at one or more of the six stages in the cycle of change for addictive behaviors.   He tried to find many small theories that were aimed at each change stage.

He and his colleagues reasoned that at least one or more of the cluster of theories at each stage would be acceptable to clients who were going to use the resulting technology.

Yet, one of our biggest concerns in this area is that many educational technologists have philosophical and belief barriers to accepting any technology based on experimental science and quantitative research. Many of these barriers get expressed as "post modernist" views of science. The postmodernists generally make two points. First, science is not the only way to view reality and that the role of politics and culture often ignored in learning. However, the cycle of technology development we propose explicitly recognizes the role of culture in the design and development of instruction and organizational realities in the implementation of instruction. Secondly, postmodernists claim that science is a flawed methodology, often citing the work of Feyerabend (1975) and other philosophers of science. While it is worthy of note that Feyerabend's logic itself is not without flaw, we think of the scientific method much the way Churchill did about democracy. It may not be a perfect system and may also be applied imperfectly at times, but it is the best system yet devised for testing claims of knowledge about the way the world works. Myhrvold (1998) in a recent issue of Science eloquently presents the potential of science to inform our technologies and strongly urges greater support for basic research

Identifying Active Ingredients for Authentic Technologies. Finally, authentic technology design requires identifying the "active ingredient" or the key element of the intervention that is necessary for success. An operational definition of the generic "active ingredient"is to be found in the independent variables and interventions that have been used in successful experiments that verified the theory being adopted. A quick example can be found among shopper strategies for purchasing the least expensive in the "over the counter" remedies available in drug stores. The active ingredient in aspirin is 5 mg of an acid compound. This ingredient is available in equal amounts, and with equal impact no matter what medium carries it (e.g. gum, tablet, or liquid suspension) or what "brand" it sells under. Thus, all aspirin products that contain the required 5mg. of active ingredient have a more or less equal biological impact, but not the same price.

For example, Clark (1983,1994) argued that the two active ingredients influencing learning from electronic media delivered instruction are the "instructional methods" and the "task knowledge" embedded in instructional frames. A method was defined as "any external representation of a cognitive process necessary for the learning and application of task knowledge of students who could not, or would not, provide the cognitive process for themselves." Task knowledge is defined as the various types of declarative and procedural knowledge necessary to perform the task but unavailable to learners before instruction. The ID technology of van Merrënboer (1998) presents many other important active ingredients.

Identifying active ingredients in research and theory is one of the most creative acts required for technology development. It is a skill that is not well understood or taught and it is seldom modeled adequately. We have been impressed with the discussions about this issue by Reigeluth (1983), Gage (1985) and Landa (1983). Clark (1983; 1988a) described an approach to this

problem for delivery and instructional technologies. Clark (In Press 1998) describes examples of the approach for motivation technologies. Essentially we must be able to identify the generic causal agents that influence the problems we are solving at the deepest, most structural level in order to develop the most effective technology.

For example, a popular craft-based solution to motivational problems in work settings is to provide "employee empowerment". This involves allowing employees to form work teams that make decisions about how they will perform a job. In some settings these empowered teams have been very successful, while in other settings they have failed miserably and expensively. An "active ingredient" analysis of the empowerment intervention suggests that when it succeeds, it does so because those applying it feel that they will gain significant control and become more successful. In some cultures and work settings, being permitted to make decisions about one's job is considered to be a speedy way to fail at the job and therefore is rejected. So the active ingredient in these studies could be called "control beliefs" and defined as any job condition that the individual or team perceives as resulting in increased control or success.
This active ingredient analysis is most obvious to researchers who understand the modern "expectancy-control" theories of motivation (Clark, In press 1998). Without active ingredient analysis, empowerment interventions run the risk of expensive failure at work, in navigation rules for Internet- and computer-based distance education, and in the classroom. Thus, ingredient analysis is also the basis for the generalizability advantage of an authentic technology over a craft. The more we are aware of the active and necessary ingredients that form the basis of the interventions suggested by technologies, the greater the chance that the technology will transfer to new settings, people and varieties of tasks. Identifying active ingredients in instructional treatments is the extraordinary strength of both the Ellen Gagne and van Merrënboer design technologies. Both of these technologists are very clear about the key elements of their interventions.

How Do We Translate Research into Effective Technologies?

