The Effect of Delivery Method on Conceptual and Strategy Development

Lisa C. Kaczmarczyk *, **Mary Z. Last** **, **Risto Miikkulainen** ***

* **kaczmarc@rose-hulman.edu**
Department of Computer Science and Software Engineering CM4007, The Rose-Hulman Institute of Technology
5500 Wabash Avenue, Terre Haute, Indiana 47804 USA
** **mlast@umhb.edu**
Department of Computer Information Systems, The University of Mary Hardin-Baylor
900 College Street, Belton, Texas 76513 USA
*** **risto@cs.utexas.edu**
Department of Computer Sciences, The University of Texas at Austin
1 University Station C0500, Austin, Texas 78712 USA

Abstract

In order to develop intellectual expertise, the student needs to learn how to perform sophisticated pattern identification, and how to employ effective study and test taking strategies. These cognitive requirements are complex and analytical, and formal instruction operates under challenging constraints. In order to help students succeed in their chosen field, we need to understand better how instruction can help develop these meta-cognitive skills. This paper reports the results of a study in which novices attempted to categorize calculus integration problems in one of three delivery methods (Drill and Test, Fully Integrated, Incremental Learning). The results demonstrate that Incremental Learners develop the most effective study and test-taking strategies, have the best conceptual development, and have the most positive reactions to learning. The results, together with a previously reported computational study, support the hypothesis that Incremental Learners develop more effective concepts and strategies.

Introduction

It is important to understand how expertise develops for two reasons. Cognitive scientists want to know how people become experts in order to understand human cognition at its most effective. Professional educators want to help students acquire the ability to become experts in their chosen field. Thus cognitive science and educational researchers share a common goal: to better understand how people think and acquire intellectual expertise. However, while cognitive science has made progress in understanding human development, we do not yet understand enough about how adult learners acquire complex concepts in response to different instructional methods. Learners' intellectual tasks are often extremely complex, and classroom environments are hard to study because they pose unique learning constraints.

The primary goal of this paper is to increase understanding of how pedagogical delivery methods can initiate the development of cognitive behaviors seen in experts. This paper presents results from a study in which novices studied calculus integration problems, and attempted to categorize them by solution strategy. The study focuses on how different pedagogical sequencing affects study and test-taking strategies, and conceptual development. The results provide insight into the role of complexity in instructional design. In particular, this study supports the prediction of a previously reported computational model (Kaczmarczyk and Miikkulainen 2004), a backprop neural network, which demonstrated that an Incremental Learning delivery method produced the best learning compared to Drill and Test or Fully Integrated learning.

Background

It is well established that in order to learn expertise, it is critical to develop successful schemas. A schema is "a construct that allows problem solvers to group problems into categories in which the problems in each category require similar solutions" (Cooper and Sweller 1987). Successful learners develop problem schemas by categorizing problems according to structural features, whereas less successful learners rely on surface features (Chi, Feltovich, and Glaser 1981; Schoenfeld and Herrmann 1982). As schemas develop more fully, categorization ability improves; Experts can categorize problems without solving them (Robinson and Hayes 1978). This process of acquiring schemas is inductive (Cummins 1992). Once a learner has accurate schemas, she or he is able to eventually recognize effective solution strategies (Cummins 1992; Owen and Sweller 1989).

Prior studies of schema acquisition in instructional settings have often focused on identifying best-case learning environments, based upon an understanding of how experts behave. For example, it is commonly accepted that expertise results from long-term practice (Hayes 1989) and strategy choices that are consciously goal-directed, self-monitoring, and self-adjusting within the setting of each particular task (Ericsson, Krampe, and Tesch-Römer 1993). These results imply that instruction should relax time limits, be individualized, and provide regular formative feedback. Unfortunately, university settings operate under restrictive constraints that challenge these pedagogical ideals directly. One of the biggest constraints on classroom learning is time. The semester system forces all students to attempt to learn at the same pace, and there is a fixed start and end to the entire learning experience. In addition, all learners receive the same material in the same sequence, and time and workload pressures often prevent instructors from providing timely feedback. These factors leave little room for individualized instruction.

