An Investigation of Learner Differences in an ITS Environment: There's No Such Thing as a Free Lunch

Valerie J. Shute, Dan J. Woltz & J. Wesley Regian
Air Force Human Resources Laboratory
Brook A.F. Base TX 78235

Abstract

Research discussed in this paper concerns individual differences in learning behaviors and learning outcomes within an intelligent tutoring system (ITS) that teaches Pascal programming. We investigated relationships among a variety of learning process and outcome measures including: volitional behaviors (e.g., asking for hints from the system), learning process errors (e.g., errors of logic, sequencing errors), learning time, and final learning outcome. We also investigated the relationships of basic cognitive process measures (e.g., information processing speed, working memory capacity) and general knowledge to these learning measures. We currently have data collected on over 200 subjects. Preliminary analyses indicate that a high propensity for hint-asking (i.e., a passive strategy in which the learner relies on the tutor for guidance rather than working out solutions) results in significantly more learning time and significantly less knowledge and skills acquired. However, these relationships changed over time in that initially asking for hints may have been helpful, but persistent use of this behavior became more adversely related to learning. The propensity for hint-asking was partly explained by differences in working memory capacity and general knowledge, suggesting that the choice of a passive learning strategy may in part result from inadequate knowledge and processing ability. These findings have implications for any computer-based instructional systems that have help options available for learners' use and abuse. In particular, ITSs must be able to offer help to learners, but also be able to respond differently to those learners who think the help represents the proverbial free lunch. This is but one example of how more intelligence is needed in ITSs to adapt instruction to individual learner styles and capabilities.

Background

The Learning Abilities Measurement Program (LAMP), at the Air Force Human Resources Laboratory, is conducting basic research on individual differences in knowledge and skill acquisition. The long-term objectives of this work are to apply cognitive science theories and methods in order to enhance (1) the predictive power of Air Force personnel selection procedures, and (2) the effectiveness of instructional systems and approaches. Two laboratories are employed to achieve these objectives. First, the Cognitive Abilities Measurement (CAM) laboratory consists of 40 Zenith 248 microcomputers which are used to conduct research on individual differences in basic cognitive processes (e.g., information processing speed, working memory capacity). Second, the Complex Learning Assessment (CLASS) laboratory consists of 30 Xerox 1186 Lisp computers which are used to investigate higher level learning as occurs in rich, real-world types of environments. CLASS uses intelligent tutoring systems (ITSs) as controlled instructional environments to obtain learning criteria measures, including behavioral measures during learning and outcome measures after learning.

Theoretical Framework

A theoretical framework used to guide research in LAMP is shown in Figure 1 (for more detail see Kyllonen & Christal, in press). The framework asserts that skill acquisition can be seen as a progression from *declarative* (factual) knowledge, to *procedural* (action-centered) skill, to *automatic* (cognitively automatized) skill.