Those who have attempted to apply research findings have learned quickly, and sometimes painfully, that such transfers are not direct or easy. Learning about positive correlations between age and height, on the one hand, and learning and time on the other hand, should not lead us to expect that we will necessarily grow taller and smarter if we only wait for time to pass. Yet this kind of twisted logic can be found in many existing instructional media technologies. Because we find learning correlated with the use of new electronic technologies, we wrongly assume that providing more technology will produce learning advantages. We need a way to conceptualize the process by which theories are developed, and active ingredients are identified and translated into technology. Our suggestion is to focus on the four stages in a "science to technology spiral".

INSERT Figure 1 ABOUT HERE
SCIENCE - TECHNOLOGY SPIRAL MODEL

We suggest a four-stage model for describing the process by which science develops and is, in turn, translated into social or behavioral technology. We briefly describe the model and then

apply it to understanding the translations that occur in the two authentic technologies we are using as examples. The four stages that characterize the movement from science to technology we will call: 1.) Descriptive research (S1); 2.) Experimental verification of theory (S2); 3) Generic Technology (T1); and 4) Contextualized Technology (T2). The first two stages are commonly associated with science. The final two stages define technology development.

Stage 1: Descriptive Scientific Research Stage (S1). In stage S1, descriptive research attempts to produce reliable measurement and description of variables and processes of interest. At this stage, very creative new ideas are explored. Nothing is rejected if it can be measured. Construct definition, measurement and hypothesis generation are the key descriptive goals. Both qualitative and quantitative methods attempt to tease out new variables and processes of interest to scientific research. In fact, it is possible that many descriptive research problems result from analyses of craft solutions. Since all craft solutions that are successful in any setting contain elements that might have been sufficient to solve the problem once, many successful technologies have resulted from descriptive reasoning about why the craft solution worked.

Yet no technology results directly and immediately from stage one inquiry. Hypotheses developed here need to be tested at the next stage before they are useful. The survey methods, naturalistic observations, thinking experiments, literature searches for construct validation, path analysis and Liseral or Structured Equation models used at this stage all serve to define key variables and provide evidence for relationships between variables. This leads to important hypotheses that need to be tested and linked into theories at the next stage. Conclusions about the factors that cause a problem cannot be drawn and no interventions or authentic technologies can be developed based on descriptive data.
The research conducted at this state is often very creative and naturalistic. For example, in the 1850's, the Viennese physician Semmelweis noticed that four times as many women died of symptoms called "childbirth fever" in a birth clinic staffed by physicians than in a similar clinic staffed by midwives. He could not describe the cause of the deaths until a physician friend cut himself with a knife he had used for dissecting cadavers and died of symptoms identical to childbirth fever. Semmelweis connected these two events and reasoned that since physicians who delivered babies in their clinic had often been dissecting cadavers just before the delivery, he suggests that "cadaverous matter" may have been causing the deaths through transmission by the physicians (an S1 description and hypothesis). He planned an experiment to validate his descriptive observation. He asked physicians to wash their hands before attending childbirth. Subsequent deaths in the physician's clinic fell to the level of the midwife's clinic. His experiment was an example of science at the S2 stage.

Stage 2: Scientific Experiments and Theory Development (S2). In stage 2, theory development and the experimental verification of theories is carried out through the testing of hypotheses suggested at the descriptive level. Theories are built, checked in experiments and revised. Theoretical paradigms compete for attention. Quantitative and qualitative methodologies cooperate to test the key hypotheses in theories. Those theories that encompass the greatest range of phenomenon and survive experimental tests, continue. Confounding and artifact are eliminated from the typical experimental designs that test competing theories. The unconfirmed or narrow theories should fail but sometimes survive out of the ignorance or ego problems of

researchers and journal editors. Rival hypotheses and alternative explanations for the effects measured in controlled studies are explored and, if possible, eliminated.   Active ingredients that produce measured results are identified. Paradigms shift and change with shifts in evidence and fashion. Advanced theories that survive this stage are ripe for the development of authentic technologies.

Semmelweis's hand washing experiment is a good example of the long history of connections between science and technology. While there was no effective knowledge of micro organisms in the 1850's, Semmelweis's instructions that doctors wash their hands before attending women in childbirth" decreased deaths. He assumed that the active ingredient causing death was "cadaverous matter" and that hand washing would eliminate the problem.