To understand better how to help students become experts, there is a need to investigate what delivery methods encourage the best conceptual and strategy development within these constraints. How can we help learners to acquire effective schemas in an environment such as a university classroom? There are two promising areas where instruction can be adjusted. The first is sequencing. In formal classroom settings an important pedagogical issue is how to sequence the material. Posner and Strike (1976) propose a framework for analyzing and describing different content sequences. For example, a "concept-related sequence" structures delivery so that material is presented consistently with how the concepts

Published in *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, 2006

themselves relate to one another. They cite traditional deductive mathematics instruction as an example of a concept-related sequencing. However, traditional mathematics instruction often assumes that a “logic of inquiry sequencing” will take place: the learner will be able to generalize fundamental principles inductively from exposure to many examples. Thus traditional instruction in mathematics embodies a possible conflict between the actual concept-related presentation and desired inductive learning.

The second promising area is in encouraging learners to focus on structural analysis of problems. Learners are more likely to focus on the structural attributes of problems, and improve their schema induction, when forced to compare problems to one another without solving them (Cummins 1992). Learners are also more likely to focus their attention on structural analysis when “goal-free” problems are presented to them. Goal-free instruction groups multiple problem examples of the same category type with different surface features; problems with similar surface features are spread across different category types (Quilici and Mayer 1996).

The study presented in this paper provides learners with goal-free problems in three sequencing modes. Drill and Test proceeds in a blocked sequence. At the other extreme, in a Fully Integrated sequence, content presentation is mixed. The third presentation sequence is Incremental Learning, which presents content incrementally. The study presented in this paper presents a comparison of conceptual and strategy development between these three sequencing modes. The experiment tests the hypothesis that Incremental Learning initiates better conceptual and strategy development than Drill and Test or Fully Integrated learning. This study differs from previous studies in several ways. First, it is motivated by a computational model of human expertise learning that illustrates the effect on learning and performance of the three delivery methods (Kaczmarczyk and Miikkulainen 2004). Second, the current study explicitly acknowledges key limitations inherent to formal instructional environments and conducts an experiment to explore how schema development can be encouraged under these adverse conditions. Third, it incorporates goal-free problem presentation of mathematical expressions (previous work on goal-free problem solving has focused on statistical word problems; Quilici and Mayer (1996). The results support the hypothesis that Incremental Learners develop more effective concepts and strategize more effectively than either Drill and Test or Fully Integrated learners.

Human Subject Study - Method

Subjects and Materials

Fifteen volunteers (age 19-51, m = 35 years of age) took part in a one-hour exploratory learning study. All of the volunteers were undergraduate or graduate students at the University of Texas at Austin. Participants came from thirteen different academic departments including natural sciences, liberal arts and education. Volunteers responded to advertisements looking for people with an interest in either analytical thinking or mathematics, but who did not know calculus integration. Selected volunteers had all successfully completed pre-calculus and were screened to eliminate those with mathematics anxiety. Volunteers were not compensated for participating in the study. Each volunteer was randomly assigned to one of three protocols, known to the researcher as Drill and Test (DT), Fully Integrated (FI), Incremental Learning (IL).

Forty-five calculus integration problems were written individually as equations on 4x6 inch index cards. The problems were equally divided between three categories, labeled A, B, C. Alphabetic category labels were used in order to ensure that the participants would not attribute meaning to the category labels. The three category labels corresponded to the solution strategies Simple Integration, U-Substitution, and Integration by Parts. Integration problems were collected from college level calculus textbooks (Lang 1986; Stewart 1995). The solution strategy for each problem was determined by the textbook and confirmed by a calculus expert (mathematics faculty or TA). Scrap paper and a pencil were provided. A set of 4 examinations was created for each of the three protocols (12 examinations total). The examinations contained calculus integration problems that were not part of the study set. The integration equations were used according to the procedure described in the next section.

Procedure

The first part was a categorization task. Each volunteer was given an identical instruction sheet. The instructions told the participant that they would be given index cards with one integration problem written on the front of each card, and one of three categories written on the back of each card. Their task was to study the problems and try to identify common properties for each category. The instructions also informed the participant that there would be four timed study periods, each followed by a test; the tests would contain additional problems to categorize.