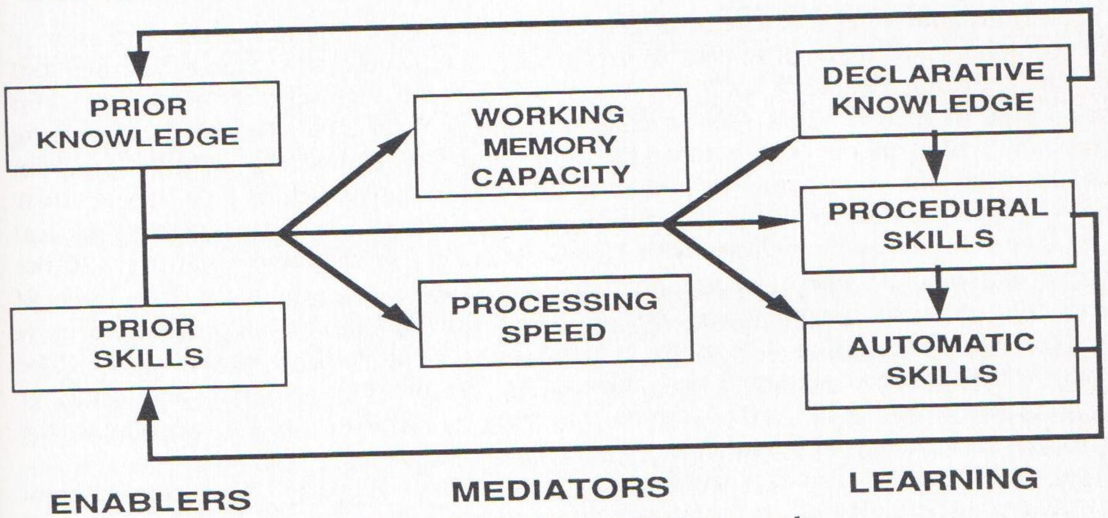

Figure 1: LAMP Research Framework

The first stage of skill acquisition is the learning of the basic facts required to support the skill. For example, in the case of learning to drive an automobile, one must learn about the locations and functions of the gas pedal, steering wheel, speedometer, and so on. The second stage of skill acquisition is the learning of procedures that must be applied to perform the skill. For example, in order to start an automobile one must depress the clutch, activate the ignition, place the transmission in first gear, and so on. Procedures are essentially action recipes for performing the skill. The third stage of skill acquisition occurs when the procedures are cognitively 'automatized' to the point where task performance is smooth, fluid, and effortless. Automatization of the task has important benefits. For example, automatized task performance allows the individual to perform other functions at the same time (such as driving and route planning) and renders task performance highly reliable under stress (Schneider and Shiffrin, 1977) and highly resistant to skill degradation (Regian and Schneider, 1986).

According to the LAMP framework, learning efficiency (individual differences in the rate and quality of skill acquisition) is mediated (limited) on the one hand by hard-wired cognitive constraints such as working memory capacity, processing speed, and proceduralization rate. On the other hand, learning efficiency is enabled by various historical characteristics of the new student, such as prior knowledge and skills. For purposes of this paper, the key point about the LAMP framework is that individuals differ along many dimensions that affect their learning efficiency (e.g., knowledge, motivation, working memory).

Main Research Question

While conventional instructional methods typically ignore learner differences such as those described above, intelligent tutoring systems (ITS) should be able to accommodate to them and optimize the instructional environment accordingly. This paper presents some preliminary findings from a study investigating individual differences in learning Pascal programming from an ITS. We are investigating how learner characteristics relate to specific learning behaviors (e.g., actions, strategies and misconceptions) that have been hypothesized to relate to both learning efficiency and learning outcome measures.

Pascal Programming ITS. The intelligent tutoring system employed in this research as our complex learning environment was originally developed by Jeff Bonar and his staff at the Learning Research and Development Center (University of Pittsburgh) and elaborated at AFHRL. This ITS consists of a curriculum of 25 Pascal programming problems which increase in complexity from simple problems (e.g., write a program which prints out your name) to problems involving complex 'while' and 'repeat until' loops. Each problem requires a three phase solution which begins with a natural language solution and concludes with Pascal code. Help is available at all times in the form of nested hints where, for any problem area in any phase, there are three levels of hints that increase in explicitness. Besides specifically asking for hints, a subject may receive hints, unrequested, if he or she is being unsuccessful in the problem solution. The hints are all context-sensitive and relevant to the current problem area. Another feedback feature of the ITS is that created 'programs' may be run, and the execution of the program can be observed in real time.

Learning Behaviors. The way in which subjects interact with the ITS reflects different learning styles, strategies and misconceptions. Some of the highest level behaviors or 'performance indicators' extracted from the ITS (per problem, per phase) include: Number of times hints were requested, number of times the solution was run (separate indices for total runs, successful runs, and crashes), number of unrequested hints the subject received from the system, number of times the problem statement was reviewed, number of errors of omission (i.e., leaving out a phrase, plan, or line of code in the solution), number or sequencing errors (i.e., when a phrase, plan or line of code exists, but is misplaced), number of errors of logic (e.g., the stopping condition of a loop is wrong or the counter plan increments from the wrong value), number of errors involving operators (e.g., using a + instead of a *) and number of 'data typing' errors (e.g., input a real number instead of an integer).

Learning Efficiency. Since each of the 25 problems in the curriculum must be solved correctly, sequentially, in each of the three learning phases, our learning efficiency measure is defined as the total time spent on the ITS (i.e., the sum of individual solution times per problem). We have currently analyzed data from over 200 subjects who have spent up to 30 hours interacting with the ITS, and total learning time ranges from 3.1 hours to 26.4 hours with a mean of 12.3 and standard deviation of 5.2 hours. This suggests large individual differences in learning efficiency.

Learning Outcome. To test the breadth and depth of knowledge and skills acquired from the learning environment, we developed an online criterion test battery. After completing all 25 problems in the ITS, subjects take three tests of increasing difficulty. The tests are of a procedural format where subjects are actually engaged in some activity (e.g., writing Pascal code) instead of simply reiterating some information as with standard declarative knowledge tests. There are 12 problems per test and the problems are all isomorphic to ITS problems, similarly graded in complexity although randomized in presentation. In the preliminary analyses to be discussed in this paper, we will be using a collapsed accuracy score as our learning outcome measure: all three tests combined.

