Developing Teachers' Decision-Making Strategies for Effective Technology Integration: A Simulation Design Framework

Ron Stevens, Ph.D.
Professor of Microbiology and Education
UCLA
Director of the IMMEX Project

Sara Dexter, Ed.D.
Research Associate,
University of Minnesota,
Director, eTIPS Project

Symposium Presentation
AERA National Meeting, 2003
Division K: Teaching and Teacher Education
PT[3] National Initiatives: Facilitating Change in Assessment and Accountability for 21st Century Teacher Quality

April 2003

Keywords: simulations, cognitive frameworks, problem solving, decision-making, strategies.

Summary
Rapid advancements in hardware, software and connectivity are helping to shorten the times needed to develop computer simulations for education. These advancements however have not often been accompanied (or driven) by corresponding theories of how to best design and use these technologies for teaching, learning, and testing. Such design frameworks ideally would be guided less by the strengths/limitations of the presentation media and more by cognitive analyses detailing the goals of the tasks, the needs and abilities of students, as well as the structure of the knowledge base in the discipline.

This article first describes a theoretical framework for the IMMEX™ problem-solving software that provides the development and delivery environment for the eTIPS PT^3 Catalyst Project. This software has been used extensively for investigating how students from elementary school through medical school select and use strategies as they solve complex science problems. The framework developed shows how the structure of the knowledge base in science, the needs of the students, and the learning/assessment goals have shaped the design criteria, tasks and the selection of the analytic tools for tracking student progress.

From this established perspective, the structure of the education knowledge base(s) is examined as it applies to the development of school-based simulations around disciplinary themes (with classroom integration of technology as an example). By contrasting the structural differences between the education and science knowledge bases characteristics are revealed that influence simulation design and/or shape the student performance and progress outcomes that could be expected to be derived from school-based scenarios.

Introduction
Understanding and learning, whether about science or science education, technology integration issues, or teaching in general, requires and follows active involvement (Spillane, Reiser & Reimer, 2002). This is true for students learning science, for pre-service teachers learning to teach, or in-service teachers and educators implementing education reform programs in the

schools. Simulations[1], appropriately designed, implemented, and evaluated, can create environments for active involvement that can promote learning by including many opportunities to solve problems and examine phenomena across contexts and perspectives in engaging and challenging ways (Bransford et al, 1999). They can provide experiences that are too risky or expensive to otherwise implement (Lane et al, 2001), and can support decision-making, systems thinking and perspective taking that are hard to provide in other ways (Resnick, 1994). With the increased use of simulations for learning, it is also likely that the dimensions of assessment will begin to change and the data collection and reporting capabilities of simulations will greatly enhance an educator's ability to personalize learning and provide incremental, targeted (and most likely immediate) formative and summative feedback to multiple audiences (Pellegrino, et al, 2001).

While simulations hold great educational promise, the discipline of simulation creation and use for the most part remains fragmented with most researchers developing for, and publishing the results of their studies in, disciplinary subfields (e.g. business, medicine, law, and education) (Bos, 2002). Thus, while guidelines and theories exist for presentation of information on computer screens (Mayer, et al, 2001), there are fewer guidelines regarding the design process of simulations that cross disciplines. The U.S. Department of Education's program for Preparing Tomorrow's Teachers to Use Technology (PT^3) has provided a unique opportunity for researchers from science and science education to collaborate with pre-service educators to extend the simulation environment of science problem-solving software into classroom-based technology integration scenarios. In this paper, rather than discussing the specific details of the software environments, the authors focus on design pattern and design detail differences between science and education simulations that begin to provide cross-discipline frameworks for simulation construction.

[1] The term simulations can broadly refer to games, models, and immersive and manipulated environments, in computer and non computer environments. Simulations, as referred to here, are online scenarios where participants adopt a functional role in a simulated environment where a problem exists and search and data interpretation is required to resolve the issue(s).

IMMEX Project Characteristics

For the past 14 years, the IMMEX[2] Project has been using emerging technologies and simulations to personalize learning, first in medical education, and subsequently in K-20 educational activities (Stevens, Kwak & McCoy, 1989, Palacio-Cayetano et al, 2000). From a software design perspective, the IMMEX Project has the following goals and approaches.

