Delta Education Journal

Table of Contents

Dean’s letter .................................................................................................................................................. 2
  Lynn Jenné House, Ph.D.

Editor’s Letter .................................................................................................................................................. 3
  Scott Alan Hutchens, Ph.D.

Thinking About All Children .................................................................................................................. 4
  Maud Kuykendall, Ed.D.

The Cart before the Horse: Thoughts on the Emergence of Educational Technology .... 9
  Reid Jones, Ph.D., Scott Hutchens, Ph.D., Heidi Eyre, Ph.D., and Darlene Crone-Todd, Ph.D.

MCT Scores Significantly Improved After Using Reading Program ......................... 22
  Lynn Walton Varner, Ph.D.

Using CAPSI to Increase Student Engagement in General Psychology .................... 33
  Heidi Eyre, Ph.D., Darlene Crone-Todd, Ph.D., Reid Jones, Ph.D., Scott Hutchens, Ph.D., and Joseph Pear, Ph.D.

Deep Structure Learning and Statistical Literacy.......................................................... 41
  David A. Swanson, Ph.D.

About the authors .................................................................................................................................. 52

Fall 2005
College of Education
Delta State University
Cleveland, MS 38733
Dear Colleagues,

We hope that you will enjoy this edition of *The Delta Education Journal*. You will note that this issue includes articles on special education, educational technology, reading, and deep learning that we feel you will find thought-provoking.

Please consider how you can make a contribution to our region and state by contributing your ideas to the next edition of the Journal. By sharing our work with one another, we have a greater chance of making a real difference in education in our state.

Please let us know if you need additional copies of this edition of the Delta Education Journal. If you have questions or comments please contact me at 662-846-4400 or e-mail lhouse@deltastate.edu.

Sincerely,

Lynn J. House, Ph.D.
Dean
College of Education
Dear Colleagues,

*Delta Education Journal* is published by Delta State University, College of Education, in the fall and spring each year. We invite manuscripts that promote teaching, learning, and educational issues. Submissions should follow APA style. For the fall issue, please submit a MS Word file email attachment to shutchen@deltastate.edu by October 1, and for the spring issue, by April 1. If a manuscript is received after a deadline, it will be considered for the next issue.

The *Delta Education Journal* has been a peer-reviewed journal since the Spring of 2005. It consists of a review board as well as the editor. Submissions will be reviewed and evaluated by the board and editor for possible publication in the Journal. Submitting authors will be notified concerning their papers’ acceptance (or rejection) and revisions may be suggested.

We would like to increase the number of submissions in the near future. Thus, we would appreciate your submissions of manuscripts related to learning and teaching. Also, please share this information about the Journal with your colleagues and encourage them to submit manuscripts to the Delta Education Journal. By doing so, your work may be shared with others, new research ideas and teaching strategies may be generated, and the quality and stature of the Journal will increase.

Sincerely,

Scott Alan Hutchens, Ph.D.
Editor
Associate Professor of Psychology
Thinking about All Children

Maud Kuykendall
Delta State University

Abstract

The field of special education is relatively young. Its first major federal legislation occurred in 1975 with the passage of P.L. 92-142 and subsequent enactment in 1977: a mere 30 years of history. Thus, the field can be viewed as still expanding and evolving. However, the standards, methods and practices of special education continue to reflect a primarily behaviorist perspective as embodied by the ubiquitous IEP with its measurable goals and behavioral objectives (Council for Exceptional Children, 2005). In my essay I present several principles of constructivism in order to familiarize future special educators with these principles so that they may gain a deeper and more meaningful way to think about all children including those with exceptional learning needs within the general education classroom.

As a profession, special education relies on the assumption that it is both legitimate and possible to expose, measure, and categorize “normal” and “abnormal” cognitive, emotional, and behavioral phenomena within individuals. Thus, a close relationship exists between psychometrics and special education….However controversial, this reliance on tests (rather than portfolios, for example) to explicate learning or failure to learn engenders a particular way of thinking about children (Reid & Valle, 2004, p. 469).

Through the use of such descriptive language, Reid and Valle (2004) highlight the scientific and behaviorist approach that the field of special education takes towards thinking about children. They invite the field, specifically the field of learning disabilities, to rethink and broaden its scope of research and practice to a more welcoming and accommodating stance towards all children.

The field of special education is relatively young. Its first major federal legislation occurred in 1975 with the passage of P.L. 92-142 and subsequent enactment in 1977: a mere 30 years of history. Thus, the field can be viewed as still expanding and evolving. However, the standards, methods and practices of special education continue to reflect a primarily behaviorist perspective as embodied by the ubiquitous IEP with its measurable goals and behavioral objectives (Council for Exceptional Children, 2005). In
In my essay I present several principles of constructivism in order to familiarize future special educators with these principles so that they may gain a deeper and more meaningful way to think about all children including those with exceptional learning needs within the general education classroom.

Positivist thinking of empiricist theories dominated the field of education from the turn of the 20th century when behaviorism, a major component of empiricism, ascended as a powerful influence in education (Kuykendall, 1997). Concurrently during the 20th century, constructivist theories of learning emerged. Heated debates arose between proponents of behaviorism and proponents of constructivism. The conceptions of constructivist theorists challenged the notion that knowledge is acquired and can be transmitted from teacher to child with the belief that knowledge is constructed through meaning-making within social contexts (Bruner, 1990; Chomsky, 1957; Dewey, 1938, 1963; Gallagher & Reid, 1983; Vygotsky, 1971, 1986). Today educational perspectives and practices continue to reflect, in varying degrees, both influences of behaviorism and constructivism and on occasion, the controversial aspects.

Perspectives and practices in the field of special education that can be readily observed from the classroom level to the national governing level continue to reflect primarily behaviorist theories. At the national level, the Council for Exceptional Children (CEC) is the largest organization representing individuals with disabilities and consequently providing an ethical and practical foundation for special education and its educators. It is dedicated to improving educational outcomes for children with exceptional learning needs and for providing professionals with high standards of teaching. The CEC standards for the preparation of special educators contain language that is so firmly behaviorist in tone, it is hard to envision how these standards might provide some insight to the application of constructivist principles. For example, consider the language in the last sentence of Standard 4 that reads “Moreover, special educators emphasize the development, maintenance, and generalization of knowledge and skills across environments, settings, and the lifespan” (Council for Exceptional Children, 2005).

I offer two key ideas embedded in CEC Standard 3, Individual Learning Differences and in the first section of Standard 4, Instructional Strategies to scaffold our understanding towards a constructivist perspective. In Standard 3, the idea is that special
educators are challenged to reach a standard that would allow them to “understand that the beliefs, traditions, and values across and within cultures can affect relationships among and between students, their families, and the school community” (Council for Exceptional Children, 2005). In Standard 4, Instructional Strategies, the idea is that they are challenged to “enhance student learning in terms of critical thinking, problem solving, and performance skills… and increase their self-awareness, self-management, self-control, self-reliance, and self-esteem” (Council for Exceptional Children, 2005). Both of these ideas embody the constructivist concept that learning involves an understanding that human behavior includes meaning-making and the intentions and emotions of people in relationship with themselves and others (Bruner, 1990).

Constructivist theories profoundly impact our understanding of the nature of learning as well as our understanding of the nature of language. During the 20th century and in many parts of the world, scholars proposed through a variety of works the theory that knowledge is constructed. The works of John Dewey, Jean Piaget and Lev Vygotsky influenced our educational beliefs by elucidating their theories. Dewey influenced American education during the first half of the 20th century by advocating that students be directly involved in their own learning, building on their current knowledge and experiences (Dewey, 1938, 1963). In Switzerland, Jean Piaget, an epistemologist, indicated that knowledge is actively constructed rather than passively acquired (Piaget, 1948, 1972). In Russia, then the Soviet Union, Vygotsky discussed the higher psychological functions and the role of both social and cultural factors including thought and language. Additionally, he recognized the need to study people as members of a particular society or community (Vygotsky, 1971, 1986, 1978, 1922, 1993).

Principles related to meaning-making form the basis for understanding how children and adults with or without exceptional needs learn, think critically, and problem-solve. These principles are tied mainly to theories of Piaget who postulated that language is neither innate nor learned, but emerges as a result of a children's constructive activity. Children’s language grows most readily when they actively pursue their own learning. Meaning (i.e., knowledge) is actively constructed by the learner. In addition, errors are critical to learning (Piaget, 1948, 1972).

Dewey (1938, 1963) contributed foundational concepts that propose open, active, and participatory learning in which the learner would develop self-control and
make choices that encouraged thinking. Students with disabilities along with all children need opportunities to develop self-control. He recognized that children are always learning and that control of learning rests with the learner. He also proposed students participate in authentic language and literacy experiences throughout the curriculum (Dewey, 1938, 1963).

Vygotsky and his colleagues posited the notion that humans learn through those around them who, in turn, are immersed in their own culture and history. Learning is both transactional and a social process that is fostered by social interaction. Children with exceptional needs along with all children engage in the complex processes of reading, writing, discussing, and thinking while they simultaneously develop language and literacy, and learn about and through these processes. He believed that children learn to read by reading and to write by writing (Vygotsky, 1978). He also emphasized that language is learned most easily when language is kept whole and authentic in the context of real speech and literacy events (Vygotsky, 1971, 1986). He as well as Dewey recognized that children learn best what is meaningful and relevant to them (Dewey, 1938, 1963; Vygotsky, 1978).