Research Method

Subjects are tested in groups of approximately 20. Each group spends 40 hours testing and learning over a seven day period. Prior to the ITS instruction, all subjects are administered cognitive process tests measuring working memory capacity and information processing speed, as well as measures of general knowledge. Subjects are from local (San Antonio, Texas) vocational colleges and are prescreened as to the following requirements: no formal Pascal programming training or experience, high school graduation (or equivalent), English literacy, and age between 17 and 30. Since this study is currently in progress, our total sample size will be 270, but data reported in this paper are from a sample of 215 subjects.

Results (Preliminary Analyses)

The learning behaviors under investigation can be seen in Figure 2 and fall into two categories: (1) those under one's volition denoting different learner styles and strategies (e.g., running a solution, asking for hints), and (2) those representing different error types or misconceptions (e.g., errors of omission, sequencing errors). Of the volitional behaviors, the only one to be examined in this paper given the page limitations (although there are data on other behaviors as well) is how frequently a subject requests hints from the system in order to solve a given problem. While this may be a necessary strategy for all learners initially, the continued use of this can be seen as a rather indolent approach to learning; that is, letting the system lead the learner by the hand toward the correct solution.

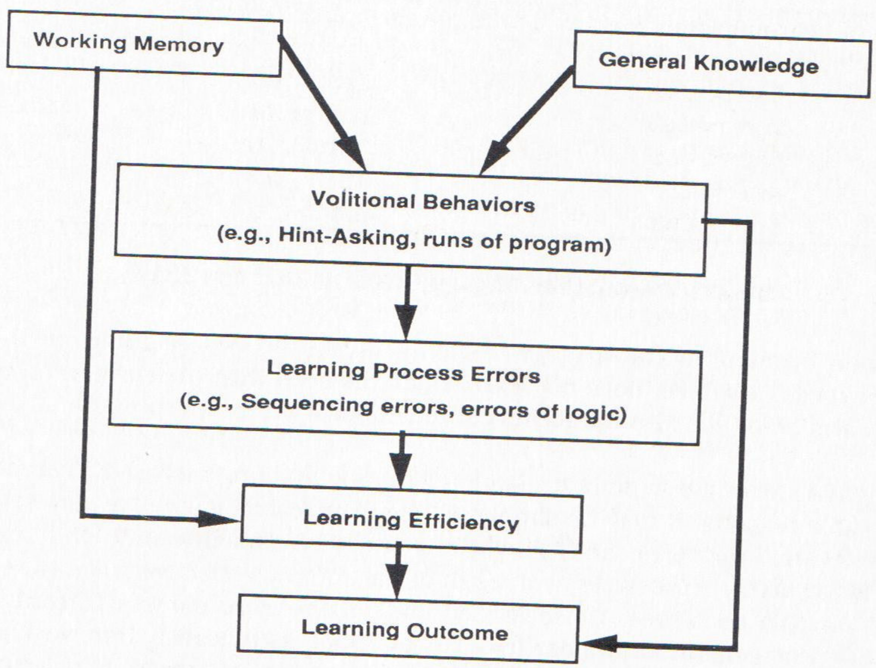

Figure 2: Learning Behaviors

Some research issues involving this hint asking behavior are: a) whether it results in more or less efficient learning (i.e., it might reduce or increase the total time needed for correct solution to each problem); b) whether it affects the quality and/or quantity of knowledge and skills acquired by the end of the ITS; and c) how it relates to different error types within the tutor problems.

The answer to the first question is that continually asking for hints seems to be a very inefficient learning behavior. The simple correlation between hints and Learning Efficiency was: $r = .74$, $p < .001$. Thus, the more hints a person requested, the longer the time spent learning from the ITS, hence, less efficient learning. Furthermore, asking more hints of the system and spending less time figuring out solutions on one's own had negative ramifications as far as ultimate knowledge and skills acquired from the ITS. The correlation between hints and Learning Outcome reflects this adverse effect: $r = -.59$, $p < .001$ where more hints were associated with less knowledge and skills acquired. Finally, the following correlations between asking for hints and frequency of different error types confirmed the inefficiency of this learning strategy: Errors of logic ($r = .95$); Errors of omission ($r = .93$); Errors of sequencing ($r = .83$); Errors involving operators ($r = .70$) and Errors of data type ($r = .62$). The high correlations between hints and error types are not an artifact of a few outliers in the population; rather, scatterplots show clean linear relationships between the number of hints asked for and the different error types. It should also be noted that the first three error types (i.e., errors of logic, omission, and sequencing) represent more conceptual types of misunderstanding while the remaining two error types (i.e., errors of operators and data type) are at a more syntactic level. So, asking for many hints has a strong negative correlation to the more conceptual errors, and lesser negative correlations to the more syntactic errors.

We also looked at the pattern of hint asking behavior over time with the hypothesis that in the latter stages of the ITS, this behavior would be most deleterious both in terms of learning outcome and learning efficiency. As expected, when the 25 problems were divided into fifths, the correlations between hint asking and the two learning measures became more pronounced (see Table 1).