1) Long-term Approaches/Contributions to Theory: *Extract Behaviors from Sequences of Intentional Actions*

The project's long-term approach has been to develop software systems for reliably extracting useful information about reasoning from sequences of intentional actions. These sequences have been clickstream data recorded from students' actions while navigating computer-based simulations[3], but they could be any form of sequential information. To reliably and rapidly derive and report this information, artificial neural networks have been our analytic tool of choice (Stevens & Najafi, 1993, Casillas et al, 2000). Using these systems we have shown that we can 1) distinguish novice/expert differences in problem solving (Stevens, Lopo & Wang, 1996), 2) build valid models of student problem solving without a priori rubrics, or other indicators of what is a "correct" performance (Vendlinski & Stevens, 2002) 3) reveal the close cognitive relationship between what students do and what they say (Chung et al, 2002), and 4) report models of student understanding in near (and soon to be) real time (Stevens, et al, 2001). A key to the success of this approach has been the development of search path mapping analytic tools that reveal the underlying structure of the models (Stevens, 1991).

2) The Problem Solving Model: *Hypothetical-Deductive Learning*

The details of problem solving are often described in terms of the hypothetical-deductive learning cycle where explanations of phenomena are investigated and refined[4]. This cycle is derived from a late stage of intellectual development characterized by students beginning to engage in combinatorial thinking, identification and control of variables, proportional thinking, probabilistic thinking and correlational thinking. From a cognitive perspective, situations

[2] Interactive Multi-Media Exercises, © The Learning Chameleon, Inc, 1991.

[3] In this paper we refer to specific simulations, problems or cases and these are equivalent terms. The term imulation draws from the computer science literature, problems from the cognitive science literature and cases from the case-based reasoning literature.

[4] For a historical perspective on learning cycles, see Lawson, 1995 pp 155-169.

requiring hypothetical-deductive thought often involve a starting condition, a goal condition, and resources to transit between these two cognitive states. In most situations this is an iterative process where intermediate goals (hypotheses) are confirmed/rejected based on the latest information available. If a student were pursuing a particular hypothesis or line of reasoning, the goal of acquiring additional information would be to increase the confidence in the validity of this reasoning chain. Conflicting data, if obtained, would instead decrease the confidence in the current hypothesis and result in the initiation of a modified search of the problem space.

3) *Strategies* - The Supporting Theoretical Design

An important aspect of the hypothetical-deductive model is that students engaged in such activities continually select and revise strategies to optimize the outcomes. Strategies, whether successful or not, are aggregates of multiple cognitive processes including comprehension of the material, search for other relevant information, evaluation of the quality of the information, drawing of appropriate inferences from the information, and the use of self-regulation processes that help keep students on track. Documentation of students' strategies at various levels of detail can therefore not only provide evidence of a changing understanding of the task, but can also provide experimental evidence of the relative contribution of different cognitive processes to the strategy. Strategies used by students can then become a phenotype, or a proxy so to speak, of the working state of a student's knowledge.

The theoretical design the IMMEX Project uses for investigating students' selection and use of strategies during scientific problem solving is based on extensive work by others (VanLehn, 1996, Schuun & Reder, 2001, Schuun et al, 2001, Haider & Frensch, 1996) and can be organized around the following principles:

- Principle 1: Each individual selects the best strategy for them on a particular problem and individuals might vary because of learning in the domain and/or process parameter differences;
- Principle 2: People adapt strategies to changing rates of success;
- Principle 3: Paths of strategy development emerge as students gain experience; and,
- Principle 4: Improvement in performance is accompanied by an increase in speed and a reduction in the data processed.

4) Simulation Design Frameworks

The following sections describe how these strategic contexts along with the experiences of over 55,000 students have shaped the design features of IMMEX and have helped to maximize the information we can derive about student understanding with these simulations. In each section the IMMEX-centered approach is followed by alternative design considerations for applying similar principles to school-based scenarios that take into account differences in the structure of the knowledge bases for science and education. These considerations draw on the experiences derived over the past two years from the eTIPS[5] PT[3] Catalyst Grant, a project that focuses on decision making around the challenges of appropriately integrating technology into the classroom (Dexter, 2001), and with the University of Minnesota, IMMEX, and the Vermont Institute for Math and Science Teaching as partners.

a. Relate the Embedded Content to the Discipline Knowledge Base(s)

Scientific problem solving builds on established knowledge bases and involves the accurate conceptual linking and manipulation of information. From a design standpoint, this requires that each piece of information in a simulation must be valid, and consistent with the contexts of every other piece of information in the problem (we term this causal coherence or consistency). Such causal coherency reinforces the organization of complex concepts through the student-directed manipulation of the artifacts, and it appears, in fact, that this is an important component of student learning and reasoning (Chen, et. al, 2001). Aside from student and teacher accounts, the evidence of this learning includes strategic improvements relating to paths of strategy selection, use, and refinement over time, which is provided by the sequence of item selections as students solve or miss the problem. Causal coherency is generally easy to establish and validate where a concrete knowledge base exists, such as in science. In education, however, the knowledge base is somewhat different. A wealth of concrete knowledge exists that is represented by the formal research literature, and this can be drawn upon for constructing simulations. A second knowledge base core also exists, that is more dynamic and emergent, and is being called practitioner knowledge (Hiebert, Gallimore & Stigler, 2002). Practitioner knowledge is closely linked with the practice, is highly contextual, and is integrated around the problems of the

[5] Educational Technology Integration Principles

practice. It is more socially grounded than the research database and influenced by values and beliefs that shape perceptions of a scenario, i.e. they are experience/belief-driven practices.

