THE JOURNAL OF THE LEARNING SCIENCES, 5(4), 329–355

Individual Differences in Patterns of Spontaneous Online Tool Use

Valerie J. Shute
Armstrong Laboratory
Brooks Air Force Base, Texas

Kevin A. Gluck
Galaxy Scientific Corporation
Lackland Air Force Base, Texas

More than 400 individuals participated in an experiment involving two versions of a computer-based tutor teaching principles of electricity. We examined the relations among elective tool use, learning environment, outcome, and efficiency. We also tested the influence of both individual differences and learning environment on tool-usage behavior. The data showed no differences between the two learning environments (rule application vs. rule induction) with regard to outcome performance or learning efficiency. In addition, neither environment significantly influenced overall tool use. There was a main effect of tool use on learning outcome, but not on learning time. We categorized learners into four groups, based on tool-usage patterns and found that (a) people tended to show stable patterns across time and (b) that patterns differed significantly in terms of learning outcome—it was most effective to use the online tools earlier in the learning process rather than later. In terms of individual differences, we identified the characteristics of learners who evidenced different tool-usage patterns. They varied according to cognitive ability, domain-related interest, and gender. We propose a causal model that takes into account all of these data sources in predicting posttest performance. The article concludes with implications of these findings for those interested in maximizing instructional effectiveness, as well as suggestions for future research directions.

Requests for reprints should be sent to Valerie J. Shute, Armstrong Laboratory AL/HRTI, 1880 Carswell Avenue, Lackland Air Force Base, TX 78236–5507. E-mail: vshute@colab.brooks.af.mil

Behaviors and attitudes manifested during the learning process can influence the quality and degree of learning that takes place, beyond that attributable to cognitive ability (e.g., Ackerman & Kyllonen, 1991; Shute & Glaser, 1990). For instance, interest in a topic is positively related to motivation, which in turn influences participation in the learning process and, consequently, learning outcome (Kanfer, 1989; Shute, 1994; Tobias, 1994). Thus, active participation appears to be an important ingredient in the successful acquisition of new knowledge and skills. But how does that relate to specific learning behaviors?

It is our goal in this article to examine one particular learning behavior—the voluntary usage of online tools by students learning from the Ohm tutor, which instructs students on the principles of electricity (see Shute, 1993). This investigation entailed the identification of tool-usage behaviors and patterns of these behaviors that are more and less likely to result in effective learning of the material. We present evidence that particular learner characteristics are related to these patterns. It should be possible to use this array of individual differences data to inform real-time instructional modifications, with the goal of increasing student motivation and (most important) learning.

The elective, online tools available in the Ohm tutor included (a) looking up definitions in the online reference dictionary, (b) taking measurements on the online circuit (ammeter and voltmeter readings), and (c) changing component values that reside on the circuit (e.g., increasing the resistance). These options were available to learners at any time, with the exception of instances in which such information would provide the answer to a problem on which the person was working. For instance, if a learner was working on a voltage problem, the tutor would allow an ammeter reading, but not a voltmeter reading. After the problem, the student could measure both. The same restriction was in place with regard to changing component values, so the system did not give the answer away. It was possible to look up definitions at any time during instruction or problem solving.

In addition to analyzing the effects of tool usage on learning outcome and efficiency, we also wanted to test the influence of a feedback manipulation on these same dependent measures. Thus, we created two learning environments: rule application (RA) and rule induction (RI). These two environments, which varied only in their feedback, instructed the same curriculum consisting of 20 electricity principles. These principles are listed in the Appendix. Each principle had to be learned before the individual could proceed to the next one, and *learned* was defined in the program as correctly solving three consecutive problems. The focus of these principles involved the interrelations among current, resistance, and voltage, and how they do or do not differ depending on the specific location of a measurement along series and parallel circuits. Figure 1 shows the presentation of an early principle, just as it appeared to learners as they interacted with the tutor.