The length of the study sessions was the same for participants in all three protocols. The length of the first three study sessions was determined using pilot studies and achieved an optimal balance between applying time pressure, and allowing time to rapidly assess the situation and make an initial strategic decision. An important goal in selecting study times was to encourage the subjects to react instinctively to each delivery method. The first three study sessions were 2, 3, and 3 minutes long. The fourth session was only 1 minute long and simulated a “cram session”.

The delivery protocol determined the order in which the calculus problems were presented to the volunteers in each study session. The DT protocol received one category of problems only in each study session: Simple Integration, then U-Substitution, then Integration by Parts. In the cram session they received all three types of problems for study. The FI protocol received all three categories of problems during every study session. The IL protocol received first Simple problems, then Simple and U-Substitution, then Simple, U-Substitution and Integration by Parts problems in the third, and again during the cram session.

The fourth examination was identical in each set and consisted of 15 problems from all three categories of problem. The first three examinations varied as follows: for the DT protocol, the first examination contained only Simple Integration problems, the second examination contained only U-Substitution problems, the third examination contained only Integration by Parts problems. For the FI protocol, all three of the examinations contained problems from all categories.

Table 1: Initial Codes - Interview Analysis

Conceptual Development
C-Category Development
C-Focus on Complexity
C-Lack of Understanding

Strategy Development	
S-Desire to Solve	S-Looking for Rules
S-Comparing Group Items	S-Reliance on Instinct
S-Looking for Patterns	S-Reliance on Memory
S-Analytical Planning	

Table 2: Emergent Codes - Interview Analysis

Affective Reactions
A-Discomfort
A-Positive Feelings

The problems were the same on each test, but the presentation order was changed. For the IL protocol, the first test contained Simple Integration problems, the second test retained the Simple problems and added U-Substitution problems, and the third test retained the Simple and U-Substitution problems and added Integration by Parts problems.

The second part of the study gathered data via structured interviews. The interviews immediately followed the categorization task. Every participant was asked the same 14 questions about their behavior and experiences. Each question elicited information about strategy use, reasoning and conceptual development. Follow-up questions were permitted if they clarified previous responses. The interviews were tape-recorded and later transcribed verbatim. Each interview lasted approximately 30 minutes.

The interviews were analyzed both qualitatively and quantitatively. The qualitative analysis used a thematic approach (Kvale 1996), focusing on three dimensions: strategy development, conceptual development, and affective issues.

Potential analytical bias was addressed as follows (Chi 1997):

(1) All of the interviews were selected for coding in order to avoid bias in selection.

(2) Prior to analysis, an initial coding scheme and operational definitions for codes was developed (Table 1). The general code categories were thematic, and derived from the hypothesis being tested: Strategy Development and Conceptual Development. Those verbal units that described actions taken by the participants, and that were intended to help them study or take the tests, were coded as Strategic. Coded as Conceptual Development were those verbal units that described a cognitive state of understanding in regards to their task. An additional category for Affective Reactions emerged during analysis (Table 2). When analysis was complete, there were 12 codes.

(3) Each individual interview was read and analyzed independently by two researchers: the principle investigator who conducted the study and the second author.

(4) The two researchers compared their coding decisions and thematic analyses. When there were divergent findings, only those codings and themes were retained in which both researchers could agree. An inter-rater reliability rating of 90% was achieved.

The quantitative analysis followed guidelines for quantifying verbal data laid out by Chi (1997). Statistical analysis using ANOVA was performed upon the coded data to confirm or disconfirm any reliable differences between the DT, FI and IL protocols. As an additional validity check of the results, final test performance was analyzed. Final score distributions were evaluated, and patterns of errors studied, to look for learning trends.