The CEC professional discourse of performance-based standards for the preparation and licensure of its professionals may be written in the language of behaviorism but offers an opportunity for future special educators to deepen their understanding about what constitutes effective teaching and how to best serve children. We may consider the two key ideas embedded in CEC Standards 3 and 4 that contain constructivist principles. These principles of learning based on the theories of Piaget, Dewey and Vygotsky present a way to interpret the CEC statement that “the beliefs, traditions, and values across and within cultures can affect relationships among and between students, their families, and the school community” (Council for Exceptional Children, 2005). Future special educators can actively construct their own meanings to understand how “primary language, culture, and familial backgrounds interact” and to “enhance student learning in terms of critical thinking, problem solving… to increase their self-awareness, self-management, self-control, self-reliance, and self-esteem” (Council for Exceptional Children, 2005). They may come to appreciate the theoretical underpinnings of learning as an understanding that human behavior includes meaning-
making and the intentions and emotions of people in relationship with themselves and others (Bruner, 1990).

References


The Cart before the Horse: Thoughts on the Emergence of Educational Technology

Reid Jones, Scott Hutchens, Heidi Eyre, and Darlene Crone-Todd
Delta State University

Abstract
The Twentieth Century was fertile ground for educational technology. In most cases, the educational concepts and pedagogy were developed before technology could take advantage of these ideas. However, the “computer revolution” of the last three decades has reversed this situation. The explosion of technologies brings capabilities for an enormous expansion of what educators can do with these tools. This article briefly describes some of the key events that shaped our present educational technology. Then the educational technology experiences of four Delta State University Psychology faculty are discussed with some thoughts about “lessons learned and learning”. Finally, the authors speculate on some of the ways new technologies may be retro-fitted to recent psychological research.

Before computers …

Since its inception, psychology has made the study of learning and memory a central focus. A great deal of modern instructional technology has grown from research and theory about these concepts. This paper outlines a brief history for some of the important ideas and discusses a few trends in that history. A Hypertext history of instructional design (S. McNeil, undated; Retrieved September 20, 2005) was used as a general source for many early contributions.

Psychology was still struggling to take on its modern form when Edison invented the first proclaimed tool of instructional technology – the motion picture. Wundt’s “Structuralism” was being abandoned for Dewey’s “Functionalism.” A catalog of instructional films was available by 1910, and films were used in Rochester, NY schools in the same year. Edison proudly proclaimed that soon books would be obsolete, since all human knowledge could be taught more effectively with the motion picture. By this time, Watson’s “Behaviorism” was rapidly gaining popularity. Thorndike (1910) summarized his early Laws of Learning (Trial and Error, Effect, Exercise, and Recency) while pointing out their obvious utility for educational design. He even anticipated programmed instructional texts by pointing out the benefits that would be gained by requiring mastery of a first task before students could proceed to the second task.

Three very important themes were developing during these early years which made it possible for technology to be applied to teaching. First, it was clear that
educational technology would be most effective if it could focus on the individual learner rather than on group learners. Showing a motion picture to a class can be effective, but it hardly approaches the flexibility of machines (computers) interacting, one-on-one with each student. Several other advantages were immediately obvious. Technology could adapt to the widely varying individual rates of learning, so that learners could make steady progress at a comfortable rate. Further, technology could evaluate performance and provide immediate feedback for individuals, something that is not possible for a teacher working with groups of students. Additionally, technology could be programmed to wait until each student demonstrated mastery of lower level objectives before that student even be allowed to attempt higher level objectives. These three themes (individualization, immediate feedback, and mastery) became much of the scaffolding that allowed our present day educational technology to develop.

By 1926 Sidney Pressey had developed a “teaching machine” which was really a modified typewriter carriage for multiple choice questions. Pressey incorporated many of Thorndike’s principles in his “Teaching Machine” to promote learning through drill and mastery. One of the key features of Pressey’s machine was its ability to provide learners with immediate feedback on the adequacy of their learning responses. Soon Pressey was calling for an “industrial revolution” in education (Pressey, 1932). In 1938, H. G. Wells (McNeil, 2005) proposed a “world brain” to replace encyclopedias, statistical databases, and journals. The world brain should be available to all students, all researchers, and all scholars. He argued that it would have to be established as some sort of network. His proposal obviously anticipates the World Wide Web. Unfortunately, the resources for Pressey’s revolution and the “World Brain” would have to wait for the “Great Depression” and World War Two to run their course.

At least, however, there were continuing developments in research and theory on learning, memory, and pedagogy. In the 1930’s, Ralph Tyler pioneered the use of general objectives for curricular design and provided evidence that objectives were most effective when stated in terms of student behaviors (Tyler, 1949). Specialized tracks (such as “college-bound” learner objectives) were also possible within each curriculum.

B. F. Skinner’s Operant Conditioning was becoming an important force in education as well as in psychology. Concepts such as immediate feedback and small learning steps were rapidly being assimilated into plans for educational technology. The
Skinnerian concept of reinforcing successive approximations ("shaping") was initially applied to such “target behaviors” as lever pressing in a rat. However, it became rapidly apparent that shaping could be applied to target behaviors such as Tyler’s behavioral objectives.

Skinner’s proposals were based generally on laboratory research, although he often suggested applications which were based more on his creativity than on hard evidence. He proposed that educators incorporate both the science of learning (laboratory research) with the art of teaching (Skinner, 1954). Subsequent publications (Skinner, 1958) showed how he proposed to develop a methodology for “Programmed Instruction,” using pre-arranged mastery-based sequences of behavioral steps to produce something very much like what Thorndike and Pressey had proposed 30 years earlier.

Concurrently, Benjamin Bloom (1956) published his taxonomy of educational objectives for the cognitive domain. This development provided an organized framework for instructional objectives and has stimulated thousands of educational research articles and applications. Bloom’s taxonomy of cognitive instructional objectives has been better operationalized (Pear, Crone-Todd, Wirth, & Simister, 1999), remaining the standard to this day.

Use of the Word Associate Test in the 1960’s by William S. Verplanck (2002) was an attempt to further develop efficient assessment tools to evaluate written test performance for college students. Students were asked to provide associations to half of the concepts on a written test and to write extended answers on the other concepts. With a counterbalanced design, high correlations were consistently demonstrated between the associations and the written answers. Verplanck was planning for the day when computers could scan essay tests, recognize key words that he had catalogued through years of associate testing, and provide immediate and objective feedback on essay tests.

Instructional technology was still greatly limited in the 1960’s by the relative dearth of information technology. Even though great advances in computers were on the horizon, some of these ideas fell by the wayside. Skinner’s Programmed Instruction and Verplanck’s Word Associate Test were pretty much abandoned as being tedious and labor intensive for students and faculty alike. Theory was so far ahead of technology, that many good ideas never got a fair chance.
It remained for Fred Keller (1968) to propose an overall plan that incorporated most of these ideas. His Personalized System of Instruction (PSI) was based on learning objectives which were sequenced in small steps, ensuring mastery. Failure to master an objective did not result in a poor grade. Instead, the student would take an alternative mastery test, and so on, until mastery was complete. Obviously, the system was individualized. Some students may need two to three times as many chances as others did for mastery to be achieved.

While educators were awaiting assistance from “Information Age” technology, the theory and practice of education added an important movement – the re-emergence of cognitive psychology and constructivism in education. Glasser (1962) developed a framework and terminology for instructional design (task analysis) for Individually Prescribed Instruction that added a higher level of cognitive objectives to the Bloom taxonomy. Robert Mager (1975) published his short, but highly influential work on instructional objectives which were specifically linked to criterion-referenced assessment. Criterion Referenced Instruction/Assessment offered far more precision than the Norm Referenced Assessment of that period.

The computer age in education …

When the Soviet Union launched Sputnik in 1957, reactions in the United States were first one of panic and then one of determination. We knew that the USSR already had the hydrogen bomb, and now we realized that they were far ahead of us in the “space race.” This event had a profound effect on our educational establishment. First, academic standards were dramatically raised in an attempt to produce a generation of highly skilled scientists. On the positive side, the federal government poured money into educational research and development. Further, a side effect of our attempts to catch up in the space race was the development of the microchip for use in satellites which has revolutionized information technology and produced the “Information Age”.

In 1963, IBM developed important alliances with educators at Stanford University to bring technology to the classroom. However, at that date, IBM promoted the use of main-frame computers and time sharing systems which were far too expensive for the individual or most school districts to afford. When Steve Jobs and Steve Wozniak developed the Apple I personal computer (1976), educators were provided with a vastly more affordable alternative. The Minnesota Educational Computing Consortium was
developed by Apple where teachers did much of the programming. Programs were polished and cataloged and then made available for K-12 teachers. IBM more or less backed out of educational computing for several years, but they had a plan. When the IBM PC finally was released, it cornered most of the educational and home markets.

Joseph Pear and his colleagues at the University of Manitoba have been using the computer-aided personalized system of instruction (CAPSI, see Kinsner & Pear, 1988; Kinsner & Pear, 1990) for over two decades. Initially, Pear used the PSI system in classrooms, which required a great deal of administrative effort. CAPSI was developed as a way to reduce the administrative burden on the instructor. CAPSI originally used email as the primary medium for assigning unit tests, reviewing, and communication between students and between students and the instructor. CAPSI also featured in-class proctoring (peer reviewers) by providing bonus points for students who had already mastered earlier assignments in the course. This increase in student-student interaction is an important way to increase student engagement.

One of the drawbacks of the original CAPSI system was that it took some time to train students how to use the system. During that era, many students had little or no experience with computers, and would express worries that a lack of typing skills may prevent them from succeeding. Over the years, CAPSI moved to a DOS-based program allowing for a menu-driven system. The new CAPSI was much easier to use and required less training time before getting to the subject matter of the course.

As DOS-based systems became obsolete, CAPSI was programmed for a web-based platform, where it now resides (www.webcapsi.com). CAPSI has been utilized by instructors and researchers at the University of Manitoba, Delta State University, and the University of Redlands. The new platform allows for instructors and students to have access to CAPSI anywhere Internet service is available. This revolution in the use of PSI also provided a better opportunity to study educational processes involving both students and instructors (see Pear & Crone-Todd, 1999).