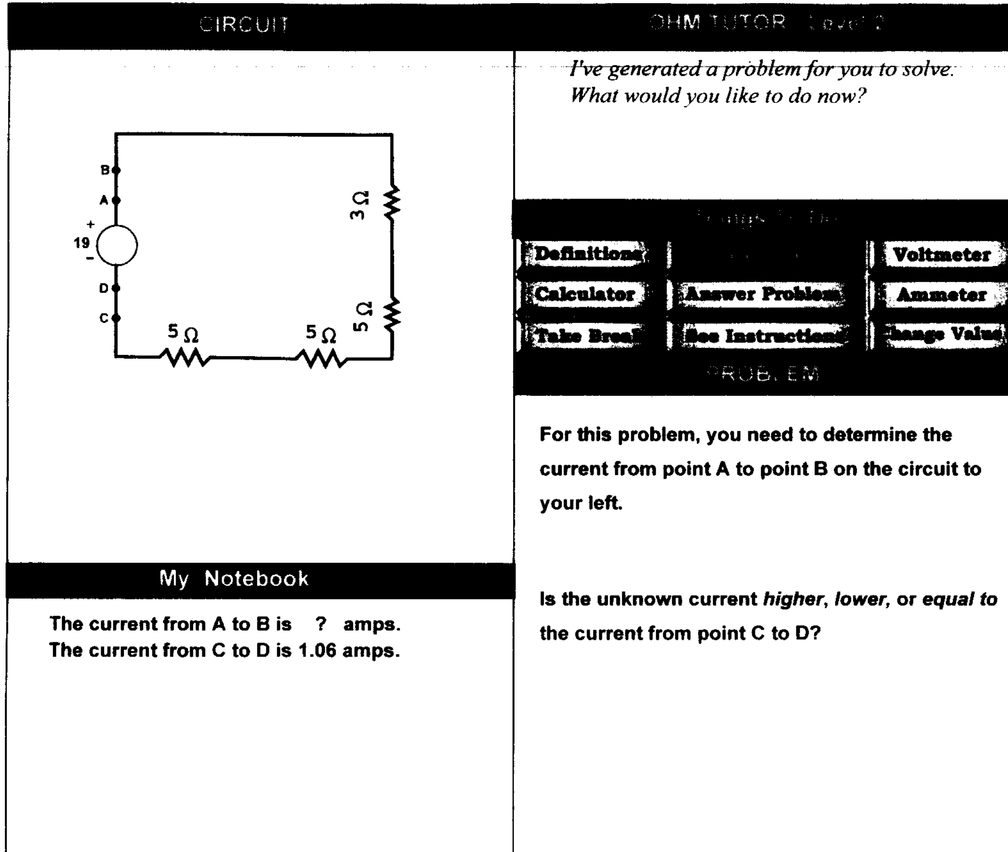

FIGURE 1 Screen display for a problem in the electricity tutor (Ohm).

We tested five main research issues in this study:

1. Is there a main effect of tool usage on learning outcome, efficiency, or both? That is, do individuals with a propensity to actively engage in elective (domain-relevant) activities acquire more knowledge and skill compared to those less inclined? Our hypothesis was that there would be a main effect of tool-usage activity on both measures—a positive relation with learning outcome and a negative relation with learning efficiency. A large body of research supports the importance of learning by doing, in which it is instructionally more effective to let students solve problems on their own, confront and work around obstacles, and then explain to themselves (and sometimes to others) what worked and what did not (e.g., Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Ohlsson & Rees, 1991; Shute & Glaser, 1990; VanLehn, 1990). The other part of this question concerned the cost associated with the application of these tools.
2. What is the relation between learning environment and our dependent measures? We believed that individuals learning from the more cognitively challenging RI environment would end up with higher outcome scores compared to those learning from the RA environment. We made this prediction based on our postulate that the RI condition would foster greater cognitive effort because it provided learners with less information compared to the RA condition. However, this greater outcome was believed to be attained at the expense of learning time.
3. Is there an interaction between tool-usage behavior and learning environment on learning outcome? Given findings from a similar study (Shute, 1993), we hypothesized that learners who were "matched" to a learning environment would perform better than would individuals who were "mismatched." Shute reported that individuals who spent more time using the online tools (i.e., demonstrating a more active, exploratory learning style) learned significantly better from an inductive environment compared to a more applied one. Conversely, learners who used the online tools less frequently learned better from a more applied, straightforward learning environment than from the inductive one.
4. Are these self-initiated activities subject to change (i.e., manipulable) through the interface, or instructional influences? We hypothesized that the inductive environment (RI) would promote greater tool usage compared to the more didactic environment (RA). Learners in the RA environment had no tenable need to engage in elective, exploratory activities. On the other hand, the RI environment required active participation in the learning process because the tutor only provided learners with parts of a principle, and they had to derive the conceptual glue (functional relations) themselves, by any means they could. Thus, we believed that the inductive environment would support (if not actively promote) the use of elective online tools so that learners could obtain information needed to solve the problems.
5. What individual differences variables might be related to tool usage? We hypothesized that aptitude and domain-related interest would show positive rela-

tions with more active behavior (Shute & Glaser, 1990; Tobias, 1994). We further posited that gender may be related to tool usage based on findings that men rate themselves as more active, independent, and interested in math and science, compared to women (Newcombe, 1982). Finally, Schiefele, Krapp, and Winteler (1992) concluded from their meta-analysis that interest accounts for 12% of achievement variance in men and 6% in women. Thus, in addition to cognitive ability and gender, one's interest in the subject matter and possibly even specific pedagogical approaches all play some role in determining the learning outcome.