Stage 3: Generic Technology Development (T1). In stage 3, generic technologies are developed based on the strongest and most successful theoretical paradigms that have survived scrutiny at stage 2. This is the state where a generic active ingredient analysis is critical. Technologies must first be developed in a generic or decontextualized state so as not to confuse the issues surrounding the many unique limiting features of the setting or context in which they are being applied. For example, we need first to know that a need to feel in control is the active ingredient that drives motivational commitment to learning goals. Only with then can we introduce the many conflicting types of control issues one finds in different application settings. (The issue of setting-specific translations of the active ingredients is introduced in the next stage). The goal at this stage is to generate a model that explains how to solve the problems that served as dependent variables in prior research.

Reigeluth (1983) notices that the reasoning at this stage is "backwards". Whereas research starts by trying to find a treatment that will "predict" changes in a dependent variable. Technology at the generic stage starts with the dependent or "problem" variable and attempts to incorporate the active ingredient into a generic treatment or intervention that will solve the undesired problem. For example, a researcher will focus on the fact that micro organisms can be the cause of increased deaths following childbirth, while a technologist begins by collecting information about the causes of death following childbirth in order to develop procedures for eliminating or reducing them.

Prochaska's change strategy exists only at the generic T1 stage. He and his colleagues wisely decided to design a generic technology that could be "translated" by psychologists in many treatment or prevention settings. This permitted the addition of local features that increased the cultural acceptability of the treatments without diluting the active ingredient. This is one of the strengths (and some have complained, one of the weaknesses) of his system. Its strength is that the generic, "active ingredient" description of the key elements of the successful interventions at every stage allows all users to translate the T1 ingredients in a way that will make them acceptable to their clients. Yet the clarity of the description of the T1 solutions permits users to make reliable and accurate applications in a great variety of culturally different settings.

Van Merrënboer used a similar strategy for his instructional design technology. He has a generic T1 system that cannot succeed without being procedurally translated for application in any

specific application. He notes that his system "does not ... focus on ready made procedures to support the process of instructional design" (van Merrënboer, 1998, p. 1).   He explains that it is to be "applied in conjunction with an ISD-model in order to receive support for the activities not treated in the model such as needs assessment, needs analysis, production of instructional materials, implementation, delivery and summative evaluation" (p. 3). Presumably, other T1 authentic technologies could be developed to accommodate the additional "activities".

Stage 4: Specific Technology Development (T2). In stage 4, generic (T2) technologies are translated for specific settings, people and tasks. An educational technology that is acceptable and effective in a North American K-12 classroom may not be acceptable in a Chilean K-12 classroom.   An educational technology developed for San Jose, California may not be acceptable in San Pedro, California. The children's television series "Sesame Street" was eventually adopted for use in over 80 nations. In many of those settings, there were controversies about the almost total focus of the program on "cognitive skills" and an absence of "social skills" training. In other countries, the scripts had to be edited and new versions of the program were produced to change elements such as gender roles, situations, relationships and songs to make them culturally acceptable.

At this final stage, the generic T1 technology is shaped and translated for the unique cultural beliefs, expectations, knowledge and value patterns found in a specific application setting. The specificity of T2 technologies vary, depending on the novelty of the cultural expectations and values held by users who will experience the technology. It is also critical to note that only T1 technologies should be transferred between different settings. Since T2 technologies are specific to a context, one would not expect the intact transfer of T2 interventions to a different context to succeed.

Translating generic T1 technologies into specific T2 applications requires a great deal of experience with the culture and expectations found in the target setting. We are not aware of any science that informs this process and so we need to rely on existing craft, systems approaches and on a "trial and revise" correction cycle. Imagine a need to transfer a T1 technology for teaching concepts to a specific setting. It is easy to see how the knowledge and learning strategies people need to learn and perform can be described in a generic way. A concept is defined the same regardless of the context in which it is taught or learned.

Yet critical elements of instructional presentations such as our specific choice of the examples, simulations, and analogies used to illustrate concepts can and should change depending on local constraints. Important among those constraints are teachers and learner's beliefs.   For example, the usefulness of certain types or format's for examples; the value placed on humor or a rejection of humor as "frivolous"; whether instructional decisions are typically made by learners or system managers; and the efficacy beliefs of learners and system managers, name only a few of many constraining beliefs. Much more work remains to be done at the T2 level to insure adequate translation of generic technologies.

Roles and Responsibilities. A few words about the various roles in this spiraling cycle of science to technology may clarify the dimensions of the type of collaboration we are calling for. First,

while no one can juggle all the roles simultaneously, many people play more than one role. For example, practitioners who are responsible for implementing and delivering solutions to clients may function as a technologists at the T2 level while adapting a T1 technology for use in their specific organizations and clients, then shift to a craft role in improvising a part of the solution set where our scientific knowledge is missing or incomplete. So people may move among various roles as their needs, interests, and opportunities allow.