This structure of the practitioner knowledge base is likely to have significant influence on the learning model around which to base the school simulation designs, shifting them away from problem solving and more towards decision-making. In decision making, cause and effect beliefs (which can play the role of conceptual linkages between information items in science simulations) can shift the criteria (the items of information or "lab tests" in the science simulations) by which decision makers evaluate alternatives (goals or solutions) as well as the alternatives actually being sought.

At one end of the decision-making continuum, there are few differences from the hypothetical-deductive learning cycle (Gredler, 1992). For instance, solving a problem by the efficient and effective gathering and interpretation of information is nearly identical to Explanation-Based Decision Theory where conditionally dependent pieces of evidence are analyzed prior to choosing a decision option. With these similarities, the tools and theoretical approaches from the science-based IMMEX perspective would apply nicely.

This is quite different, however, from the Expected Utility Theory of decision-making that assumes humans will make choices, which maximize personal utility or the Conflict Theory of decision-making where a decision of some consequence to the person must be made. Personal utility and consequences for teachers could include a host of options, involving substantial components of beliefs that could be explored in the context of school-based simulations.

Irrespective, the simulation designer would need to define at the outset, in clear terms, what understanding(s) they hope to derive about student behavior before the construction of the problem space begins (Mislevy, et al, 1999). Is the understanding to be derived mostly evidence-based, asking questions like "Did the decision-making construct encourage the inclusion of all relevant data? Did it encourage the consideration of alternatives? Was the most accurate cause/effect information evaluated?" If so, then perhaps the problem spaces should be more research oriented.

Alternatively, if the focus of understanding is more on the beliefs that helped shaped a decision then the questions become: "What is the nature of these beliefs? How do they manifest themselves in the gathering and interpretation of information? Are these beliefs that we wish our students to have, and if not, how strong are they and what are the interventions needed to change them?" With these central questions, the problem spaces should contain more practitioner-related knowledge.

Finally, can a single simulation encompass both poles of the decision-making continuum, i.e. investigative and experience/belief directed? With careful thought regarding the design, probably, and while examples exist in the IMMEX science problem-solving series (Scott, 1994), they do not currently exist in school-based scenarios.

b. Construct Challenging and Varied Problem Spaces

However, the Learner-Centered Psychological Principles also emphasize motivational and affective factors that also affect learning (http://www.apa.org/ed/lcp.html). These, along with the cognitive complexity issues discussed below, shape the presentation of information for each item on the screen, within a case scenario, or across a series of cases in a problem set[6]. There are two major variables that are considered here: the reality of the simulation/information, and the value that each item possesses relative to the domain outside the simulation.

The reality (or fidelity) of the simulation can be broadly grouped as real, realistic or fictitious. In science, instructional activities that would be considered as real would be hands-on, inquiry-based activities, whereas in pre-service education, reality exists in the classroom induction experiences. There are places and allowances for each degree of reality group within a simulation; the critical design feature is to decide the balance. For instance, Figure 1 portrays a flame test for sodium in a chemistry simulation. The data behind this information item is real, the portrayal is realistic and the simulation from which this is derived (Hazmat) is fictitious.

[6] A problem set is a series of simulations that share the same problem space, yet contain different data and possible solutions to the case. The *Hazmat* simulation described above has 39 different cases in the problem set which provide students multiple problem-solving experiences and supplies progress data (Section 4.4). Problem sets do not currently exist for *eTips* simulations as each case is its own entity. There are, however, 48 of these simulations and teachers can sequence them as desired.

In the eTIPS simulations, the questions posed are real (i.e. derived from accepted NETS standards (http://www.iste.org)), the data behind the simulation is fictional, and, the data presentation is realistic (see Figure 2 which shows student achievement across grade levels).