METHOD

Participants

A total of 431 persons participated in the experiment, but approximately 20 failed to complete the entire study; thus, their data were excluded from analyses. Of the remaining participants, 59% were men, and 41% were women. Participants had a minimum of a high school diploma (or equivalent), and they ranged in age from 17 to 30 years, with a mean age of 22. Participants were obtained from two local temporary employment agencies, and none had any prior electronics instruction or training. All were paid for their participation.

Hardware

All computer-based materials were administered on Compaq 486/33 computers with high resolution SVGA displays on 15 in. monitors, standard keyboards, and Logitech mice.

Software

Pretest and posttest. An online, four-section pretest was administered to assess incoming domain-related declarative knowledge (e.g., characteristics of resistors) and conceptual understanding of Ohm's and Kirchhoff's laws. The posttest began with four sections that were identical in structure to the pretest, but were composed of questions with slightly different content. Three additional sections were included in the posttest. Sections 5 and 6 measured quantitative understanding of the electricity principles. Section 7, the most difficult section, measured the transfer of knowledge to novel situations (e.g., designing circuits to do specific things).[1]

[1] See Shute (1993) for more detailed information on the content of the pre- and posttests.

Learning environments. The curriculum was identical in both versions of the Ohm tutor, and the environments differed only in terms of the feedback, either explicating the specific principle being taught (RA) or requiring its induction (RI). Feedback was provided after problem solution, beginning with "Correct" or "Incorrect" and followed by the principle being instructed. For instance, in the RA environment, feedback clearly stated the variables and their relations for a given problem, communicated in the form of a rule. The feedback that accompanied the problem in Figure 1 was "The principle involved in this kind of problem is that current is the *same* before and after a voltage source." Learners then proceeded to apply the rule in the solution of related problems. In the RI environment, the tutor provided feedback that identified the relevant variables in the problem, but the learner had to induce the relations among those variables. For instance, the following inductive feedback relates to the same principle embodied in Figure 1, "In order to solve this type of problem, you need to understand the relationship between current measured before and after a voltage source." Participants in the RI condition, therefore, generated their own interpretation of the functional relations among variables comprising the different rules.

In both environments, the program was designed such that it monitored every action made by the student. Thus, tool-usage behavior was measured by tracking the amount of time spent in relevant activities that could result in some insight into the principle being instructed at that time. The global tool-usage behavior index was computed as a proportion involving the total time spent engaging the three different online tools divided by the total time spent learning from the tutor. A further distinction was made between *declarative* and *procedural* tool-usage behaviors. Declarative behaviors involved accessing the online dictionary to look up terms and concepts, whereas procedural behaviors related to actively changing component values and metering on parts of the circuit.

Design and Procedure

The 1st day began with the completion of the pretest, followed by an online demographic questionnaire that asked questions about educational and scientific background and interests. Participants then started learning from the tutor.

To familiarize learners with the Ohm interface, everyone started the tutor in the same environment. That is, learners spent about 10 min engaged in directed activities presented in the RA environment. Then, half the participants remained in the RA environment for the duration of the curriculum, and the other half were assigned to the RI environment. Participants spent the rest of Day 1 and all of Day 2 completing the curriculum and the online posttest. They returned on the 3rd day for 8 hr of cognitive abilities testing, using a subset of the Cognitive Abilities Measurement (CAM–4) battery (Kyllonen, in press; Kyllonen et al., 1990).

RESULTS

Preliminary Analyses

Criterion data. All pretest and posttest data are shown in Table 1. In general, scores increased from pretest to posttest across the four isomorphic tests (i.e., Sections 1 to 4), suggesting that learning occurred. A repeated measures multivariate analysis of variance (MANOVA) was computed on the four matched tests (i.e., the within-subjects factor = pretest–posttest gains), and the difference was significant, $F(1, 402) = 421.34, p < .001$.

Data reduction. To simplify the outcome data, we computed a factor analysis (principal axis factoring) on the posttest data and saved the regression scores for each person. This resulted in the extraction of a single posttest factor score (Postfac) with a mean of 0 and a standard deviation of 1 (part 7 of the battery was excluded from analysis because the mean proportion correct data for these items was too low). The percentage of Postfac variance accounted for by these six test data was 71.5%. The respective factor loadings were all high: Test 1, .86; Test 2, .82; Test 3, .80; Test 4, .86; Test 5, .86; and Test 6, .88. This factor score (Postfac) was used as the dependent measure in subsequent analyses.