Lessons learned and learning …

The psychology faculty at Delta State University have all been involved rather heavily in educational technology. In the following section, they will briefly describe some of their experiences and uses of instructional technology.
Reid Jones: In the late 1970’s, I was Department Head of Psychology at Knoxville College – a small private college for African-American students supported by the Presbyterian Church. I was able to obtain two grants (at $250,000 each) from the National Science Foundation to develop the social science curriculum and to dramatically increase the number of students who went on to earn doctoral degrees. One of the major activities of the project was to set up a computer-managed general psychology sequence that lasted for a full academic year. This was based on the Keller system of Personalized Instruction with a strong assist from Robinson’s (1970) SQ3R study techniques (Jones & Shrader, 1978, 1980, 1983). Students could prepare for clear instructional objectives and then be tested at any time in two computer labs on the campus. Multiple choice items were randomly selected from 1200 items criterion referenced to chapter and instructional objectives. Data was also collected on which objectives were giving students the most trouble. What we had was a Computer Assisted Personalized System of Instruction (CAPSI) based on multiple choice items.

I was able to achieve the objectives of those grants, and the multiple choice CAPSI yielded an additional and surprising dividend. Psychology majors had to complete the three term CAPSI course before they could register for popular upper division psychology classes. By the time they had finished a year of self managing their learning behavior, as a group, our psychology majors were among the most organized and conscientious students I have ever had. While I greatly prefer the way Darlene Crone-Todd and Heidi Eyre are now using CAPSI at Delta State, I was very pleased to be far ahead of the curve in the 1970’s. CAPSI at Delta State is based on far more use of higher level cognitive (behaviorally defined, of course) objectives and emphasizes written answers.

A few years later, IBM came out with their long-range plan that incorporated personal computers, relegating Apple to a small share of the educational technology market. There are a few applications (music and art) where Apple still may have the edge, but most educational computing is Windows-based (essentially IBM). We are also now able to have both the advantages of a personal computer and one that is connected to the World Wide Web (H. G. Well’s “World Brain”).

PSI is now widely used because of educational technologies such as CAPSI. Further, there is substantial evidence that PSI and CAPSI are more effective than
traditional college level instruction. Many studies that support this notion, but perhaps the most influential was the meta-analysis by the Kuliks at Michigan State (Kulik, Kulik, and Cohen, 1979). PSI classes produced higher performance and were preferred by students to traditional lecture classes. The individualization of instruction was a necessary adaptation that had to precede computer assisted personalized systems (CAPSI). It has also made possible a wealth of “distance learning” options and on-line courses.

**Darlene Crone-Todd:** As an undergraduate student, I had several experiences with the CAPSI system at the University of Manitoba that helped me realize the importance of learning at the mastery level. It enabled me to learn the material at a deeper level and at my own pace, while allowing increased interaction with students and the instructor. Later, as a teaching assistant for Joseph Pear’s undergraduate courses, I realized that the CAPSI system was a way to allow students flexibility in scheduling but still have deadlines for completion of supervised examinations. During my doctoral studies, and currently, I have collaborated with a number of students, peers, and colleagues to study various research questions using CAPSI. There are numerous lessons that I have learned as a result of this research endeavor.

Lesson 1: It is important how you give feedback. I learned early on that students like it when someone said, “You did a good job!”, or “Nice example!”. However, when students are receiving constructive criticism, it is important to depersonalize such feedback. For example, instead of “You did not include an example”, the feedback could be written as, “The answer did not include an example.” As an instructor and a researcher for these courses, the complaints increase if a mentor or reviewer is personalizing constructive feedback.

Lesson 2: It is important to be able to specify what is required, and how you will grade the answers. My collaborative research on modifying Bloom’s taxonomy established a new scoring method that is more reliable (see Crone-Todd & Pear, 2001; Pear, Crone-Todd, Wirth, & Simister, 2001). I found that as I continue to work with the taxonomy, that students like to learn how to learn, and that the flowcharts developed for assessment help them learn how to think.

Lesson 3: Sometimes one needs to add back in components that were lost. In my use of CAPSI, the mastery outcome was always a “pass” or “restudy”. I have since learned that when students are developing as writers, it is important to give “close”
answers a “conditional pass” and to allow students to revise their answers. This changes the course climate from one of frustration on the part of the students (and sometimes the instructor) to one of hopeful improvement. My experiences thus far indicate that the use of the conditional pass is a useful feature to incorporate into CAPSI, particularly at the undergraduate levels of higher education.

There are of course many more lessons from researching and using CAPSI, but space does not permit an exhaustive review. Suffice it to say that using CAPSI is a valuable way in which to increase student writing and engagement, and that the beneficial learning outcomes certainly make the work worthwhile.

Heidi Eyre: Not long after I joined the psychology faculty at Delta State, I started thinking about how to increase student engagement in my general psychology courses. I knew Darlene Crone-Todd had been researching student learning issues. I approached her about collaborating on a research project designed to vary the type of assignments given in general psychology and measure subsequent student engagement and performance. Crone-Todd asked what my goals were in terms of increasing student engagement and performance. I responded that I wanted students to be able to demonstrate more understanding of psychology through the use of frequent knowledge-based assessment measures and engage in a greater number of course-related activities (as opposed to only using one or two short essays and exams).

Fortunately, the Internet version of CAPSI was almost completed and would be very helpful in achieving these goals. We modified the University of Manitoba general manuals for use at Delta State (which contained the course syllabus and instructions on how to use CAPSI) and also created a course manual (with unit study questions) for my general psychology course. After a successful trial run of CAPSI in Spring 2005, we modified the course slightly to allow students to take course exams at their own pace in Fall 2005. Also, I developed a course manual for developmental psychology and began offering that course as a CAPSI-based course in Fall 2005 as well.

Thus far I have been impressed with the CAPSI system. It requires a fair amount of preparation and set-up time at the front end of a new course, but that tends to be true for all technology-intensive courses. However, once the course is set up, it requires about the same amount of time and effort to manage as any other course. The advantage for students is that they can work at their own pace and they are in control of what they learn
and earn. They can write unit tests for points and can gain bonus points through peer reviewing if they like. Many students enjoy this aspect of the course because they feel like they are in control of their course grade, rather than endorsing the common misconception that their course grade is dependent on the instructor. In short, it appears that many students really enjoy using the CAPSI system because it challenges them to think critically, improve their writing skills, rewards them for not procrastinating, and gives them immediate and visible payoff (in terms of grades) for their effort.

More recently, these CAPSI technologies have been widely used for distance learning. However, the big development of the 1990’s was the appearance of the “World Brain” – the Internet. Just as H. G. Wells had proposed, virtually every student now has access to an omnibus encyclopedia, where much of the information is being sent in “real-time.” Streaming videos, chat rooms, and online discussion boards are widely used at almost all universities. Scott Hutchens has been using these tools within web-enhanced, hybrid (i.e., course that meets half of the time and the rest is online) and online courses.

**Scott Hutchens:** I have been using instructional technology (i.e., WebCT, Blackboard, PowerPoint, etc.) in the classroom and converting traditionally-taught courses into technology-based courses since 1999. As a result, I have learned through experience that hybrid and online courses cannot be taught and assessed in the same ways as traditionally-taught courses (see Hutchens, Jones, Crone-Todd, & Eyre, in press). My role has shifted significantly from one primarily delivering course material in a traditional course to one of guiding or mentoring students in their learning of the material (Bourne, McMaster, Rieger, & Campbell, 1997). I have designed hybrid and online courses so that students must be active participants, become more responsible for their own learning, and have or develop good communication skills (Miller, King, & Hutchens, unpublished manuscript). In such courses, students cannot passively listen to course lectures and rely on the instructor to make sure they learn the material as they may do in some traditionally-taught courses.

Thus, I have designed my hybrid and online courses to create an online collaborative learning community in order to establish communities of learners, facilitate social interaction, active participation, cooperation, exploration, and collaboration (Hiltz, 1997; Miller, King, & Hutchens, unpublished manuscript; Spitzer, 1998). My hybrid and online courses are also student-centered, self-regulated, self-directed, and constructivist
(Hutchens, Jones, Eyre, & Crone-Todd, in press; Miller & Hutchens, under review). I have also learned that instructors should strive to enhance their personal touch in hybrid and online courses in order to increase a student’s sense of belonging and establish a collaborative learning community (see Hutchens, Jones, Crone-Todd, & Eyre, in press).

Furthermore, I have found that I can appropriately engage students and assess student learning in hybrid and online courses by effectively utilizing the tools (i.e., online discussion boards, email, assignments, quizzes, tests, group presentations, progress tracking, and grade books) in online management programs such as WebCT (see Hutchens, Jones, Crone-Todd, & Eyre, in press). Over the years, using hybrid and online courses to create an online collaborative learning community has been successful for me. Quantitative data have demonstrated that students perform better and learn more in web-enhanced than traditionally-taught general psychology courses (see Hutchens, 2004b). In fact, recent quantitative data have indicated that student learning and performance in general psychology courses did not significantly differ between web-enhanced and online courses (Hutchens, under review).

Moreover, qualitative and quantitative data have indicated that students are more satisfied in technology-based courses than traditional courses (Hutchens, 2004a). Specifically, students reported that they enjoyed engaging in online discussions, assignments, simulations, demonstrations, PowerPoint shows, and the ability to check their grades using the course management program WebCT (Hutchens, 2004a). Students also reported that they appreciated the fact that course material was available anytime, anywhere (Hutchens, 2004a). It has also been demonstrated that students learn more technology skills in technology-based courses than traditionally-taught courses (Hutchens, 2005a). Students are also more likely to learn technology skills when they are required to use technology as a learning tool, and they observe a model or mentor demonstrate the appropriate use of instructional technology (Hutchens 2005a, 2005b).

Some possibilities . . .