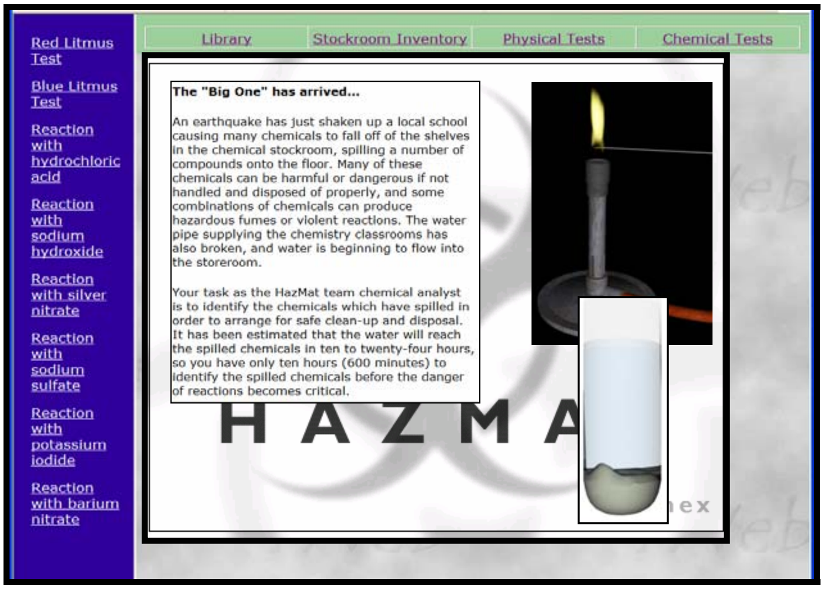

Figure 1. HAZMAT- A High School Problem Set. The opening screen of the problem is shown in the upper left corner of the window and the main and selected submenu items are shown on the top and left side of the screen. Also shown are two of the information items available. The item in the upper right corner shows the result of flame testing the unknown and the frame at the lower right of the figure is the result of a precipitation reaction.

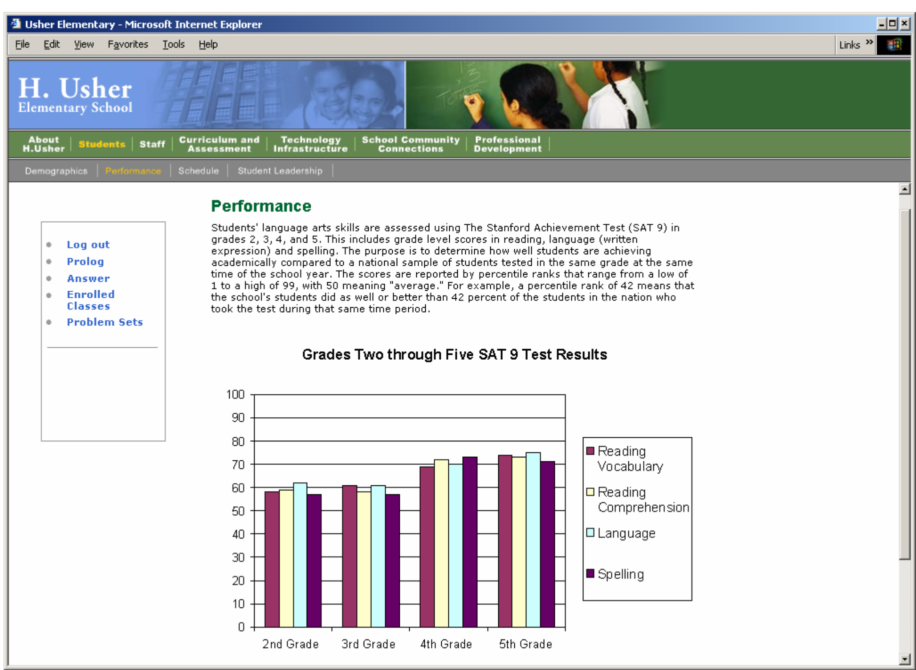

Figure 2. Sample Data Screen from an eTIPS Technology Integration Simulation

In deciding the balance of reality, it would appear important not to create a simulation that is totally fictitious, as this would decrease the value of the simulations for both students and faculty.

Value is defined in terms of the degree of external relevance a piece of information, or the simulation itself, has for the user outside of the immediate simulation, and across different perspectives. Externally valuable items are those that are valid, at least realistic, and instructionally informative without relying solely on context of the simulation, and when viewed/manipulated by a student would contribute to something s)he already knew, or teach him/her a new concept or linkage. The use of ABO blood typing evidence in a simulation is an example of an item with high external value as this topic/concept links to health, genetics, biochemistry, forensic science, etc. Another example in science would be the properties of an element in the periodic table as this would be predictive of molecular structure, industrial usefulness, reactivity, etc. In education, simulations based on comprehensive school system data or case studies, or extensive reform implementations would also have high external value. So would simulations for a local group of students that are based on local demographics and student data.

Items with lower external value have information that is primarily useful within the simulation. Low external value in and of itself is not necessarily bad to the extent it is deliberately designed for. For instance, the Roger Rabbit IMMEX problem set provides both a belief-driven solution to the case as well as a data-driven solution (Scott, 1994). Here the forensic data has high external value and the conjecture of bystanders has low external value, existing only within the simulation. This combined high/low item value approach is particularly revealing about the scientific decision making of middle and high school students (as well as teachers) when real-world alternatives to data are provided. Here some students adopt and persist with a conjecture approach, while other students choose a more data-driven approach (Stevens & Underdahl, 2003). However, problem sets with such thematic complexity are more difficulty to design, as they require that each thematic strand have its own causal coherency.