TABLE 1
Pretest and Posttest Descriptive Statistics: Proportion Correct Data

	M	*SD*	*Minimum*	*Maximum*	*N*
Pretest					
Part 1: T/F	.64	.10	.34	.93	406
Part 2: MC	.49	.17	.07	.93	406
Part 3: Qual/pix	.41	.15	.07	.93	406
Part 4: Qual/no pix	.45	.15	.00	.93	406
Posttest					
Part 1: T/F	.70	.12	.39	.96	406
Part 2: MC	.55	.21	.00	1.00	406
Part 3: Qual/pix	.62	.19	.13	1.00	406
Part 4: Qual/no pix	.59	.21	.13	1.00	406
Part 5: Quan/pix	.48	.25	.00	1.00	404
Part 6: Quan/no pix	.47	.28	.00	1.00	405
Part 7: Transfer	.11	.19	.00	1.00	404

Note. T/F = true/false; MC = multiple choice; Qual/pix = qualitative reasoning where items contained pictures of circuits; Qual/no pix = qualitative reasoning where items contained no pictures and only text; Quan/pix = quantitative items required computation where items had accompanying pictures; Quan/no pix = quantitative items required computation without pictures; and Transfer = difficult problems that required integration of all knowledge and skills, pictures of circuits presented.

We similarly simplified the pretest data using the same procedure, that is, we computed a principal axis factor analysis on the four pretest scores and saved the regression scores, per person. A single pretest factor score was extracted (Prefac) with a mean of 0 and a standard deviation of 1. The variance accounted for by the individual tests was 49%. The factor loadings were as follows: Test 1, .81; Test 2, .75; Test 3, .55; and Test 4, .66.

Incoming differences by learning environment. To determine whether there were any incoming differences among learners in terms of domain-specific knowledge between our two learning environments (RA vs. RI), we computed an analysis of variance (ANOVA) predicting Prefac from our dummy-coded environment variable (0 for RA, 1 for RI). This was not significant, $F(1, 405) = 0.01$. Thus, we did not need to include pretest data as a covariate in subsequent analyses.

Tool Usage and Learning Environment (Global Analyses)

Predictors of learning outcome and efficiency. The four questions addressed in this section map onto the first three research issues presented in the introduction. First, do people who choose to employ the online tools achieve higher learning outcome scores compared to those not using the elective tools? This question relates to the issue of whether active learners, in general, are more successful compared to more passive ones. Second, is there a cost in tutor-completion time for using the tools? Third, are there differences in the dependent measures as a function of the two learning environments? Recall that we originally hypothesized that the RI environment may produce greater outcome scores, but these scores may be at the expense of learning time. Fourth, is there an interaction between tool usage and environment on learning outcome, efficiency, or both? We postulated that persons tending to spontaneously use the online tools would be better suited to the RI environment, whereas those engaging the tools to a lesser degree would be better suited to the supportive RA environment.

We computed a MANOVA on Postfac and tutor time with the overall tool-usage index (divided into high and low categories resulting from a median split on the data) and learning environment as the independent variables. The results showed a main effect of tool usage, $F(2, 373) = 3.42, p < .05$, and results from the univariate tests showed this was due to a significant main effect on learning outcome (Postfac), $F(1, 385) = 6.67, p = .01$ (low-usage $M = -0.11$ and high-usage $M = 0.15$, effect size = .28), but not in relation to tutor time (RA tutor time $M = 3.11$ hr, RI tutor time $M = 3.30$ hr, $F < 1$). In general, outcome scores were enhanced by tool usage, at no cost to learning time. There was no main effect of learning environment on either dependent measure (Fs < 1), nor was there a significant Tool Use × Environment interaction.

However, when the overall tool-use data were decomposed into declarative and procedural tool data and the analyses were recomputed, the story became more interesting. We predicted the same two dependent variables (Postfac and tutor time) across four possible conditions of tool use by environment (i.e., 1 = low-RA, 2 = high-RI, 3 = low-RI, and 4 = high-RA).[2] Results (for both declarative and procedural data) showed a significant effect of condition on Postfac, but not on tutor time. The univariate test results for declarative tool usage were, $F(3, 375) = 3.40, p = .02$; and for procedural tool use, $F(3, 391) = 2.97, p = .03$. Comparing participants who were either matched (Conditions 1 and 2) or mismatched (Conditions 3 and 4) to learning environment, the respective effect sizes were as follows: declarative, .27; and procedural, .15. These data, shown in Figure 2, indicate that for declarative and procedural tool use, learners who readily employed the online tools performed much better on the posttest having learned from the RI condition compared to the RA condition. The converse was true for individuals less inclined to use the tools.