Second, it is clear that all the roles are necessary for optimal solutions of real world problems. Scientists add to our knowledge of the world, but it is not their job to apply that knowledge. Technologists use knowledge as the raw material to forge solutions at both the T1 and T2 levels. Practitioners deliver these solutions in the real world and employ craft to fill in the gaps of our knowledge. Craftspeople are those ingenious people who devise ways to make things work, even if they do not know why it works. For example, before there was a technology of anesthesia, a craft approach to pain relief was to give a patient as much whiskey as they could consume. At that time, whiskey was a responsible and compassionate (if dangerous) intervention. Today, however, when we have a technology of anesthesia based on scientific knowledge, the use of whiskey is not an acceptable practice. We deplore the persistence of craft approaches when the state of our knowledge permits the development of authentic technology.

The process is similar to the application of architecture our buildings. No one would dream of simply erecting a big, expensive building in an urban area without planning and careful architecture (nor would our laws permit it). In architecture, the laws of physics are essential first considerations and will always apply to the structure as long as it stands. The plans also need to reflect what type of building it is to be and how it is to be used an apartment building will be designed differently from a department store or a school and they must contain all the elements to perform their basic function (T1). Finally, the building must also be planned to account for the community, its role in that community, the existing cultures and styles of the community, and the lives of the individuals who will live and work in and around the building, while doing all of this within the tastes, budgets, building codes, and aspirations of the community (T2).

We are not advocating some set of teaching methods which we are labeling "scientific" or some set of content we claim is "scientifically" derived. We do claim that human learning and motivation can be studied through the scientific method and basic principles discovered. Further, the application of these principles to the design and development of instruction, in the context of the cycle of authentic technology, results in more effective instruction and achieves more of the intended goals of the intervention.

How Do We Determine that a Technology has "Solved" an Performance Problem?
Since substantive, multi-dimensional evaluation is seldom performed on educational interventions, the case for authentic technologies is more difficult. If we require adequate evaluation of interventions, the clear superiority of authentic technology over craft will be evident. Too many practitioners experience research as a complex, error filled process where competing claims are voiced constantly. Whereas information about craft solutions are most often based on "personal or organizational experience testimony" where the positive is emphasized and failures are not mentioned. To paraphrase an old saying in the media area,

"The invisible virtues of craft are compared with the very visible mistakes made as we develop research-based authentic technologies".

We are attracted to a combination of formative evaluation (in progress check's accompanied by a revise cycle) and the adoption of a more elaborate version of the Kirkpatrick (1994) four level evaluation model used in business settings for summative evaluation. Kirkpatrick's model permits us to distinguish among 1) motivation (like it?), 2) learning (learn it?), 3) application (use it?), and 4) results (payoff?). This is a check on the needs analysis and problem analysis that began the design cycle in the first place.

Formative evaluation can be performed at all of the four stages in the development of a technology. Outcome evaluation can most persuasively be performed only on T2 technologies but because of the indeterminate nature of the T1 to T2 translations, considerable uncertainty always remains about what has caused any measured gain (or lack of gain) at any of the four Kirkpatrick evaluation levels. It is discouraging that so little evaluation is performed on any educational innovation. This pattern makes it difficult to compare the effectiveness of authentic technologies with craft solutions. When evaluation is performed, it is often badly designed and focused mainly on reactions (which can be inversely related to learning, behavior and results) and secondarily on learning (and most learning evaluation is focused on memory for facts). We seldom evaluate transfer or results, and yet our goals are almost always to achieve transfer and to solve the problem that began the development process in the first place.

The Partnership Needed to Develop Authentic Technologies

One of the important implications of the Spiral Model of Science and Technology is the need for the collaboration of practitioners, technologists, and scientists. We feel that much of the controversy in this area results from two mistaken beliefs. One is that a good professional should do it all from basic, descriptive research to implementing the polished T2 solution set. The second is that some roles in this cycle are inherently more worthy of respect than others.

The first belief is a common, yet impossible problem. No one person can perform all the roles in the cycle simultaneously. First, the demands and opportunities of the organizations we work in put us all in different situations. Universities afford time for research and reflection; school classrooms and business settings demand swift, effective action, and consultancies provide an overview of the same problem in many settings and the bridge between theory and practice. Second, we all have different strengths and skills as a result of our aptitudes, interests, and focused experiences. The special skills and instincts we develop that make us so good in one role may be counterproductive in another role. A dispassionate, deeply analytical, and reflective approach will serve the researcher very well, but will be fatal to the classroom teacher and business trainer who need to respond fluidly and immediately.