Individual items with low external value may assume greater importance by a variety of means. For instance, items with inherently low external value could be enhanced in a simulation by linkage to other items, concepts or tools with higher external value. Returning to the example in Figure 1, the simulation designers could have simply stated "The flame test indicates sodium," which would have less external value than by saying "The flame test gave a yellow color", which has less value than showing a needle in a flame with a resulting yellow color. With each increment, the external value increased as the number of external linkages the student could make increased. The final image, for instance shows "how" one would derive this identifying information about sodium, a concept that can be used in other contexts. Similarly, the data in Figure 2 which is fictitious and with lower external value, could be shifted more externally by linking it to standard analytic methods and presenting trend lines, distribution characteristics, error estimates, etc.

Individual items with low external value may also assume a higher value by aggregation across established research relationships. An example would be a fictitious teacher (Item 1, low value) with low content knowledge (Item 2, low value) who also has lower attitudes towards reform (Item 3, low value). When aggregated through the research base, these three items begin to assume a higher value for both students and teachers. (http://www3.interscience/wiley.com/cgi-bin/fulltet?ID=7400031&PLACEBO=IE.pdf). It would be important, however, to make these linkages explicit within the simulation, and especially at the closing of the simulation for those students who did not conduct a thorough enough exploration of the problem space to uncover them for themselves, or who examined the information, but missed the inherent linkages. Efficient data modeling of the problem solving process should be able to capture in real-time the students who are missing these linkages and provide formative feedback.

Finally, the internal value of an individual problem or scenario can be increased by purposefully relating the current problem to others within a problem set. Here the same considerations would apply as per the above example, but they would exist across cases. For instance, the spectrum of eTIPS cases contains primary as well as secondary school examples, as well as examples of technology-rich or technology-lean environments. Although these cases do not exist as a problem set per se, faculty with a good knowledge of the different simulations available could choose and order them as needed.

Indicators of the overall value of the simulation or problem set by which designers can gauge the success of their efforts include: 1) use of the problem set by teachers/faculty, 2) the active engagement of the students in the simulation, and the thoughtful performance of many cases, 3) student progress and mastery, and 4) user feedback. More extensive measures of the value outside the perspective of the immediate users would rely on an external evaluation team, and a detailed analysis of the cognitive complexity of the simulations.

c. Provide and Document Cognitive Complexity

The first two specifications that have been presented help developers make a "best guess" about the construction of a simulation problem space. Once student performance data becomes available it is then possible to derive a more refined perspective through research studies.

Given problem spaces of sufficient size and scope, problem solving should require an integration of domain knowledge and cognitive process skills, and research studies have shown that IMMEX provides a rich cognitive environment for problem solving. For instance, in a study of over 150 undergraduate students, concurrent verbal protocol analysis has indicated that over 90% of student verbalizations while solving a typical IMMEX case could be mapped into twelve distinct cognitive and metacognitive processes including the need for accurate cause-effect inferences, accurate evaluation of information, clarification of gaps in knowledge, monitoring of problem solving behavior, etc. (Chung et al., 2002). As expected, outside evaluators have also documented that students and teachers perceive problem sets more as a tool for reasoning and integrating information than as a system for learning new facts (Chen et al., 2001).

The cognitive complexity of a case can be influenced by 1) the degree of specification of the task, and 2) the alignment of the simulation with the ability of the users. If what is expected from a student is overly defined, or if there are overt "hints" in the framing statement that direct or prompt the student, the resulting problem space search becomes artificially focused and less revealing of the working knowledge and process skills of the student. Such directed search of the problem space will generally show a dominant and preferred approach by the majority of students when analyzed by search path mapping. Similarly, problem spaces that exceed, or underestimate the abilities of the user will often show very limited or overly extensive search of the problem space as students struggle to formulate a problem solving and/or decision making strategy.

However, as hinted at above, the cognitive complexity of the simulation can also be influenced by the overall goals of the simulation regarding the balance of assessment (Pellegrino et al, 2001) and learning (Bransford et al, 1999). While it is true that a good assessment is also a good learning experience, simulations that are directed primarily at learning a task, and assume little prior knowledge/experience on the part of the student, will have different cognitive complexities than those used more for assessment. With directed instruction the focus also begins to switch more towards the within-task monitoring of student understanding and progress, with less emphasis on the across-task monitoring (however, see Section 4. d).