Learning Outcome	Learning Efficiency
1st $r = -.39$	1st $r = .60$
2nd $r = -.43$	2nd $r = .61$
3rd $r = -.51$	3rd $r = .70$
4th $r = -.60$	4th $r = .72$
5th $r = -.63$	5th $r = .72$

Table 1: Correlations Between Hint Asking and Learning

In summary, by adopting the strategy of simply having the system guide one through the solution process, a subject does not learn much (as seen with the correlation to learning outcome), and he or she also manifests certain conceptual errors throughout the learning process.

Why might some subjects embrace such a lethargic learning strategy? There are at least three possible reasons. It may be simply an acquired learning style, it may result from a lack of relevant knowledge and/or working memory capacity such that more active learning is too taxing, or it could be a result of the subject's affective state. In terms of our cognitive process measures, those individuals with a more passive learning style had significantly less general knowledge ($r = -.51$, $p < .001$), significantly less working memory capacity ($r = -.50$, $p < .001$), and were significantly slower on information processing speed tests ($r = .31$, $p < .001$). These process indices were obtained from subjects' performance on a battery of computerized tests measuring various abilities and skills. We also have measures of positive and negative affect from each subject. On each of the seven testing days, subjects completed a Mood Survey (PANAS, Watson, Clark & Tellegen, 1988) both in the morning, before the day's events, and in the afternoon, subsequent to the testing and learning. Two independent dimensions can be extracted from the subjects' responses to the 20 items in the survey: Positive Affect (PA) and Negative Affect (NA). Briefly, PA corresponds to a subject's self perception of enthusiasm and alertness versus lethargy and disinterest. The orthogonal NA dimension reflects a subject's feeling of distress and unpleasurable engagement, as opposed to the attributes of calmness and stability.

The relationship between general mood dimensions and number of hints a person requested from the tutor showed that average PA (collapsed across days and sessions) was not significantly correlated with the number of hints asked for ($r = .07$), but NA was ($r = .30$, $p < .001$): subjects with high negative affect requested more hints from the system. The reason for this finding may be due, in part, to the high NA subjects having low self-confidence, translating into a more passive learning style.

Discussion

One of many possible volitional learning behaviors has been explored in relation to learning efficiency and learning outcome from an ITS. The preliminary findings regarding this learning behavior suggest a number of modifications that could be implemented to make our Pascal ITS more "intelligent" and thus, more effective (Note: these findings may well be generalized to other ITSs that have help options available to the user). First, it was demonstrated how asking for numerous hints is an inadequate strategy in this environment, particularly in the later problems after subjects have become familiar with the ITS. Consequently, the tutor should emphasize or engender more independence in problem solution for learners likely to employ that particular strategy. This requires the identification of passive learners and the subsequent fostering of their independence. One way this could be accomplished is by attenuating the available assistance over time, or at least associating some cost with using hints too frequently. Statistics could be maintained on the frequency of hint usage (all three levels) and a criteria established beyond which the students had to solve problems themselves.

Another finding is that working memory capacity, existing knowledge, and to some extent, information processing speed may play an important role in the choice of learning strategy. Although these cognitive constructs are not easily modified within the human system, the ITS could use the information about an individual's working memory and knowledge to provide more appropriate instruction. For example, low working memory capacity could be compensated for by externalizing and extending it as part of the interface.

Building such intelligence into a tutor would represent a critical test of the Aptitude-Treatment Interaction model of education. That is, intelligent tutoring systems provide a perfect environment for evaluating the potential benefits of different instructional treatments for subjects who differ on certain learning aptitudes. For more passive learners with fewer incoming skills, an initial support or scaffolding may be required, then progressively less assistance if those individuals are ultimately going to learn. Conversely, more active or exploratory learners with more initial skills should have the option of proceeding at their own pace with trial and error learning available whenever possible.

Finally, analyses such as described in this paper and other analyses not reported (i.e., structural equations modeling, factor analyses, etc.) are important for both basic research on individual differences in acquiring knowledge and skills from complex learning environments, as well as for suggesting modifications to intelligent tutoring systems to make them more like the individualized teaching systems they have the potential to be.

References

Kyllonen, P.C., & Christal, R.E. (in press). Cognitive modeling of learning abilities: A status report of LAMP. In R. Dillon &J.W. Pellegrino (Eds.), *Testing: Theoretical and applied issues.* San Francisco: Freeman.

Regian, J.W., & Schneider, W. (1986). Assessment procedures for predicting and optimizing skill acquisition. Paper presented at ETS conference on diagnostic monitoring, July.

Schneider, W., & Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1-66.

Watson, D., Clark, L.A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, (6), 1063 - 1070.