This complexity suggests that the many types of expertise required to produce authentic educational technologies development will require team efforts. Highly collaborative teams composed of researchers, designers, and representatives of the target audience will most likely be

more successful than individuals who try to represent very different and diverse experience. Prochaska, and van Merrënboer discuss their own collaborations with colleagues possessing different but complementary knowledge.

The second misconception is related and deeply troubling. Unless we understand the roles of our collaborators in different phases of the cycle and can develop true respect for their contributions and the constraints they work under, we will not be able to collaborate effectively. Practitioners are not less worthy or less valuable than researchers; they are not simply people who could not make it in research. Researchers are not impractical nerds doomed by their timid souls to spend their cloistered lives proving the obvious at great expense to the taxpayers. Each role has a unique contribution to make and a need to partner with all the other roles to add to our knowledge and to solve real problems.

Role for ISPI

ISPI can play a crucial role in this cycle of authentic performance technology development. While there are many societies that cater to the needs of one quadrant of the cycle, ISPI is best positioned to serve a forum and meeting ground for technologists who design generic technologies (T1) and practitioners who design, develop and deliver interventions for real clients in the real world (T2). A key point we make is that T1 is a vital linking role in developing effective performance interventions, yet there is little institutional support for this role. Universities focus on basic research while schools and industry focus on near-term delivery of solutions. There are few accolades for T1 generic technology, sometimes called "applied research,"we believe because most people do not understand the importance of this linking stage. T1 technology provides a stepping stone that enables scientific knowledge to cross the chasm into practical application; lacking this stepping stone, the gulf is too broad and that is why we see so little science in our interventions despite our great scientific advances.

Specifically, ISPI can promote a model such as this one, which shows the relationship of science, technology and practice and insist on mutual collaboration and respect. Through its research and publications, insist on clearly identifying and differentiating between technology based in science and craft, and promoting the replacement of craft when there is a clearly superior technology. Along with this comes the need to promote and insist on high standards of analysis and evaluation in both technology development and technology application.

Summary and Conclusions

We are proposing collaboration between educational researchers and craftspeople on the development of authentic performance technologies. We believe that only with a commitment to authentic technologies can we hope to solve persistent problems with the utility and relevance of our performance technology interventions. We argue in this article, and an earlier discussion (Clark and Estes, September-October, 1998) that most of our current work can be described as limited and non-generalizable craft. We define an "authentic educational technology" as educational solutions resulting from a systematic analysis that identifies the problem being solved, selects and translates appropriate, well-designed research and applies it to design culturally appropriate educational solutions. While educational technologists want to take advantage of the extraordinary advances over the past two decades in the social and behavioral

sciences in order to solve pressing educational problems, we show how that the lack of support for the design and development of generic technologies (T1) hinders that transfer of knowledge.We propose a role that ISPI could play in supporting T1 technology and several specific recommendations.

We are deeply aware that space considerations required us to give only cursory descriptions of ideas and approaches that are very complex and, in a few cases, controversial. Our purpose in this article is to shift the direction of our dialogue a bit. We invite questions, suggestions and alternative views. We only hope that we can be constantly clear about the criteria we will use to determine which ideas will prevail. Our criteria emphasize measured gains that reflect a solution to the problem targeted by our technology. We know that this discussion is not finished. In our next installment, we will describe some of the replies and ideas that have been proposed about these ideas.

References

Chalmers, A.F. (1999). What is this thing called Science? Indianapolis, IN: Hackett Publishing.

Chi, M. T. H., Feltovich, P.J., & Glaser, R. (1981). *Categorization and representation of physics problems by experts and novices.* Cognitive Science, 5, 121-152.

Clark, R.E, & Estes, F. (1999). *The development of authentic educational technologies.* Educational Technology, 39(2), 5-16.

Clark, R.E., & Estes, F. (1998). Technology or craft: What are we doing? Educational Technology, 38(5), 5-11.

Ford, M. (1996) Motivating Humans. LEA

Myhrvold, N. (1998). *Supporting science.* Science, 282(23), 621-622.

Rossett, A. (1999). First Things Fast.

Rossett, A., & Czech, C. (1996). *They really wanna but...The aftermath of professional preparation in performance technology.* Performance Improvement Quarterly, 8(4), 114-132.