The process of defining the cognitive complexities of the cases would appear similar between the science and school-based simulation environments involving student models, task models and evidence models that focus on problem identification or framing, evidence gathering and problem closure (Mislevy, et al, 2001). It could be speculated that simulations with low external value or with undiscovered linkages between related low-valued items (as may occur with novices), may reveal more superficial cognition-related events focused more on the immediate task rather than the rich understanding of the domain of study.

d. Provide Repeat Problem-Solving Experiences

The second and third principles of strategy selection and use suggest that with time, either within the first simulation, or through multiple performances of parallel tasks, students should be able to learn which strategies are "best" and modify their initial approaches as needed. This implies that opportunities should be available for students to improve and that these changes (or lack thereof) should be documented to show when, and under what conditions, these changes occur.

The design principle of repeat performances builds strongly however, on the prior three design principles. If the simulations are not solidly linked to the knowledge base of the domain, or possess little external value, or do not have a cognitive complexity aligned with the target audience, then it is unlikely that students will be motivated to perform multiple simulations and observing progress will not be possible.

Given that caveat, IMMEX addresses this need by designing problem sets that contain between 5 and 50 different instances (or clones) of a problem that share the same problem space but contain different solutions and different data. These cases can be sequenced by the teachers either randomly, or by degrees of difficulty established from prior performances and Item Response Theory. These parallel forms have the further advantages of providing multiple perspectives of a problem space, and reducing unintended interactions among students during learning and/or assessment episodes. Examples of student performance and strategic progress in problem solving can be found in Vendlinski & Stevens, 2002, Stevens & Palacio-Cayetano, 2003, and at http://www.immex.ucla.edu.

There are many opportunities to explore performance and progress in a practitioner-based simulation setting. The first is simply conveying an understanding of the scope of the professional practice. Experience-based type of knowledge is difficult to formally "teach" but is critically important for most professions (medicine as another example). Real-life reasoning and problem-solving behavior is almost never original, and solutions to new problems are often adaptations of previous problem solutions (Kolodner, 1997). Whether this is through the recall of exemplar cases (either representative or contradictory) (Berry & Broadbent, 1988), or by mental model generalizations (Johnson-Laird, 1983) (or scripts (Hudson, Fivush & Kuebli, 1992)) across a number of cases is less clear as some aspects of strategic reasoning may involve the use of compiled knowledge or implicit memory, i.e. for the most part unconscious (Reder & Schunn, 1999). Irrespective of the mechanisms of understanding, many aspects of professional practice (technology integration as in eTIPS, or leadership or substance abuse issues) could be made explicit through the presentation of a large number of exemplar simulations from different perspectives.

Strategic progress, similar to that we have described in IMMEX simulations depends also however, on the consistency of the decision-making process across cases in a problem set. It would be difficult, for instance, to follow a student's decision making progress in a school-based problem set that contained a mixture of both research/data oriented and experience/belief oriented problems as the tasks would tap different strategic processes.

e. Shape and Extend the Problem Space with Constraints

For strategic improvement, some measure(s) of success should be a component of the task (the second strategic principle), and while such measures can be diverse (a "score", a "solution", successfully stabilizing a simulation, submitting an essay to be scored, personal feedback, etc.) they should be present and students should be rapidly informed of their progress. Such measures of success fall into the category of constraints, which can have powerful influence on student's strategies, but are perhaps the least well understood of the design specifications. As an overall guide to constraining the simulation design, it would appear best to not overly constrain the simulations in the early phases of development either through embedded hints, prompts, or by restricting the scope of the question(s) posed. The most successful (research and student interest-wise) IMMEX simulations have been those where the teacher designer teams were surprised when the first student performance data was obtained.

While some constraints, such as problem solving or decision-making success act at the problem level, other constraints can act at the individual data selection level. A common constraint is to include a cost (or risk, or time penalty) for each item requested or for each incorrect "guess" at a solution. Not only does this help validate the intentionality of the act (Section 4. b.), but also serves to focus student attention on each decision during the performance. Another constraint is the time allowed to solve a case and this can be naturally encouraged through a 1-2 hr. course or lab format, or can be more structured (i.e. solve 3 cases in 60 minutes) in an assessment situation.

Constraints can also be relaxed by having students, who normally work individually on cases, collaborate in groups (Soller et al, 2002). Evidence from Case et al (2002) suggests that students who work on IMMEX problems in groups can jog some students out of persistently poor strategies.