Results - Qualitative Analyses

Drill and Test

The Drill and Test participants' actions were characterized by extreme nervousness. All participants in this group expressed discomfort and anxiety throughout the course of the study. This discomfort was expressed through behavior and language. For example, two of the students nervously asked the experimenter if she was going to use her masking tape "on them" (the tape was for hanging a Do Not Disturb sign on the door). During the interviews, most students were so anxious that they frequently had trouble expressing themselves:

Student: "and again, I'm, I'm not, I'm a little shaky even on how you, separate them into, these problems, how you separate, what's, you know, where do you put the [making swooping figures with her hands]"

Interviewer: "Parenthesis? That's what you are doing with your fingers there?"

Student: "yeah, yeah"(DT-05)

DT participants lacked organized strategies for studying the categorized problems. Instead, they relied on memory-based strategies, which they were aware were ineffective:

Student: "I never feel like I had really committed the entire category to visual memory...then I was trying to memorize, you know, what the different sets, because within each category it seemed like there were similar cards, sets. So then I was just trying to remember..."(DT-06)

In reflecting during the interviews on the failure of their strategies, three participants were convinced that they must have misinterpreted the instructions, one wondered if she was being tricked, and another complained that the task was unfair. In taking the tests, DT participants took one of two approaches: they either gave up and guessed randomly, or they chose all the same answer on each test. This same answer was the most recently studied concept.

The analysis revealed that DT participants lacked an understanding of how to categorize the integration problems. All DT participants said that they were unsuccessful at learning how to categorize problems; three of them said that they guessed, and had little confidence, on all the tests. One participant acknowledged her lack of understanding as follows:

Student: "I'm aware that my criterion, my criteria are very superficial, and not, I'm, I mean I can tell that they don't work appropriately. Like when I did the test I can tell that it's just, it's not the right criteria." (DT-02)

In summary, all of the members of the Drill and Test group displayed strong negative reactions to the task, relied upon ineffective memory strategies, and developed superficial understandings of the categories.

Fully Integrated

The Fully Integrated participants' actions were characterized by nervousness initially, with some decrease in anxiety over time. When the study began, all the subjects were extremely frustrated and overwhelmed. At the end of the first study session, one student burst out into hysterical laughter, one yelled that the task was "impossible! disaster! hopeless!" (FI-01). Another student froze during the first test; she simply sat and waited for the experimenter to return (because the tests were not timed, it was close to 15 minutes before the experimenter went to check on her). Another participant had this reaction:

Student: "[My] impulse, on the first test was to choose all As, because, partly out of frustration...I ended up just going across aesthetically A,B,C,B,A, making a zig-zag." (FI-02)

The analysis of the interviews revealed that three FI subjects gradually evolved a deliberate strategy to look for similarities within groups. Their strategies began to develop sometime after their initial anxiety had partially abated in the second or third study session. The other two FI participants had no specific strategy other than to "just look at them and see if there is anything like a pattern." (FI-05)

All FI participants reported that they mostly guessed on the tests, and predicted that they were not performing well. They did not believe that they understood the categorization task very well. However, four of them were confident that given a lot more time, they could learn to distinguish the categories.

In light of their reported success, four members of the FI group demonstrated an increasing awareness of how the integration problems were categorized. Analysis of the interviews revealed that this understanding was more implicit than explicit. For example, one subject reported that by the end of the study, she was noting regularities on the tests, although she was unsure what to do with this awareness. Another subject showed the beginning of intuitive understanding:

Student: "I noticed...sometimes with the cards, I was having some luck, like I would, after I sort through them all...by the third [study session] I would look at it and I would say "ok I think this is going to be an A: and it WOULD be, you know...but when I looked at the test...I didn't feel confident that I was able to identify those."(FI-03)

In summary, all members of the Fully Integrated group found the task frustrating, by the final study session three of them had developed search strategies, and four of them showed signs of increased understanding of the categories.