Some conclusions seem justified from this brief history of pedagogy and educational technology. This seems to be a “horse before the cart” situation. First, early theories of educational technology were greatly limited by our lack of the technology to actually accomplish these proposals. Second, it was important for the theory and research to proceed, even though the technology was not available to accomplish an
application of these concepts. Third, educational technology has recently advanced beyond our pedagogy. Perhaps you could say now that “the horse is running away from the cart.” If you are really serious about seeing some of the enrichment activities that are available, I recommend that you “log on” to Delta State University psychology courses available on the Internet.

Fortunately, there continues to be creative development in research and theory that can challenge us to use the remarkable tools that technology has given us. We will just mention a few ideas which have yet to be adapted to digital technology. I believe that A. S. Kauffman’s concept of “simultaneous processing” (2000) could be adapted to the multi-tasking capabilities of computers. Most of our current educational technology depends on the more traditional left-brain sequencing of concepts. Another underdeveloped area is computer adaptive testing, based on “branching” routines in the selection of testing and allows assessment and learning to be automatically fitted to the learner’s capabilities. Perhaps we can find ways to more thoroughly explore Gardner’s eight types of intelligence (Gardner & Hatch, 1989) and how technology can be harnessed for educational applications which sample all of these “intelligences.”

Author Note
Portions of this paper were presented at the Reaching Out to Mississippi: Education in Action faculty development conference held in Cleveland, MS, September 28, 2005.

References


MCT Scores Significantly Improved After Using Reading Program

Lynn Walton Varner
Delta State University

Abstract
After using a computer-managed supplemental program to focus and aid instruction, students’ scores significantly improved on the Mississippi Curriculum Test (MCT). For the third grade (98 students), researchers compared state test score gains after students in the treatment schools had been exposed to the program for seven and one-half months. Groups were matched for socioeconomic status, ethnicity, gender, and achievement levels at the beginning of the study. Using a One-Way Analysis of Covariance there was a significant difference F (2, 96) = 44.15, p < .0001; η² = .479 in reading score gains. However, the students’ scores in mathematics and language arts did not improve; the classroom teachers reported not using the program in those two subject areas, although they had been trained to do so.

Introduction
Because of the critical consequence that the ability to read effectively has upon school success and quality of life in general, great emphasis has been placed upon various instructional and assessment methods for elementary reading. Increased emphasis on accountability places more importance on implementation of programs that work. Schools are searching for scientifically-based programs that have been shown to be effective. Various methods of instruction are constantly appearing on the horizon with the latest fix for students being espoused. This is especially true in the area of technology, which in itself is not enough to improve instruction. However, technological gains have provided tools which can be of great assistance to good teaching. One of the most helpful aspects of computers is the ability to individualize instruction to the learner’s specific needs.

Four schools located in the Mississippi Delta participated in a research study conducted from November 2002 to May 2003. Using cluster sampling, seven third grade classrooms in four school districts which were matched for socioeconomic status, ethnicity, gender, and achievement levels participated. Access to teacher knowledge of technology as well as teacher experience and qualifications were reviewed and found to be equivalent for the comparison and treatment groups. End of year second grade Mississippi Curriculum Test (MCT) scores were compared to ensure equity in achievement levels. The MCT is a criterion-referenced statewide test, which was developed by multiple experts and has reliability coefficients ranging from .88 to .99 for
the reading portion (Mississippi Department of Education [MDE], 2003). Additional information regarding the MCT may be found on the MDE website (http://www.mde.k12.ms.us/). The sample consisted of 48 students in the treatment group and 50 students in the comparison group. MCT gain scores (third grade compared to second grade for the same students) were analyzed with the initial differences (second grade scores) accounted for by the use of covariance.

**Review of Literature**

High stakes testing is now a ubiquitous part of the accountability systems in many states; school districts, administrators, teachers, and students are judged based upon scores on annual academic tests. Price and Burton (2004) explain that many schools have been designated as underperforming according to the federal No Child Left Behind Act and the resulting state-adopted accountability models. In these schools, test results reveal an inability to show acceptable levels of achievement for all groups of students. However, in many schools students are achieving at acceptable and even greater than expected levels in terms of academic performance on state tests. In these effective schools, Chapman and Harris (2004) found a central and pervasive focus on teaching and learning and on program evaluation among administrators and classroom teachers. School leaders must collect and analyze school data to make informed decisions and develop improvement strategies. “Data richness has long been found to be an important component of effective and improving schools” (Chapman & Harris, 2004, p. 225).

Traditionally, most statewide assessment data only show group progress (Olson, 2005). In an effort to provide data which promote the understanding of individual student progress, many states have changed from norm-referenced to criterion-referenced achievement tests. However, teachers need more specific student assessment information in order to individualize classroom instruction to meet students’ particular needs than is usually provided by state assessment data (Popham, 2005).

New approaches to instruction often involve the use of technology, specifically computer-based programs. The ability of these programs to provide formative assessment that allows the computer to map out individual plans of instruction for each student is seen as a great strength (Mioduser, Tur-Kaspa, & Leitner, 2000). Technology-based programs are generally classified as computer assisted instruction (CAI). Data provided by CAIs can give support to show student mastery of skills, which adds objectivity to
analysis of student progress (Popham, 2005). Computer assisted assessment allows efficient and detailed information about an individual student’s level of competence in each specific skill within every subject area (Olson, 2005). Students can then be grouped and re-taught the specific skills they have not mastered until mastery is achieved. Each child gets only what he/she needs.

Few efficacy studies are available for many commercial technology-based programs. State departments of education review programs with the purpose of providing schools a list of state-approved programs to assist with decision making (www.mde.k12.ms.us). Often, federal and state funding cannot be used to purchase programs that are not on the state’s lists of scientifically based programs. New versions of software are continually being released, new technologies are continually being developed, and educators are attempting to be on the cutting-edge of the latest methodologies. Published success stories on vendor websites often reveal results of a particular school or school district that has gathered and analyzed its own data. A review published by North Central Regional Educational Laboratory (Eisenberg & Johnson, 1996) stated that integrated learning systems help students learn skills in isolation, but fail to help students understand how to integrate or transfer those skills to other experiences.

In a review of literature (Parr & Fung, 2004), integrated learning systems were examined and rated by degree of cost effectiveness and academic results. Most of the programs included a computer-managed system that offered assessments, prescriptions for instruction based on assessment results and continual reports of progress. Some of these effective programs were very costly to school districts. One factor was evident; the correlation between the school’s curriculum and the CAI’s content, both assessment and instruction, were vital to the effectiveness of the technology-based program.

The particular computer assisted instructional program used in this study was the LeapTrack Assessment and Instructional System (LeapFrog SchoolHouse). A validity study conducted in Henry County, Georgia, revealed that the LeapTrack end-of-year assessments (3rd grade) were correlated with the Stanford-9 (SAT-9) in six classrooms (LeapFrog SchoolHouse. The Research).
Methods

Four schools in the rural Mississippi Delta region were chosen to participate; two schools were classified as low-achieving and two were classified as high-achieving according to the state’s standards (based upon 2001 performance on the Mississippi Curriculum Test). One high and one low-achieving were used as the comparison group; in these schools, the students were given the assessment portions of the program only. In the other two schools, the assessment and the instructional materials as prescribed by the program’s learning paths were utilized. This allowed teachers to prescribe specific skills for students to practice according to the diagnostic assessment provided by the program.

Third grade students in two treatment schools used the *LeapTrack System* (*LeapFrog SchoolHouse* website) for approximately 30 minutes a day, two to three days per week, during independent study time. In these classrooms, students rotated through a guided reading center, a computer station, independent board or sheet work, and the *LeapTrack System*. Procedures were the same in the comparison classrooms except for the absence of the *LeapTrack* center. The program uses the *LeapPad*, a near touch technology; a writeable/readable cartridge; and printed materials to administer at least three formative assessments per school year and then prescribes learning activities following each assessment. The media used is a book that is placed into the base accompanied by a data cartridge that gives audio instruction and records feedback as writable data. The data are then transferred to a computer’s hard drive or network drive through a universal serial bus connection (USB).

All classrooms in this study followed the state’s curriculum framework, provided interventions recommended by the state, and used the same state adopted textbooks. Teachers in the participating schools were designated as *highly qualified* according to the *No Child Left Behind Act*; those in the treatment and comparison schools were matched for levels of degree, certification, and professional development experiences. The classroom environment and methodology were similar in the comparison schools without the exposure to the *LeapTrack* instructional system. The comparison students participated in the *LeapTrack* assessments, which were administered by an assessment team contracted by Delta State University. Regular visits were made to the two treatment schools to observe classroom implementation in order to ensure treatment fidelity. Teachers were interviewed in order to determine the support level.
needed, and suggestions were discussed with teachers in the areas of classroom management. It was determined by observers and interviewers that adequate time had not been dedicated to the math and language instructional skill activities. According to observational and interview data, the teachers in the treatment schools used the program in reading an average of 30 minutes two to three days per week. However, the use in mathematics and language arts was far less than that amount.

Treatment and comparison classrooms were matched demographically by free and reduced lunch, minority enrollment, achievement levels, access to technology, and teacher qualifications. All treatment and comparison students were measured using the Mississippi Curriculum Test (MCT) in May 2002 (pre-test) before treatment students used the LeapTrack system, and again in May 2003 (post-test). The MCT was developed by the state of Mississippi through committee representation in conjunction with CTB/McGraw-Hill in 1998 and 1999; it was then pilot tested in 2000 and 2001. The grade level tests statistics for the three operational forms originally developed are available on the MDE website (http://www.mde.k12.ms.us/acad/osa/Technical_Manual_for_MCT.pdf). Comparison of the group’s spring 2002 scores showed there was no significant difference between the sample students, $t(80.98) = .183$, $p = .855$. Students who did not have MCT pre and post test results in all three subject areas were not included in the results. Data were analyzed by examining scaled scores. Therefore, it was concluded that the treatment and comparison groups’ initial scores were not significantly different at the beginning of the study.