The problem spaces and the challenges for the students can be extended beyond the simulations by requiring prior related reading, or providing different text and online references, or by concluding the cases in more open-ended ways. On most IMMEX simulations students choose an answer from a list of possible solutions which a) demonstrates an intentional act, and b) allows a quick database matching allowing immediate feedback on success that students can use to refine their strategy on subsequent cases of a problem set. A disadvantage of a list is that some students begin to overly rely on this list and some students even develop alternative strategies for maximizing the gain from the solution list (Chung et al, 2002). This is another example (although unintended) of thematic complexity and has been useful in revealing gender differences in the way that cases are closed with males being more risk-prone than females (Paek et al, 2001). Removing list selections by having students enter free-text problem solutions eliminates these alternative strategies, but apparently increases the cognitive complexity of the cases. On a medically-related problem set this resulted in a ~50% decrease in overall solution frequency for comparable groups of students.

The eTIPS school-based technology integration cases close with the students submitting an online essay to a question posed in the prologue. This essay is captured in the IMMEX environment where it is electronically submitted via Web Services to the Vermont Institutes for Science and Math Teaching (VISMT) where teachers score it with an online scoring tool (Gibson et al., 2003). Here teachers rate the essays according to a pre-defined scoring rubric (which can be locally modified/enhanced), assign a global score and enter comments for the students. This is electronically transferred back to IMMEX where students and teachers can print out reports.

Online essays have the advantage of combining student's search of the problem space along with a written description of the final decision, providing two rich and independent lines of evidence regarding student understanding. Long term, automated modeling of the students search strategy, combined with a method of automated essay scoring could provide multidimensional and near real-time assessment and feedback potential regarding student performance and progress on complex simulations.

5) Validation

Careful prospective design of the simulations will minimize the number of revisions during field-testing and accelerate the accumulation of appropriate validating data. Multiple validation criteria potentially exist for the types of simulations that have been presented, and those regarding cognitive complexity and issues of value have already been discussed in the context of the above frameworks. As further examples, IMMEX has also acquired validation data regarding the strategic principles in Section 3. For instance, students do change their strategies based on their success/failure at solving cases, they do show preferred pathways of strategic change when doing so and they do decrease their solution times with experience. Differences in these criteria can also be seen with populations of different abilities/experiences (Stevens & Cayetano, 2003, Vendlinski & Stevens, 2002). As a result, an increasing number of these simulations are beginning to assume predictive validity in the sense that it becomes possible, at the group level, to predict what students would do on subsequent simulations after as few as 2-3 simulations. This provides a rationale, and suggested points in time, for applying interventions.

The current generation of eTIPS cases is still evolving and the primary validation is of the tools (tasks, prompts, rubrics) themselves, and of broad performance differences across test sites. A prior iteration of the cases, which underwent extensive testing with both in-service and pre-service teachers however, suggested differences in the decision-making approaches between these two groups.

Summary

Simulations will continue to assume increasing importance in education during the 21st century, probably in ways that are difficult to speculate on currently. It is likely that they will significantly help to improve teacher quality, motivation and retention through their ability to improve student learning, refine and accelerate assessments, and contribute to and help organize an expanding knowledge base of practice. It is also likely that careful attention to simulation design will be an important aspect to making this a reality; without careful attention to design and value, many simulations will go unused.

In closing, it is important to point out what issues of effective simulation generation and use that have not been detailed here (Kanowith-Klein et al, 1998). Such issues include the composition of the design and development teams, student and teacher preparation for maximizing the classroom use of the simulations, implementation formats such as the use of simulations for homework vs. testing vs. in class instruction. These are all issues that are important to consider when deciding to embark on a simulation project as they relate to the concept of product design which, as described by Zaritsky et al (2002), is "essentially a prediction that its creation and adoption will provide increased leverage to a significant number of users across a range of contexts."

Acknowledgements

The core design, validation and research components of the IMMEX Project have been supported in part through funding from the Public Health Service, National Science Foundation (NSF-ROLE 0231995, DUE Award 0126050, ESE 9453918) the Howard Hughes Medical Institute, the Atlantic Philanthopies, and The Learning Chameleon, Inc. The simulation design linkages between science and education described in this paper have resulted through funding from the U.S. Department of Education's PT[3] Program to Dr. Stevens (PI Implementation Grant, P342A-990532), Co-PI, eTIPS Catalyst Grant) and Dr. Dexter (PI EduTech Implementation Grant , PI eTIPS Catalyst Grant #). Special thanks to Dr. Mary McNabb for helpful discussions.

References

Berry, D.C., and Broadbent, D.E. (1988). Interactive tasks and the implicit-explicit distinction. British Journal of Psychology, 79: 251-272.

Bos, N. (2002). Personal communication.

Bransford, J.D., Brown, A.L. & Cocking, R.R. (1999). How people learn: Brain, mind, experience, and school. Washington, DC: National Academies Press.