Incremental Learning

The actions of the Incremental Learning group were characterized by confidence and focus on the task. During both the study sessions and the interview, IL learners made few emotional comments. There was no evidence of fear or anxiety, expressed directly or indirectly through tone or body language. They described the study with words such as "insightful", "fun" (IL-01), "amused" and "stress-free" (IL-03). When pressed by the interviewer, IL participants admitted to being nervous at the start of the study, but reported that these feelings rapidly diminished. The IL group and the DT group shared the same first study session (Simple integration problems only), so it is reasonable that both groups were stressed at first. However, in contrast to the deteriorating attitudes of the DT group, the IL subjects' attitudes improved rapidly. When asked how well they felt they performed on the final examination, four students in the IL group replied with a positive numerical estimate (e.g. 75%). This response is in marked contrast to the DT and FI groups in which all but one participant gave negative verbal estimates (e.g. "pretty bad"). The IL subjects spent the bulk of their interviews confidently describing detailed analytic strategies that they employed to tackle the categorization task. Even when they were not confident that they had succeeded, they were generally confident that they had made solid progress and that given time they would be able to figure out how to categorize the problems. For example:

Student: "The first one [test], I was completely lost...and then [second study session], I was able to compare it and make the correlation...just understand how they were different from one another...the third test, it was insightful, it was a learning experience, I figured out that I still understood A and B...And so I knew, in the final study session...I knew I needed to focus on group C."(IL-01)

As the above quote demonstrates, study and test-taking strategies in the Incremental Learning group were highly organized and efficient. The participants developed individualized systems that identified sub-sets of problems on which to focus. They adjusted these sub-sets in response to new information and insights, or in order to focus on some features they were less sure about. Two participants systematically moved back and forth between comparing within a group and comparing between groups. They used this process to test and clarify understanding and to reinforce previous conclusions. A third subject devised a system in which she started analyzing the outer edges and general symbols of each problem and moved step by step into the center of the problem and more complex feature combinations. A fourth subject systematically chose two groups at a time to compare, removing from her sight those cards she wished to ignore.

One IL subject differed from the four participants just described, by choosing several successive strategies which relied on memorization and speed. In the interview, this student reported that she knew her strategies were not working. She claimed however, to be noticing some regularities on the tests, although she was unable to explain what she saw.

Analysis of the interviews revealed that the four "successful" incremental learners were gradually forming a deeper understanding of the integration problems. One student described her progress as follows:

"As we got deeper and deeper into what's a category B versus a category C, it started getting clear." (IL-02)

Another subject said that her understanding was "a little better...[then] a little better..."(IL-03). A third participant felt that she was on the cusp of a breakthrough: "I was looking for [describes features]...but I couldn't quite find that." (IL-01)

Another subject gave an example of her categorization when she described how to categorize a sample problem:

Student: "I would definitely put that in a C [Interviewer: why?] Because C was the ones that had e's in them. And, and besides, this is also a more complex of an equation, with both the co-efficient and the exponent. So, the one thing I was noticing about C, was at least to me, Cs had the e's and , and, any, like if you were raising it to a tan, like if the exponent was a tangent or something. Anything that started getting even more complex dealing with e's especially, I would put that in a C. For those reasons." (IL-02)

In summary, all members of the Incremental Learning group had a positive reaction to the categorization task, and all but one of them demonstrated a non-superficial understanding of the categorization. The successful group members employed highly efficient and analytical strategies which reduced the cognitive demands of the task.

Results - Quantitative Analyses

Learning Results

As expected, subjects in the Drill and Test (DT) and the Fully Integrated (FI) protocols made less cognitive progress than subjects in the Incremental Learning (IL) protocol. Subjects in the IL protocol showed statistically significant differences on several measures of Strategy Development, Conceptual Development and Affective Reactions. A One-Way ANOVA was conducted to examine the differences between the three delivery methods on each of the codes measured in the Qualitative Analyses. Mean values of analytical planning differed significantly between delivery method ($F(2,12) = 9.33$, $p < .01$). Post hoc Tukey HSD tests indicated that IL subjects had a statistically greater number of analytic strategies than either DT ($p < .01$) or FI subjects ($p < .01$). There was no significant difference between number of analytic strategies used by the DT and FI subjects. These results support the hypothesis that learners in an IL learning environment develop better meta-cognitive planning skills than either DT or FI learners.