Differences in tool usage by learning environment. We next investigated whether there were any differences in tool usage as a function of learning environment (the fourth research issue mentioned in the introduction). That is, did either environment promote or thwart such behaviors? Our original postulate was that the RI environment would engender tool use compared to the didactic RA environment. We computed an ANOVA on the overall tool-use index by environment, and the difference was not significant, $F(1, 403) = 0.01$. Thus, the environment in which a student learned did not seem to affect tool-usage behavior.

Having conducted these global analyses, we then tested more specific hypotheses related to different patterns of tool use, the relation of these patterns to outcome performance, and learner characteristics related to each pattern.

Patterns of Tool Usage

Tool use across time. As mentioned, there were 20 principles that were instructed by the Ohm program, each more difficult than its predecessor. The principles all related to Kirchhoff's and Ohm's laws (see the Appendix). We divided the curriculum into equal sections, the first 10 principles comprised the *early* segment, and the last 10 principles comprised *later* learning. This division provided a more sensitive way to examine tool-usage patterns compared to the across-the-curriculum scores used in the previous global analyses. Each participant's tool-us-

[2]Explicitly testing the condition variable, rather than analyzing two separate independent variables (environment and tool use), represents a slightly different test than was used before. Specifically, we tested the main effect of condition rather than the two-way interaction.

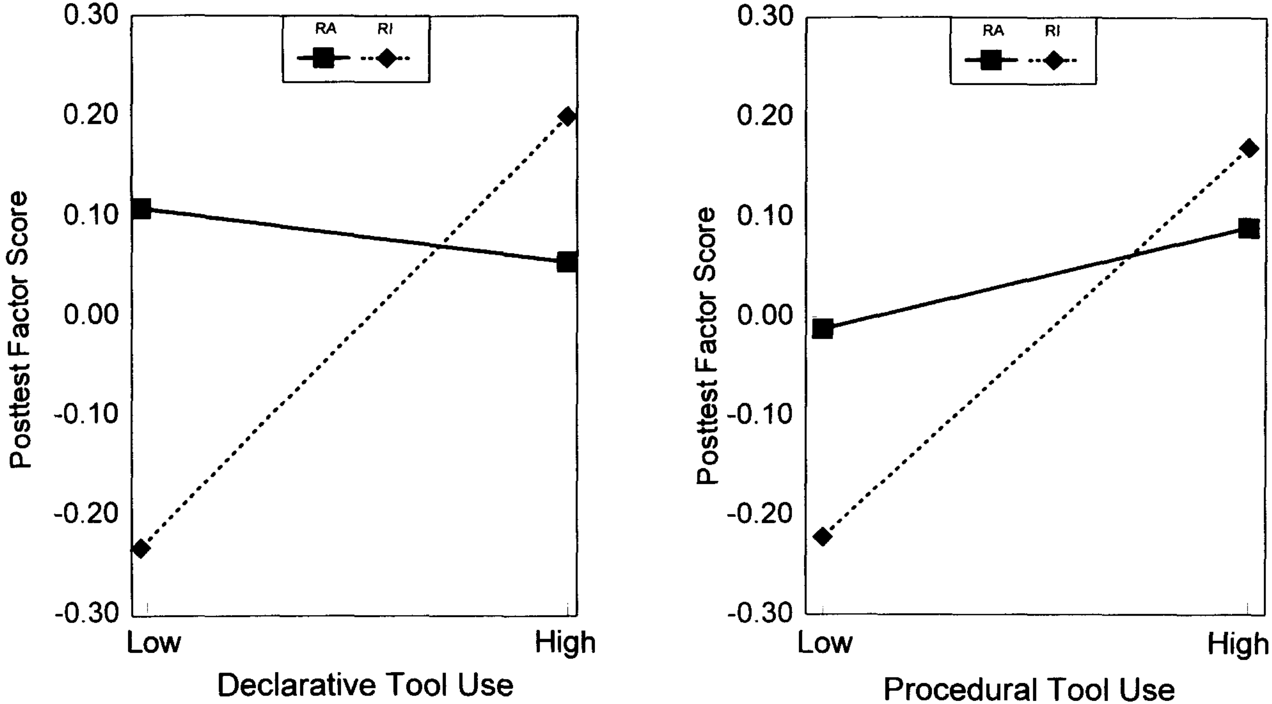

FIGURE 2 Interaction between tool-use data (declatative and procedural) and learning environment on learning outcome.

age data were first separated into declarative and procedural variables, then they were averaged across the early and late segments of the curriculum. This resulted in four new variables: declarative-early, declarative-late, procedural-early, and procedural-late.