Results

In order to test the significance of the results of the MCT for the two groups, an independent samples $t$ test was performed. Results of an independent $t$ test on MCT language arts scores for the two groups did not indicate a significant difference between the treatment and comparison groups’ scores for 2003 [$t(95.016) = -1.350$, $p = .180$]. The same was true for the MCT mathematics scores for 2003 [$t(94.853) = .076$, $p = .940$].

The results did indicate a significant difference between the treatment and comparison groups’ scores for 2003 MCT in reading [$t(94.027) = 3.166$, $p < .002$]. When another $t$ test was conducted on the reading growth score for each group between 2002 and 2003, a significant result also occurred [$t(93.405) = 4.313$, $p < .000$] (see Table 1).
Considering all three of these variables at once (2002 score, 2003 score, and growth score), a One-Way Analysis of Covariance with 2002 Reading score as a covariate, revealed a significant difference between the spring 2003 reading scaled scores (MCT) and the spring 2002 reading scaled scores, $F(2, 96) = 44.15, p = .0001; \eta^2 = .479$ (See Table 2 & Figure 1). This was a particularly interesting finding considering observational and interview data determined that although training was given in all three subject areas and the original intent was for the program to be implemented in all three areas, the program was not used in language arts or mathematics. Perhaps teachers are more accustomed to using centers in reading and therefore implementation in this subject was natural. The 2003 and the 2002-2003 gain scores were only significantly different between the treatment and comparison groups for the subject in which the teachers actually used the program.

Table 1

MCT Reading Scores

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>Std D</th>
<th>Error</th>
<th>t</th>
<th>df</th>
<th>sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison</td>
<td>49</td>
<td>471.78</td>
<td>41.27</td>
<td>5.90</td>
<td>.183</td>
<td>80.98</td>
<td>.855</td>
</tr>
<tr>
<td>Treatment</td>
<td>50</td>
<td>470.50</td>
<td>26.17</td>
<td>3.70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison</td>
<td>49</td>
<td>485.82</td>
<td>40.07</td>
<td>5.72</td>
<td>3.166</td>
<td>94.03</td>
<td>.002*</td>
</tr>
<tr>
<td>Treatment</td>
<td>50</td>
<td>514.26</td>
<td>48.97</td>
<td>6.93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth (2003 less 2002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison</td>
<td>49</td>
<td>14.04</td>
<td>30.38</td>
<td>4.34</td>
<td>4.313</td>
<td>93.41</td>
<td>.000*</td>
</tr>
<tr>
<td>Treatment</td>
<td>50</td>
<td>43.76</td>
<td>37.84</td>
<td>5.35</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p< .01
Table 2

Univariate Analysis of Variance

<table>
<thead>
<tr>
<th></th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tx/Con (1=treatment; 0=control)</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>50</td>
</tr>
</tbody>
</table>

Levene’s Test of Equality of Error Variances
Dependent Variable: MCT Rdg 03 - scaled score

<table>
<thead>
<tr>
<th>F</th>
<th>df1</th>
<th>df2</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.465</td>
<td>1</td>
<td>97</td>
<td>.066</td>
</tr>
</tbody>
</table>

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+MCTR02+GROUP

Tests of Between-Subjects Effects
Dependent Variable: MCT Rdg 03 – scaled score

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>102821</td>
<td>2</td>
<td>51410.64</td>
<td>44.152</td>
<td>.00</td>
<td>.479</td>
</tr>
<tr>
<td>Intercept</td>
<td>5250</td>
<td>1</td>
<td>5250.86</td>
<td>4.510</td>
<td>.04</td>
<td>.045</td>
</tr>
<tr>
<td>MCTR02</td>
<td>82799</td>
<td>1</td>
<td>82799.53</td>
<td>71.110</td>
<td>.00</td>
<td>.426</td>
</tr>
<tr>
<td>GROUP</td>
<td>21564</td>
<td>1</td>
<td>21564.79</td>
<td>18.520</td>
<td>.00</td>
<td>.162</td>
</tr>
<tr>
<td>Error</td>
<td>111781</td>
<td>96</td>
<td>1164.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>24982606</td>
<td>99</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>214602</td>
<td>98</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R squared = .479 (Adjusted R squared = .468)

Fixed = Treatment/Control; Covariate = MCT 02 Reading
Figure 1. Comparison of MCT reading score for treatment and control (comparison) groups for 2002 and 2003.

Mississippi Curriculum Reading Test Scores 2002 & 2003

A chi square was also done to compare the four proficiency levels (See Figures 2 & 3) of the MCT for 2002 and 2003 [Pearson $\chi^2(2, N = 98) = 7.69, p < .05$]. A Cramer’s V (.2801) indicated a small effect. However, when the lower two categories (minimal and basic) and the upper two categories (proficient and advanced) were combined, the $x^2 = 9.96$ and the Cramer’s V indicated a medium effect (.3188). Although many factors could be responsible for the significant difference in reading score change, the results do suggest that the LeapTrack program may indeed serve as an effective tool for improving student learning.
Figure 2. Frequencies for comparison (control) group proficiency levels of MCT reading scores for 2002 and 2003.

Control Proficiency Levels

Figure 3. Frequencies for treatment group proficiency levels of MCT reading scores for 2002 and 2003.

Treatment Proficiency Levels
Conclusions

The LeapTrack system was shown to be an effective technology-based program when using a state-designed test as a measure of growth. Further analysis of time-on-task reports would allow researchers to measure gain and time-on-task in high implementation groups compared to groups for which less time is dedicated to the program. A longer period of instruction over an entire school year would provide even more data for review. It would be interesting to repeat this study with greater control over teacher use in all three subject areas and to track time-on-task for program use to examine achievement increase or decrease as related to actual amount of interaction with the program time.

A longitudinal study would help determine whether or not the novelty effect is present as well as allow some insight into how students would perform over several years using the LeapTrack system. Since the program is available for grades K-5, a larger study, which covers more grade levels, is in order. It is clear that technology is a part of our educational system in the twenty-first century. A well-designed, computer-based instructional program can be a valuable addition to current methods of teaching.

Computers cannot, however, replace the time and expertise of a well-qualified teacher. Teachers are just as important in teaching children as they were one hundred years ago, now they simply have another way to analyze and understand student learning needs. CAIs allow teachers and administrators to step back and make an unbiased analysis of instruction and student progress (Fox, 2004).

References


LeapFrog SchoolHouse. The Research Basis for the LeapTrack Assessment and Instruction System: Formative Assessment and Student Learning.

LeapFrog SchoolHouse. Website: http://leapfrogschoolhouse.com/do/findpage;jsessionid=487B117F808886F0C9CEF650633013ED?pageKey=home


Imagination is more important than knowledge.

- Albert Einstein

32
Using CAPSI to Increase Student Engagement in General Psychology

Heidi L. Eyre
Darlene E. Crone-Todd
J. Reid Jones
Scott A. Hutchens
Delta State University

Joseph J. Pear
University of Manitoba

Abstract

Because General Psychology is a course in which nearly all DSU students enroll, it allows the psychology faculty a unique opportunity to increase campus-wide student engagement through a variety of innovative educational approaches such as the use of technology in the classroom. In Spring 2005 two sections of General Psychology used a modified PSI web-based course management system called CAPSI (Computer-Aided Personalized System of Instruction). The course centered on sets of study questions designed to guide students' readings of the textbook. Student-generated answers to these study questions were discussed in class and students could also receive feedback on their study question answers through a series of mastery-based unit tests within CAPSI. Subsets of these study questions were also used for the exams. Feedback from students indicated that they enjoyed using CAPSI and were very engaged in the course. The course modifications for Fall 2005 to further increase both student involvement and investment are also discussed.

University professors often look for new methods of teaching that will simultaneously increase student engagement and focus students on relevant content areas. One method of instruction that has consistently been shown to both increase student engagement and increase learning of content knowledge is Keller’s Personalized System of Instruction (PSI; Keller, 1968). Kulik’s (1976) meta-analysis of PSI classes revealed that it was not simply the high aptitude students who were outperforming other students in PSI classes; rather students with low aptitudes also outperformed their counterparts by nearly seven points. Students enrolled in PSI courses typically score significantly higher than students enrolled in traditional lecture classes with regard to multiple choice examinations (a 7-point average difference), essay examinations (a 12-point average difference), and final examinations (an 8-point average difference) (Buskist, Cush, & DeGrandpre, 1991).

With regard to student engagement, students report spending more time on PSI courses, being more satisfied with PSI courses, and having higher rates of course
completion (Kulik, Kulik & Cohen, 1979). Students enrolled in PSI courses also tend to be more involved in the course because in a self-paced system students must remain engaged in learning the material more consistently over the course of the term. In contrast, in courses where term tests or exams are the main mode of assessment, students tend to cruise-and-cram, with studying to learn material only occurring a few times per semester (Greenspoon, 1974).