Mean values of focusing on complexity differed significantly between delivery method ($F(2,12) = 4.56$, $p < .05$). Post hoc Tukey HSD tests indicated that subjects in the IL protocol reported a significantly greater number of conceptual descriptions that relied on complexity analysis than DT ($p < .05$) and FI users ($p < .05$). There was no significant difference in the use of complexity between DT and FI subjects. Mean values for lack of understanding differed significantly between delivery method ($F(2,12) = 11.03$, $p < .002$). Post hoc Tukey HSD tests indicated that subjects in the DT and FI protocols mentioned significantly far more times that they did not understand the problem than did IL subjects ($p < .05$). These results support the hypothesis that the IL delivery protocol supports cognitive development of complex concepts better than the DT or FI protocol.

Mean values of discomfort differed significantly between delivery method ($F(2,12) = 13.44$, $p < .001$). Post hoc Tukey HSD tests indicated that subjects in the DT and FI protocols showed significantly more expressions of discomfort ($p < .01$) ($p < .01$) than subjects in the IL protocol. There was no significant difference between expressions of discomfort between DT and FI subjects. These results support the results from the qualitative analyses that the DT and FI delivery techniques are highly stressful for learners, whereas the IL delivery technique is not.

Mean values of positive feelings differed significantly between delivery method ($F(2,12) = 5.57$, $p < .01$). Post hoc Tukey HSD tests indicated that subjects in the IL protocols showed significantly more positive reactions than subjects in the DT ($p < .05$) or FI protocol ($p < .05$). There was no significant difference between expressions of positive feelings between DT and FI subjects. These results support the results from the qualitative analyses that the IL delivery strategy produces a better environment for learning difficult concepts than DT or FI delivery strategies.

Performance Results

An analysis of score distribution on the final examination confirmed that IL subjects were making greater cognitive progress than DT or FI subjects. Although all of the final scores were somewhat low, the median final examination score for IL learners was highest (IL median score 53.33% compared to 46.67% FI and 40.00% DT). Median and IQR are reported because the data was slightly negatively skewed, and these statistics are the best index of typical performance under these conditions. Overall the IL learners performed more consistently than FI learners, as reflected in the interquartile range (IQR) of 30.00 for IL learners compared to 36.66 for FI learners. DT learners had not only the lowest median score (40.00%), but the smallest interquartile range (19.99), reflecting the homogeneous poor nature of their performance.

A frequency analysis of patterns of error made on the final examination provided additional confirmation that IL subjects were learning better than DT or FI subjects. There was no discernible pattern to type of error made by the DT subjects. This lack of pattern confirms their assertions that they were guessing randomly. Errors made by FI subjects confirmed their claims that they could identify most of the A category (Simple problems). Most of the errors made by the FI subjects were confusions between the more complex problems: U-Substitution and Integration by Parts (categories B and C). However, FI subjects often appeared to be fooled by the length of a problem. They often assumed incorrectly that longer problems had to be more complex. Finally, errors made by IL learners were spread fairly evenly across problem types. IL subjects were somewhat less likely than DT or FI subjects to be fooled by the length of a problem. This finding in particular indicates that the IL learners were beginning to acquire a deeper understanding of the structure of the problems. The final scores were low because the study sessions were short; however, the IL learners showed a clear trend towards starting to study and learn more effectively. These performance results complement the learning results, and support the hypothesis that IL learners are acquiring the best meta-cognitive skills for learning complex concepts.

Discussion

This study supports the hypothesis that Incremental Learning initiates better conceptual and strategy development than either Drill and Test or Fully Integrated learning. Perhaps the most important result is that under adverse learning conditions, an incremental sequencing of material appears to be correlated with the best schema development. Verbal data and statistical analysis show that IL learners analyze problems in terms of their structure, whereas DT and FI learners focus on surface level features. DT and FI learners are much more likely to admit that they do not understand the concepts. On examinations, DT and FI learners are more likely than IL learners to make random errors or to be fooled by surface-level features such as the length of a problem. This evidence of more advanced structural understanding implies that IL learners have the most well-developed problem schemas.