To determine the degree to which each variable predicted outcome performance, we computed a stepwise regression analysis predicting Postfac from the four new variables. All four variables remained in the equation (Multiple $R = .34$), $F(4, 381) = 12.10$, $p < .001$. The results were as follows: (a) declarative-early, $t(380) = 3.04$, $p < .005$; (b) declarative-late, $t(380) = -1.97$, $p < .05$; (c) procedural-early, $t(380) = 4.06$, $p < .001$; and (d) procedural-late, $t(380) = -5.29$, $p < .001$. These results suggest that for both declarative and procedural tool usage, it is more beneficial to engage in such behaviors earlier rather than later in the curriculum (note the direction of signs on the individual t tests).

Categorizing tool-usage patterns. To answer questions about individual differences in tool-usage patterns and how those patterns may affect learning, we needed to categorize learners based on their early–late activities. That is, we wanted to be able to distinguish individuals who were fairly consistent in their tool-usage patterns from those who changed their pattern over time. We began by computing a median split on the four variables listed earlier, resulting in divisions of high and low data. For both the declarative and procedural data, this resulted in four new categories of tool-usage patterns: (a) High-early, Low-late (High–Low); (b) Low-early, High-late (Low–High); (c) Low-early, Low-late (Low–Low); and (d) High-early, High-late (High–High).

Which of these patterns is the most efficacious with regard to learning? Table 2 shows the frequencies of each of the four patterns of tool-usage behavior, separately by learning environment.

One should note the unequal frequencies among learners in each of the four categories. In the RA environment, on average, 61% of the learners reflect unchanging, or stable, patterns of tool usage (i.e., Low–Low and High–High), and 39% show changing patterns (High–Low or Low–High). Similarly, in the RI environment, 60% of the learners show stable patterns, and 40% show changing patterns. Thus, regardless of the characteristics of the learning environment, individuals tended to demonstrate consistent behaviors across the tutor.

Predicting outcome: Tool-usage pattern and learning environment. Having made the classification into pattern types, we next tested the degree to which these patterns influence learning outcome. In other words, do certain behavioral patterns enhance learning of this subject matter, and is there an interaction between pattern and environment in relation to learning outcome? Two ANOVAs were

TABLE 2
Frequencies of Learners by Tool-Usage Pattern Per Environment

Declarative			*Procedural*			*Average*			
Early	*Late*	*N*	*Early*	*Late*	*N*	*Early*	*Late*	*N*	*%*
				Rule Application					
High	Low	33.0	High	Low	48.0	High	Low	40.5	Mixed
Low	High	42.0	Low	High	30.0	Low	High	36.0	39
Low	Low	54.0	Low	Low	62.0	Low	Low	58.0	Stable
High	High	68.0	High	High	58.0	High	High	63.0	61
Total		197.0			198.0			197.5	
				Rule Induction					
High	Low	50.0	High	Low	26.0	High	Low	38.0	Mixed
Low	High	49.0	Low	High	44.0	Low	High	46.5	40
Low	Low	60.0	Low	Low	68.0	Low	Low	64.0	Stable
High	High	48.0	High	High	75.0	High	High	61.5	60
Total		207.0			213.0			210.0	

computed on the declarative and procedural tool-usage data separately, where the dependent variable was Postfac and the independent variables were pattern (High–Low, Low–High, Low–Low, and High–High) and environment (applied and inductive):

- Declarative data: Results showed a significant main effect of pattern on learning outcome, $F(3, 385) = 4.52$, $p < .01$. There was neither a main effect of environment on outcome ($F = 1$) nor a significant Pattern × Environment interaction ($F < 1$).
- Procedural data: These results were similar to the declarative data. That is, there was a significant main effect of pattern on outcome, $F(3, 403) = 5.10$, $p = .002$, but neither a main effect of environment on outcome ($F < 1$) nor a significant Pattern × Environment interaction ($F < 1$).

These data, representing both declarative and procedural tool usage, clearly indicate that the optimal pattern is High–Low, and the least effective pattern is Low–High. The stable patterns (Low–Low and High–High) show about average outcome performance. These data are shown in Figure 3.

The next series of analyses examined specific learner characteristics believed to relate to differential patterns of tool usage, the fifth research question examined by this study.

Individual Differences and Tool-Usage Pattern

Aptitude, interest, and gender. What are the characteristics of individuals who manifest the different behavioral patterns? Specifically, (a) do high-aptitude individuals employ the online tools differently than do the cognitively challenged? (b) does greater interest in the subject matter promote different kinds of behavioral patterns? and (c) do men show different patterns compared to women?