Case, E.L., Cooper, M., Stevens, R., and Vendlinski, T. (2002). Using the IMMEX system to teach and assess problem solving in general chemistry. ACS Annual Meeting Presentation, Orlando, Florida. http://www.acs.org/portal/PersonalScheduler/EventView.jsp?paper_key=200723&session_key=34422

Casillas, A.M., Clyman, S.G., Fan, Y.V., and Stevens, R.H. (1999). Exploring alternative models of complex patient management with artificial neural networks. Advances in Health Sciences Education 1: 1-19, 1999.

Chen, E. et al (2001). How teachers use IMMEX in the classroom. http://www.immex.ucla.edu/TopMenu/WhatsNew/EvaluationForTeachers.PDF.

Chung et al. (2002). Cognitive process validation of an online problem solving assessment. Computers and Human Behavior, 18: 669-684.

Dexter, S. (2002). eTIPS-Educational Technology Integration Principles. In P.L. Rogers (Ed.). *Designing instruction for technology-enhanced learning.* New York: Idea Group Publishing

Gredler, M. (1992). *Designing and evaluating games and simulations: A process approach.* Houston, TX: Gulf Publishing Co.

Hudson, J., Fivush, R., and Kuebli, J. (1992). Scripts and episodes: The development of event memory. Applied Cognitive Psychology, 6: 625-636.

Hurst, K. C., Casillas, A. C., & Stevens, R. (1998). Exploring the dynamics of complex problem solving with artificial neural network-based assessment systems. (Tech Rep. 446). Los Angeles, California: *University of California, Center of Research on Evaluation, Standards and Student Testing.*

Johnson-Laird, P.N. (1983). Mental Models. Cambridge: Cambridge University Press.

Kanowith-Klein, Carolyn Burch & R. Stevens. (1998) Sleuthing for science. National Staff Development Council Journal of Staff Development vol. 19, No. 3 p.48-53.

Kolodner, J. L. (1993) Case-Based Reasoning. San Mateo CA: Morgan Kaufmann Publishers.

Lane, L., Slavin, S., and Zie, A. (2001). Simulation in medical education: a review. Simulation and Gaming, 32: (3), 297-314.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. Journal of Educational Psychology, 93, 187-198

Mislevy, R.J, Steinberg, L.S., & Almond, R.G. (2001). *On the several roles of task model variables in assessment design.* In S. Irvine & P. Kyllonen (Eds.), Generating items for cognitive tests: Theory and practice. Hillsdale, NJ: Erlbaum.

Mislevy, R.J., Steinberg, L.S., Breyer, F.J., Almond, R.G., & Johnson, L. (1999). A cognitive task analysis, with implications for designing a simulation-based assessment system. Computer and Human Behavior, 15, 335-374.

Paek, P. (2002). Problem solving strategies and metacognitive skills on SAT mathematics items. Unpublished thesis, University of California, Berkeley.

Pelligrino, J., Chudowski, N., & Glaser, R. (2001). Knowing what students know: The science and design of educational assessment. Washington DC: National Academies Press.

Resnick, M. (1994). Turtles, termites and traffic jams: Explorations in massively parallel microworlds. Cambridge, MA: MIT Press.

Stevens, R.H. (1991). Search Path Mapping: A versatile approach for visualizing problem-solving behavior. *Acad. Med.* 66(9):S72-S75.

Stevens, R. H., Sprang, M., Simpson, E., Vendlinski, T., Palacio-Cayetano, J., and Paek, P. (2001). Tracing the development, transfer and retention of problem–solving skills. American Educational Research Association Symposium presentation 44.38. Seattle Washington

Stevens, R.H., and Najafi K. (1993). Artificial neural networks as adjuncts for assessing medical students' problem-solving performances on computer-based simulations. Computers and Biomedical Research 26(2), 172-187.

Stevens, R., Wang, P., Lopo, A. (1996). Artificial neural networks can distinguish novice and expert strategies during complex problem-solving. Journal of the American Medical Informatics Association vol. 3 Number 2 p 131-138.

Stevens, R.H., Ikeda, J., Casillas, A., Palacio-Cayetano, J., and S. Clyman (1999). Artificial neural network-based performance assessments. Computers in Human Behavior, 15: 295-314.

Stevens,R.H., and Palacio-Cayetano, J. (2003). Design and performance frameworks for constructing problem-solving simulations. Cell Biology Education (June, 2003, in press).

Vendlinski, T. and Stevens R. (2002) "A Markov model analysis of problem-solving progress and transfer." The Journal of Technology, Learning and Assessment. **1**(1).

VanLehn, K., (1996). Cognitive skill acquisition. Annu. Rev. Psychol 47: 513-539.

Zaritsky, R., Kelly, A.E., Flowers, W., Rogers, E., and O'Neill, P. (2003). Clinical design sciences: A view from sister design efforts. Educational Researcher, 32: 32-34.