The IL learners develop the most sophisticated study and test-taking strategies as well. Incremental learners follow more individualized and successful learning paths than either Drill and Test or Fully Integrated learners. The IL learners are more self-aware and meta-cognitive than either the DT or FI learners. The IL learners develop analytical strategies for study and test-taking. IL learners make comparisons based upon the structural complexity of problems. They adjust these plans as they acquire more data. These strategies benefit performance because the IL learners are the only learners who improve by the end of the study. They also enjoy their learning experience, in spite of the adverse learning conditions. This result is important and supports the proposal by Tennyson and Nielsen (1998) that the role of affect should be explicitly included in studies of instructional sequencing and cognitive development.

Drill and Test learning appears to correlate with ineffective short term memory strategies. The Drill and Test learners have problems keeping more than one concept at a time in their memory. This result suggests that complex learning tasks such as college-level mathematics should avoid instructional sequencing that encourages students to rely on short term memory. In other words, the results reported in this paper support ceasing to teach using Drill and Test.

Fully Integrated learners initially experience cognitive overload. They are frustrated with the categorization task. Over time, there are signs of improvement: The subjects become more aware of feature differences, and become more comfortable with the material. These findings have implications for the college classroom. The high drop-out rates seen in many introductory mathematics and science courses may be due to students being overwhelmed with new concepts. If so, instructional delivery should avoid sequencing that mixes too many complex concepts. In other words, the results reported in this paper support not teaching with full immersion.

In the context of the literature on schema development and expertise learning, the appearance of improved strategies followed by improved performance, supports the claim that IL learners acquire more advanced categorization schemas (Cummins 1992; Owen and Sweller 1989). The IL learners are developing conceptual understanding and study strategies that should enable them to become successful at the categorization task, even under less than ideal learning conditions. Once these cognitive attributes are in place, they should continue to be successful learners even after formal instruction has ended.

Conclusion

The results reported in this paper provide new insight into how people learn complex cognitive tasks. In particular, they provide new insight into how instructional delivery method can affect initial strategy development, conceptual understanding, and emotional perceptions while learning. This empirical evidence supports the hypothesis that Incremental Learning initiates more effective conceptual and strategy development than either Drill and Test or Fully Integrated learning. These data provide cognitive scientists with additional knowledge about how cognition can be affected in complex learning environments. Professional educators now have greater knowledge about how they can help students to acquire the meta-cognitive skills necessary to be highly successful learners.

References

Chi, M., Feltovich, P., and Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5:121–152.

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, 6:271–315.

Cooper, G., and Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*, 79(4):347–362.

Cummins, D. (1992). Role of analogical reasoning in the induction of problem categories. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18(5):1103–1124.

Ericsson, K. A., Krampe, R., and Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100:363–406.

Hayes, J. R. (1989). *The Complete Problem Solver*. Hillsdale, NJ: LEA.

Kaczmarczyk, L. C., and Miikkulainen, R. (2004). The acquisition of intellectual expertise: A computational model. In *Proceedings of the 26th Annual Conference of the Cognitive Science Society*.

Kvale, S. (1996). *Interviews: An Introduction to Qualitative Research Inquiry*. CA: Sage.

Lang, S. (1986). *A First Course in Calculus*. Berlin: Springer. Fifth edition.

Owen, E., and Sweller, J. (1989). Should problem solving be used as a learning device in mathematics? *JRME*, 20(3):322–328.

Posner, G. J., and Strike, K. A. (1976). A categorization scheme for principles of sequencing content. *Review of Educational Research*, 46(4):665–690.

Quilici, J. L., and Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology*, 88(1):144–161.

Robinson, C. S., and Hayes, J. R. (1978). Making inferences about relevance in understanding problems. In Revlin, R., and Mayer, R. E., editors, *Human Reasoning*. Washington, DC: V.H. Winston and Sons.

Schoenfeld, A. H., and Herrmann, D. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, Cognition*, 8:484–494.

Stewart, J. (1995). *Calculus*. Brooks Cole. Third edition.

Tennyson, R. D., and Nielsen, M. (1998). Complexity theory: Inclusion of the affective domain in an interactive learning model for instructional design. *Educational Technology*, November-December:7–12.