What is it that makes PSI such an effective teaching method? It is primarily the following components that contribute to its success: (1) student self-pacing, (2) mastery-based units of study, (3) use of lectures and demonstrations primarily for motivational purposes, (4) stress upon the written word, and (5) the use of undergraduate mentors (Keller, 1968). The self-pacing feature of PSI courses allows students to proceed at their own pace within the curriculum so they can move through material they understand more quickly and spend more time on areas they find difficult. This feature also allows students to flexibly allocate their time between courses. The unit mastery component requires that students learn a small quantity of information (about a week’s worth of material) and pass a test over this information with some criterion percentage of the answers correct (85%, 90%, 95%, etc.). If the student does not reach the mastery criterion then he or she restudies the material and retakes the unit test as many times as it takes for him or her to demonstrate mastery of the material. Course credit is awarded when the unit has been mastered and there is no penalty imposed for not passing a unit test on a given attempt. Within Keller’s system, the instructor is seen as the facilitator of learning in that classroom meetings should be used to help clarify answers for students and motivate them to be engaged learners. The actual learning of the material takes place outside of classroom meetings through students’ active reading of the textbook and supplemental materials. As mentioned above, students demonstrate their understanding of this material through mastery of the unit tests and instructors and mentors typically provide individualized written feedback on these unit tests to help students learn where their answers are on-target and off-target. Finally, undergraduate student mentors, typically upper-division students, are an important part of the PSI system. These students provide valuable one-on-one feedback to students about their unit test performance and provide individualized tutoring in areas where the student needs to improve.
Given the preponderance of evidence that PSI is one of the most effective forms of instruction, why have the number of PSI courses in the country steadily dwindled since the late 1970s? Many reasons have been suggested, but one of the most viable concerns is the labor-intensive nature of PSI courses. In the PSI courses of the 1960s and 1970s, one might imagine a room set up where there are many undergraduate mentors swarming around marking unit tests, providing feedback to students, running individual tutoring sessions, and so forth while the instructor is faced with a mountain of paperwork to organize in order to keep track of where each student is in the course. Although this created logistical nightmares for instructors of the time, this problem has essentially been solved by the modern desktop computer and aided more recently by the widespread use of the Internet (Brothen & Wambach, 1999-2000; Eppler & Ironsmith, 2004; Grant & Spencer, 2003; Pear & Crone-Todd, 1999; Pear & Kinsner, 1988).

The computer-aided personalized system of instruction (CAPSI) was developed by Pear and colleagues at the University of Manitoba, and in its current form, CAPSI is an Internet-based program that allows students to log on to the program from any computer with Internet access (Kinsner & Pear, 1988; Pear & Crone-Todd, 1999; Pear & Kinsner, 1988; Pear & Novak, 1996). At Delta State University, CAPSI is being used in several courses including General Psychology (PSY101), Research Methods in Psychology (PSY201), Developmental Psychology (PSY307), and Applied Behavior Analysis (PSY420). CAPSI incorporates all of the fundamental principles of PSI through a variety of functions.

In General Psychology, students are assigned a series of guided study questions to help them learn the textbook material. These short-answer study questions are broken into units. Although the answers to the study questions are reviewed during the seminar meetings, it is the responsibility of the student to have prepared answers to these study questions on his or her own time. Whether or not students are engaging in this activity is measured by a seminar performance mark they earn when their name is randomly drawn from a box and they are asked to give their answer to a given question (grades are earned based on whether an answer is complete and correct, close but not quite on-target, or off-target/not prepared). When students feel they know the material covered in a given unit, they log on to the CAPSI system and request a unit test. The computer generates a given test for the student based on three randomly selected questions from that unit of study.
The student answers these questions (or cancels the test if he or she does not feel prepared for the unit test) and submits the unit test for marking. The unit test then either goes to the instructor, one of the senior mentors in the course, or two other students who have previously mastered that unit (called “peer-reviewers” within the CAPSI system) for marking. When the test is sent out, the student is anonymous in that their name and course ID number is not revealed to reviewers, and the reviewers anonymously mark the test and provide written feedback based on whether they believe the student has demonstrated mastery of the material. In General Psychology mastery is defined as all three answers being complete and correct (realistically about a 90-95% correct criterion). If one or more answers do not meet the mastery criteria, the student is granted a restudy. The student is also provided with substantive feedback about where his or her answer was off track and how the answer can be improved. After 30 minutes of study time, the student becomes eligible to take another unit test over the same material. Only after a pass is granted, meaning all answers are assessed as complete and correct, may the student move on to the next unit. As mentioned above, once a student passes a given unit he or she may sign on to be a peer-reviewer. Each test review earns the reviewer bonus points in the course. Exams were given in Spring 2005 at set times during the course (each exam covered three units of material) and they were created from the unit study question pool that formed the basis for the unit tests and class discussion. Total course grades in General Psychology were determined based on number of unit tests passed, seminar points, exam points, peer review points (counted as bonus points), and several other activities that counted for a small number of points.

There were 94 students who initially enrolled in the CAPSI taught sections of General Psychology in Spring 2005 and 72 students who completed the course. The first author was the instructor of both CAPSI taught sections. Self-reported student demographic information indicated there were 46 males and 48 females in the course, with 40% of the students self-identifying as African Americans and 57% self-identifying as Caucasian.

In examining students’ feedback to the CAPSI - taught course, the feedback was overwhelmingly positive. The mean overall course enjoyment rating on a 5-point scale was 3.63 with a median of 4.00. When asked whether they would consider taking another CAPSI-taught course in the future, 76% of students indicated they would. Thus it
appears that the majority of students indicated that they enjoyed the course and would consider enrolling in another CAPSI-taught course.

One of our future endeavors within the CAPSI system will focus on how to use this framework to raise the course performance level of the lower performing students. Although lack of motivation may be a factor, there are many other factors to consider as well. For example, lower performing students may not have the same degree of computer savvy as higher performing students and thus may feel intimidated by the use of technology in the classroom. This is an important issue to address and rule out as a potential cause because of Delta State University’s heightened focus on the use of technology in the classroom. Additionally, although the restudies are not meant to be punishing to students, it appears that students who have not yet developed their writing skills appear to find restudies to be punitive and fail to persist in pursuing a pass on a unit test (Crone-Todd, Eyre, Hutchens, Jones, & Pear, 2005). This may be especially true for the lower performing students who are more likely to get restudies on their first attempt.

In order to increase student engagement with CAPSI even further, the General Psychology course was modified slightly for Fall 2005. One modification was based on the finding described above that restudies tend to decrease some students’ persistence at passing CAPSI unit tests. Crone-Todd and colleagues (2005) demonstrated that when the mastery criteria were changed to include a revision option, student persistence at passing unit tests increased. Thus, in Fall 2005, the use of a “conditional pass” option for “close” answers was adopted by the first and second authors. This option allows the instructor to give a “conditional pass” if one of the answers is not quite complete and correct, but is only missing a minor detail or two. The student is given a restudy decision, but is encouraged to use the appeal function to submit a revised answer. One advantage of this option is that it allows the instructor to assist students who understand the basic concepts, but whose writing skills need to be improved.

A second modification centered on the point scheme. In Spring 2005 the course was based on a 100-point scale. Many students misunderstood the impact of unit tests and peer reviewing incentives. The unit tests had been worth 1 point each and each completed peer review was worth 0.5 bonus points. The instructors adjusted the point scheme in Fall 2005 to be based on 400 points. To acknowledge the work students put in to the passing of each unit test, the percentage was raised from 1% to 2%. However,
peer-reviewing points were decreased (from 0.5% to .0375% or 1.5 points per review). Thus far, students seem satisfied with the amount of points they are earning for both passing unit tests and peer reviewing.

The self-pacing of examinations was the third modification made to the course. In the previous semester, the unit tests were self-paced but the exam schedule was rigid. The instructor-controlled examination pacing is not consistent with the PSI criteria set forth by Keller because it requires that students conform to the schedule set by the administration and the instructor which may conflict with the students’ schedules. In Fall 2005, the instructors made the modification such that students who had completed the unit tests associated with a given exam before the scheduled examination date may take an early attempt at the exam. They may take multiple attempts at this examination up until and including the day of the regularly scheduled exam. If the student meets mastery criteria (an 85% or higher on a given exam), they may continue moving through the course at their own pace. If they do not meet the mastery criteria then they may either take another early attempt or wait until after the regularly scheduled exam date to attempt the next examination. This allows those high performing students to move quickly through the course and potentially finish it before the end of the semester. Lower performing students may also take advantage of this because it allows them multiple attempts at the examination and only their highest score is recorded. This should discourage student procrastination because many want the opportunity to either take multiple attempts or try to finish the course early. However, it does require that students finishing the course early have an “A” in the course (an 85% or higher plus the unit test points will average out to a 90% or higher). Although this innovation is self-paced in the sense that it allows students to move ahead in the course at their own pace, it still has regularly scheduled examination times during which all students must complete the exam. This keeps students from falling behind in the course and keeps them on track with the university schedule.

In conclusion, CAPSI appears to have been strongly endorsed by General Psychology students at Delta State University. These students engaged in a great deal of writing in this course as well as critical thinking (many of the study questions assigned are higher-order thinking questions; Bloom, 1956; Crone-Todd & Pear, 2001; Crone-Todd, Pear, & Read, 2000). Future research will focus on specific long-term learning
outcomes associated with CAPSI such as performance in future writing-intensive courses and other CAPSI-taught courses. Additionally, the instructors will continue to modify the General Psychology course to help all students be successful in this course and master the material.

References


Author Note

Portions of this paper were presented at the Reaching Out to Mississippi: Education in Action faculty development conference held in Cleveland, MS, September 28, 2005.

*The only limit to our realization of tomorrow will be our doubts of today. Let us move forward with strong and active faith.*

- Franklin D. Roosevelt
Deep Structure Learning and Statistical Literacy

David A. Swanson
University of Mississippi

Abstract

The acquisition of statistical literacy represents deep structure learning, a process that leads to reflective thinking, which is the ability to reason about unstructured questions that have no absolute answers. Although there are introductory statistics courses that have been developed as a path to deep structure learning, this possibility is still not being fully exploited by those teaching these courses, especially service courses taught within specific disciplines such as economics, psychology, and sociology. I recommend that faculty not employing the deep structure learning process in their statistics courses, inspect material from statistics courses organized explicitly to facilitate deep structure learning and consider using these ideas. An important benefit of this approach appears to be greater satisfaction among students taking a statistics course designed using these principles.

Deep Structure Learning

As Peter Facione (1998) has observed, a major goal of liberal education is the development of critical thinking, a set of skills essential for a rational and democratic society. And it is this same rational and democratic society that forms key assumptions serving, in turn, as the foundations of liberal learning (Anschicks, n.d.).