Aptitude was the single factor extracted from a principal components analysis of all CAM–4 test data (spatial, quantitative, and verbal): working-memory capacity, information-processing speed, general knowledge, inductive reasoning, fact learning, and skill learning. Six variables accounted for 81.1% of the aptitude factor

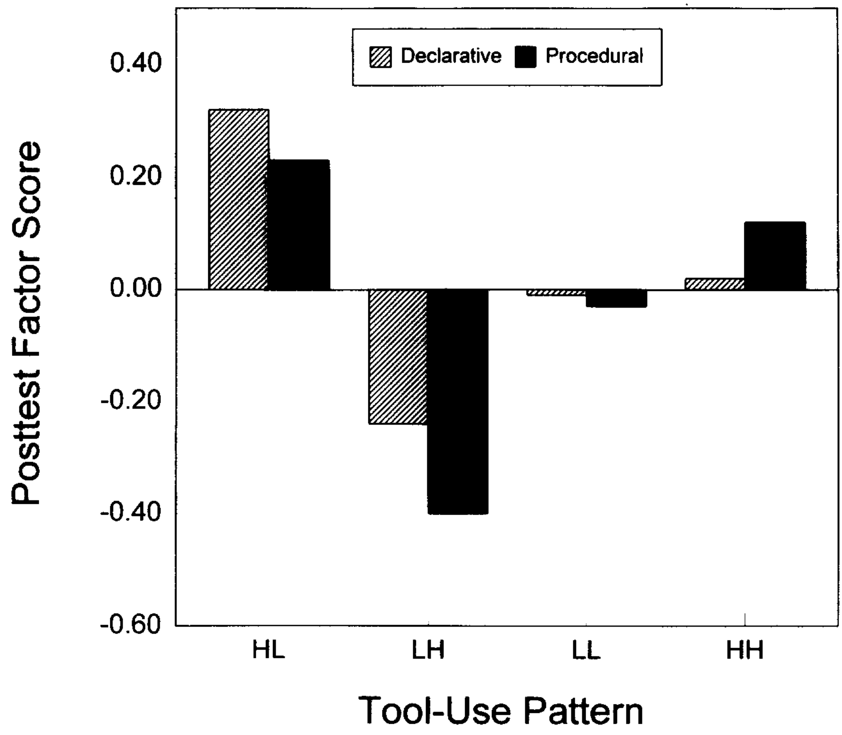

FIGURE 3 Learning outcome by tool-use pattern: declarative and procedural data. HL = high–low; LH = low–high; LL = low–low; HH = high–high.

TABLE 3
Descriptive Statistics From the Demographic Questionnaire

Variable	*M*	*SD*	*Min.*	*Max.*	*N*	*Label*
Interest						
Electronics interest	0.70	0.45	0.00	1.00	408	Interest in electronics? (1 = *yes*, 0 = *no*)
Science interest	0.81	0.39	0.00	1.00	407	Interest in science? (1 = *yes*, 0 = *no*)
Take apart device	0.52	0.50	0.00	1.00	407	Took apart electronic item? (1 = *yes*, 0 = *no*)
Build device	0.10	0.29	0.00	1.00	407	Built an electronic item? (1 = *yes*, 0 = *no*)
Education						
School	12.82	1.40	8.00	18.00	408	Number of years of school
Science	2.61	1.33	1.00	7.00	376	Number of prior science courses taken
Math	3.01	1.48	1.00	12.00	404	Number of prior math courses taken
English	3.19	1.62	1.00	8.00	357	Number of prior English courses taken
General						
Age	22.39	3.48	17.00	30.00	408	Age in years
Gender	0.40	0.49	0.00	1.00	408	0 = *male*, 1 = *female*
Use PC	0.92	0.27	0.00	1.00	408	Ever used a PC? (1 = *yes*, 0 = *no*)
Own PC	0.25	0.43	0.00	1.00	408	Own a PC? (1 = *yes*, 0 = *no*)
Preference	1.54	0.46	1.00	2.00	402	1 = *discovery*, 2 = *guided environment*

Note. Min. = minimum; Max. = maximum; PC = personal computer.

variance, and factor loadings were all high, ranging from .94 to .85. *Interest* similarly resulted from a principal components analysis of items from an online demographic questionnaire (see Table 3, top section). Four variables accounted for 41% of the interest factor variance: interest in electronics (.76), interest in science (.67), took apart electronic item (.65), and built an electronic item (.46). A median split on aptitude and interest resulted in low–high categories for both variables. Gender was a dummy-coded variable: 0 for men, and 1 for women.