In attempting to develop critical thinking skills, many of the courses comprising the liberal education curriculum are either implicitly or explicitly designed to implement a process that has been termed “deep structure learning” by Keith Roberts (2002). Using a framework developed by Patricia King and Karen Kitchener (1994), Roberts views the development of critical thinking as a progression that starts with a stage labeled “pre-reflective thinking,” moves to “quasi-reflective thinking”, and culminates in the third and final stage, “reflective thinking.” King and Kitchener (1994) describe the final stage of the process - reflective thinking - as the ability to reason about unstructured questions that have no absolute answers.

King and Kitchener (1994) organize each of the three major stages in the process according to the manifestation of three issues: (1) epistemology; (2) an understanding of causality; and (3) evidence and its connection to logical reasoning and conclusions. Outlined in Table 1 are the three major stages and the manifestations of the issues that characterize them.
Table 1. The Three Major Stages of Deep Structure Learning and Their Issues

<table>
<thead>
<tr>
<th>STAGE / ISSUE</th>
<th>Epistemology</th>
<th>Causality Understanding</th>
<th>Evidence, Logical Reasoning and Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Reflective</td>
<td>Authority-oriented</td>
<td>Dualistic</td>
<td>Anecdotal, no connections to logical reasoning and conclusions</td>
</tr>
<tr>
<td>Quasi-Reflective</td>
<td>Relativism</td>
<td>Context-specific</td>
<td>Idiosyncratic, often unconnected to logical reasoning and conclusions</td>
</tr>
<tr>
<td>Reflective</td>
<td>Commitment to an understanding of epistemology</td>
<td>Complex</td>
<td>Probabilistic Commitment to a particular interpretation</td>
</tr>
</tbody>
</table>

Statistical Literacy

What is the relationship of deep structure learning to statistical literacy? First and foremost, to be statistically literate is to be able to reason analytically, an activity that has been identified as the most important intellectual process in liberal education (Atkinson, Swanson, and Reardon, 1998). Milo Schield (1999) states that statistical literacy is a liberal art – not a mathematical science. He equates critical thinking directly with statistical literacy and differentiates between the two by using the term “critical thinking” when an argument is primarily based on words and “statistical literacy” when it is primarily based on numbers. In both cases, he points out that while the method of reasoning includes deduction, it is primarily and fundamentally inductive. Seen in this light, the acquisition of statistical literacy represents deep structure learning, a possibility sometimes not fully exploited by those teaching it and often unrealized by others.

Statistical literacy can be viewed from the standpoint of six related but distinctive conceptualization domains that I argue foster the development of critical thinking: (1) alternative explanations and solutions to problems, (2) manipulation of symbols, (3) ambiguity and uncertainty, (4) causality as a multivariate process, (5) probabilistic interpretation, and (6) internal and external dialogues needed to
communicate analytical results. In terms of the three stages of the deep structure learning process, each of the six domains has a characteristic feature, as shown in Table 2.

**TABLE 2. Conceptual Domains and Deep Structure Learning Stages**

<table>
<thead>
<tr>
<th>DOMAIN/STAGE</th>
<th>Pre-Reflective</th>
<th>Quasi-Reflective</th>
<th>Reflective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternative Explanations</td>
<td>Not sought</td>
<td>Sometimes sought</td>
<td>Actively sought on a routine basis</td>
</tr>
<tr>
<td>Manipulation of Symbols</td>
<td>Poorly done</td>
<td>Somewhat well done</td>
<td>Very well done</td>
</tr>
<tr>
<td>Ambiguity and Uncertainty</td>
<td>Highly stressful</td>
<td>Somewhat stressful</td>
<td>Not very stressful</td>
</tr>
<tr>
<td>Causality as a Multivariate process</td>
<td>Rarely considered</td>
<td>Sometimes considered</td>
<td>Routinely considered</td>
</tr>
<tr>
<td>Probabilistic Interpretation</td>
<td>Rarely used</td>
<td>Sometimes used</td>
<td>Routinely used</td>
</tr>
<tr>
<td>Internal and External Dialogues</td>
<td>Rarely used</td>
<td>Sometimes used</td>
<td>Regular and on-going</td>
</tr>
</tbody>
</table>

**Developing a Deep Structure Statistics Course**

Unfortunately, as many readers will recognize based on their own experience, far too few of our introductory statistics courses have been designed using the principles of deep structure learning (Patten and Swanson, 2003). However, it is worthwhile to point out to those who have not taught an introductory undergraduate statistics course for some years that the implementation of deep structure principles in these courses has started (Bennett and Briggs, 1998; Patten and Swanson, 2003; Rossman, 2002; Schield, 1999; Swanson, and McKibben, 1998; Wolfe, 1993). This implementation is not often made explicit, however. It is usually implied in the form of “judgments” that students must learn how to make. In the syllabus for his introductory undergraduate course, Allan Rossman (2002), for example, tells students:

“In contrast to most mathematics courses, we will be using such phrases as ‘there is strong evidence that…’ and ‘the data suggest that…’ rather than ‘the exact answer is…’ and ‘it is therefore proven that.’”
To assist in the explicit identification and incorporation of these principles, Table 2 presents the development of critical thinking on the part of students using a set of identified conceptual domains. As such, it can serve several purposes. For example, it provides an informal organizational framework for course development and can be used in formally assessing the effectiveness of different pedagogies designed to develop critical thinking. As an example how this can be used, consider the following case study (Swanson & McKibben, 1998) that describes in detail an introductory statistics course that is aimed at increasing student learning and retention. Exhibit 1 shows the structure of the case study.

EXHIBIT 1. Case Study Structure of the Example.

1. The Decision-maker: The Decision-maker is the entity responsible for selecting a course of action - making the decision. The decision-maker could be a person, or a group of persons, informally or formally constituted. It could also be something designed by people - computer algorithm, for example.

2. The alternative decisions (courses of action): The decision involves selecting one of two or more identified courses of action. The goal is to choose the course of action that is "best."

3. Events: These are occurrences that are beyond the control of the decision-maker, but yet can have an effect on the course of action selected. The events are subject to uncertainty but ideally they are mutually exclusive and exhaustive so that one and only one can occur.

4. Return or Payoff: is a measure of net benefit to the decision-maker.

5. Uncertainty: is measured by the probabilities assigned to the identified events. This can include both objective and subjective assessments.

The Problem is set up to the students as follows.

You have been recently employed by Mätitahna OY, which is investigating the optimal pricing for Baltic roe paste that it plans to sell in Singapore. Your company knows that there are approximately 2,000,000 households in Singapore (This figure represents the total size of the market). It also has the results of a random survey of 600 households in Singapore. In this survey, a set of questions was asked about two different prices that people would be willing to pay for a 190 gr. tube of roe paste. You know that your company follows a policy of maximizing gross sales revenue in terms of selling price when it opens a new market but it does so using a conservative approach. Thus, in Singapore,
your company would select a price that generates the maximum amount of sales revenue, given the probability of expected sales at a given price. You also know that Máttahna OY uses the "95 percent confidence interval" in establishing three potential levels of gross sale revenues (high, average, and low) when using survey data to forecast expected sales revenue. The company's executives then apply a "rule of thumb" that assigns probabilities to the attainment of high, average, and low gross sales revenues. This rule of thumb is based on the cumulative experience gained from entering new markets and has worked well. The rule of thumb is comprised of three decision points: (1) there is a 20 percent chance of obtaining high gross sales revenues; (2) a 50 percent chance of obtaining average gross sales revenues; and (3) a 30 percent chance of obtaining low gross sales revenues.

Your first assignment is to use the information from the survey in conjunction with company policy and your analytical skills to recommend which of the two prices found in the survey should be used in initially selling roe paste in Singapore. This recommendation is to be written in the form of a brief (approximately 2-3 pages) memo to the Director of Marketing and Sales. The memo should also include the justification for selecting the particular price you recommend. The data are shown in Exhibit 2.

EXHIBIT 2. Case Study Data

"How much would you be willing to pay for a 190 gr. tube of high quality Baltic Roe Paste?

<table>
<thead>
<tr>
<th></th>
<th>95% C. I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion Willing to Pay</td>
<td>Proportion of Households That Would Buy, By Price</td>
</tr>
<tr>
<td>This Amount</td>
<td>High</td>
</tr>
<tr>
<td>$1.00</td>
<td>.75</td>
</tr>
<tr>
<td>$1.50</td>
<td>.53</td>
</tr>
</tbody>
</table>

A Solution. In this problem, the "alternative acts" are the two different selling prices, S$1.00 and S$1.50. The "events" are represented by the high, average, and low estimates of gross sales revenues generated by the proportion willing to buy at each price as measured by the high and low ends of the 95 % confidence interval of the estimated (average) market share in the survey. For example, the "low" estimate of gross sales that would result from pricing the roe paste at S$1.00 is: ( 2000000) *(.75 - (2*.04))*S$1.00 = S$1,340,000. This value reflects the objective probability (sampling error) of the low value of market share – we
are 95% certain that at least this gross level of sales will be achieved. From the preceding information, we can construct a table that shows the "Expected Gross Sales Revenues Under Uncertainty" as shown in Exhibit 3. This table will yield the Expected Gross Revenues under each of the alternative prices, taking into account the statistical uncertainty resulting from the use of a sample.

**EXHIBIT 3. Expected Gross Sales Revenues Under Uncertainty**

<table>
<thead>
<tr>
<th></th>
<th>Alternative Act 1 (Price= S$1.00)</th>
<th>Alternative Act 2 (Price= S$1.50)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gross Sales Revenues (GSR=E*S$1.00)</td>
<td>Weighted Gross Sales Revenues (WGSR)=P*GSR</td>
</tr>
<tr>
<td>Event</td>
<td>Probability (P)</td>
<td></td>
</tr>
<tr>
<td>(High Sales) 200000* .83</td>
<td>.20 S$1660000</td>
<td>S$332000</td>
</tr>
<tr>
<td>(Average Sales) 200000* .75</td>
<td>.50 S$1500000</td>
<td>S$750000</td>
</tr>
<tr>
<td>(Low Sales) 200000* .67</td>
<td>.30 S$1340000</td>
<td>S$402000</td>
</tr>
</tbody>
</table>

Expected Gross Sales Revenues: S$1,484,000

<table>
<thead>
<tr>
<th></th>
<th>Alternative Act 1 (Price= S$1.00)</th>
<th>Alternative Act 2 (Price= S$1.50)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gross Sales Revenues (GSR=E*S$1.00)</td>
<td>Weighted Gross Sales Revenues (WGSR)=P*GSR</td>
</tr>
<tr>
<td>Event</td>
<td>Probability (P)</td>
<td></td>
</tr>
<tr>
<td>(High Sales) 200000* .63</td>
<td>.20 S$1260000</td>
<td>S$252000</td>
</tr>
<tr>
<td>(Average Sales) 200000* .53</td>
<td>.50 S$1060000</td>
<td>S$530000</td>
</tr>
<tr>
<td>(Low Sales) 200000* .43</td>
<td>.30 S$860000</td>
<td>S$258000</td>
</tr>
</tbody>
</table>

Expected Gross Sales Revenues: S$1,040,000
The events are very simple, as are the acts and remain so even when coupled with the uncertainty associated with the “Estimated Market Share” and “Rule of Thumb Probability.” One can see from the table that no information is available for certain combinations of events and acts because of how the survey was constructed and the criteria used by the company. For example, no gross sales revenues are "expected" in excess of S$166,000 if the price is set at S$1.00. In principle, of course, an unlimited number of events could occur that would affect the payoffs of alternative decisions. However, to make the problem manageable, the company has established a range of potential prices that it believes is realistic and subjected this range of prices to an empirical examination - a sample survey of 600 households randomly selected from the 2 million households in Singapore. Consequently, one can think of "demand" as being represented by only three levels - high, average, and low - that are taken from the sample of households and translated via the 95% confidence interval apparatus into the numbers expected to purchase the paste at the two prices pre-selected by the company as being realistic. Note that these levels of demand are set up to correspond with the “rule of thumb probability” used by the company to gauge potential gross sales revenues. How does one compare the alternative acts?

A number of different criteria exist for selecting the "best alternative course of action," given uncertainty. If the company knew with exact certainty what the payoff outcomes would be, the decision procedure would be very simple: Select the act that yielded the highest payoff. Unfortunately, "certainty" does not exist. This is where statistical inference can help in combination the policies used by the company and the rule of thumb probabilities used by its executives. We may not know with exact certainty what will happen but we do have some guidance from the sample survey and the experience summarized in the rule of thumb probabilities.

One choice that could be made is for the decision-maker to assume that once a course of action is selected, things could turn out to be terrible and the "real world" may turn out to be the worst possible for the decision-maker. If this was the company's policy, then the choice would be based on the assumption that only 43% of all households would purchase (the worst set of events), which happens when the price is set at S$1.50. This yields an expected payoff of S$860000 in gross sales revenues.
Because we are using information from a random sample, we are not limited to assuming the worst. We can be very confident that, in fact, more than 43 percent of the households will purchase the roe paste, given that we sell it at S$1.00 instead of S$1.50. By selling at S$1.00, we are 95 percent certain that at least 67 percent of the households will purchase. At this sales level, we minimally expect gross sales revenues of S$1340000, which exceeds the minimal gross sales revenue of S$860,000 expected if we price the paste at S$1.50. Given company policy, it appears that choosing a price of S$1.00 is the best "act."

A quick example of yet another way in which the choice could be made is to take into account the probabilities (based on prior experience) assigned by the company executives to the high, average, and low sales. By summing up the "weighted gross revenues" by sales level, we can obtain an "average" for "expected gross sales revenues" for each of the two prices. Doing this, we see that by setting the price at S$1.00, we "expect" an average gross sales revenue of S$1,484,000. This figure is higher than the expected gross sales revenues that result by summing up the weighted gross revenues by sales level at the S$1.50 price, which yields S$1,040,000. This suggests that by taking into account both the objective probabilities associated with the survey and the more subjective probabilities used in the rule of thumb by the executives to estimate sales levels, the price to select would be S$1.00 because it yields a higher average gross sales revenue figure than would the price of S$1.50, given the experience of the company as summarized in its rule of thumb for estimating sales in new markets.

This example case study is of interest to consider as a teaching device because even with the “reduction of complexity” allowed by the experiential rule of thumb and the nature of the problem, there is no single answer. Not surprising, when this case was assigned to teams in an in-class “case competition” environment for solution and presentation (for grades) before the instructor, there have been equally strong, but different solutions. In turn, this leads to more in-depth discussion of the case, which contributes to an even deeper understanding of the principles and issues underlying it. This is why this type of case is supportive of deep structure learning – the ability to reason about open-ended questions that have no definite answer. This is an optimal learning outcome for a university education, whether in liberal arts, science, or a professional field.
The process identified in Table 2 also can be used as a point of departure for refining ideas about the development of critical thinking (Swanson & McKibben, 1998). For example, in the introductory statistics course I regularly taught undergraduate students in business administration from 1995 to 2003 at the Helsinki School of Economics, the syllabus stated that the final exam (open-book, open-note) required an answer in the form of a memorandum that outlines the reasons underlying a decision regarding a course of action that should be taken in regard to the problem provided and its data. The final exam always involves a multivariate problem, uncertainty, and probabilistic interpretations. On the review day for the exam, I would inform students that there is no single right answer to the problem they will encounter and that they must connect evidence to their conclusions through logical reasoning. The final exam was explicitly designed to force students to reason about unstructured questions that have no absolute answers – that is, to make them reflective thinkers.

**Concluding Remarks**

The acquisition of statistical literacy represents an important pathway to deep structure learning and the introductory courses in statistics commonly taught in schools such as applied sciences, business administration, behavioral sciences, education, health sciences, life sciences, physical sciences and social sciences, represent important vehicles on this pathway. I encourage administrators, faculty, and students to consider this journey and to employ critical thinking in evaluating it.

There is, in addition, an important benefit that appears to accrue from the Deep Structure Learning approach: Students view it more positively than statistics courses not designed using it. Part of this reason may be that the Deep Structure Learning approach naturally accommodates other features associated with the development of critical thinking skills, such as the use of real world problems, the emphasis of concepts over mechanics, writing and presentation skills, active cooperative learning and the **worthwhileness** of a course (Patten and Swanson, 2003; Swanson and McKibben, 1998). For example, over a recent six year period at the International Business Program of the Helsinki School of Economics, the introductory statistics courses designed around deep structure learning received an average rating of 8.9 (out of a possible 10 points, n=9) compared to an average rating of 6.6 points for the same introductory statistics courses not designed in this manner (out of 10 possible points, n = 8). Further, informal
evidence collected by the author indicates that the students not only learn concepts but retain a *structural* picture of statistics that allows them to quickly regain a good understanding of individual aspects of statistics in subsequent courses and in working life after graduation.

In conclusion, keep in mind that the deep structure learning process needs to be considered *reflectively*. That is, deep structure learning is an uncertain, complex process that requires epistemological understanding and probabilistic commitments to its interpretation and that implementing the deep structure learning approach in an introductory statistics course involves working on a real world problem in an active cooperative learning environment.

**Endnotes**

1. This is an important point because statistics is a required course for students in a broad range of disciplines at both the undergraduate and graduate levels (Gordon, 1995) and, as noted by Sowey (1995), it is in these service courses that the vast majority of statistics students are found. Unfortunately, however, evidence strongly suggests that the students in these courses are not particularly happy to be there (Dillon, 1982; Hogg, 1991; Snee, 1993; Watts, 1991) and, in fact, students often rate them as the worst course or the most useless course they have ever taken (Romero et al., 1995). These assessments are typically associated with courses not explicitly designed to accommodate the deep structure thinking approach described here (Patten and Swanson, 2003; Swanson and McKibben, 1998). Even more unfortunate, the available evidence is that the students both learn and retain little of the statistics they do encounter in the statistics courses not designed along the deep structure thinking approach (Hogg, 1991; Snee, 1993; Sowey, 1995; Yilmez, 1996).

2. The author would be pleased to supply a syllabus and all materials (e.g., power point slides) used in this course on request.

**References**


**About the Authors**

Darlene E. Crone-Todd, Ph.D., University of Manitoba, Assistant Professor of Psychology, Delta State University. dctodd@deltastate.edu

Heidi L. Eyre, Ph.D., University of Kentucky, Assistant Professor of Psychology, Delta State University. heyre@deltastate.edu

Scott Alan Hutchens, Ph.D., Texas Tech University, Associate Professor of Psychology and Psychology Program Coordinator, Delta State University. shutchen@deltastate.edu

J. Reid Jones, Ph.D., University of Tennessee at Knoxville, Professor of Psychology and Academic Research Coordinator, Delta State University. rjones@deltastate.edu

Maud A. Kuykendall, Ed.D., University of Northern Colorado, Assistant Professor of Special Education, Delta State University. mkuykn@deltastate.edu

Joseph J. Pear, Ph.D., Ohio State University, Professor of Psychology, University of Manitoba. pear@cc.umanitoba.ca

David A. Swanson, Ph.D., University of Hawaii, Professor of Sociology and Chair of the Department of Sociology and Anthropology, University of Mississippi. dswanson@olemiss.edu

Lynn Walton Varner, Ph.D., University of Mississippi, Assistant Professor of Leadership and Research and Director/Chair of the Center for Rural School Leadership and Research, Delta State University. lvarner@deltastate.edu