We computed two discriminant analyses with the grouping variable as either declarative or procedural tool-usage pattern. The predictor variables, in both cases, included the aptitude and interest scores and the gender variable. Results from the declarative discriminant analysis showed that only aptitude resulted in significant differences by tool-use pattern, $\chi^2(3, N = 332) = 10.68, p < .02$. The procedural data were much more interesting. Each of the three predictor variables showed significant differences in relation to pattern: (a) aptitude, $\chi^2(3, N = 335) = 16.47, p < .001$; (b) interest, $\chi^2(3, N = 335) = 9.55, p < .03$; and (c) gender, $\chi^2(3, N = 335) = 17.53$ $p < .001$. These data are shown in Figure 4.

The aptitude data are similar to the outcome data (from Figure 3) in that the highest and lowest aptitude levels were associated with the High–Low and Low–High categories, respectively. The interest data showed that high interest levels were associated with learners in the High–High category, whereas the other patterns showed only moderate to relatively low interest in the subject matter (as assessed at the outset of instruction). Finally, men and women differed by category type, with more men (i.e., white bars falling below the 0 line) in the High–Low and High–High categories, and women greatly outnumbering men in the Low–Low category.

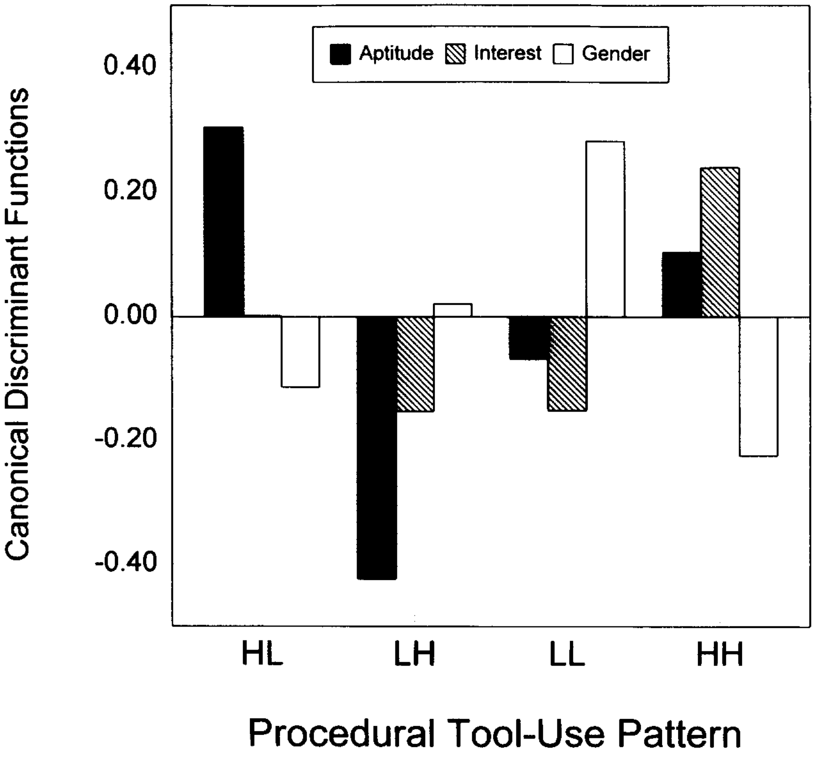

FIGURE 4 Relations among aptitude, interest, and gender by tool-use pattern. HL = high–low; LH = low–high; LL = low–low; HH = high–high.

Relations among aptitude, interest, and gender. Finally, we were curious about collinearity among the independent variables. In other words, what is the correlation among aptitude, interest, and gender? There was no significant correlation between aptitude and interest ($r_{xy} = .05$) or between aptitude and gender ($r_{xy} = -.08$). However, there was a significant correlation between interest and gender ($r_{xy} = -.36, p < .001$). We had no a priori reason to posit any aptitude differences between men and women, and we were gratified to find no correlation between these two variables. But we had believed that men would report themselves to be more interested in the topic compared to women (based on the proportion of men to women enrolled in electronics courses). To test this hypothesis, we computed an ANOVA predicting interest level by gender, and the results were significant, interest $F(1, 274) = 35.13, p < .001$. These data are shown in Figure 5.

The number and variety of research questions in which we were interested during this project resulted in a series of results that may be difficult to follow. It became evident that a parsimonious representation was sorely needed, one that summarized the complex interplay among factors shown to influence elective tool use and subsequent outcome score. We therefore constructed a causal model that combined

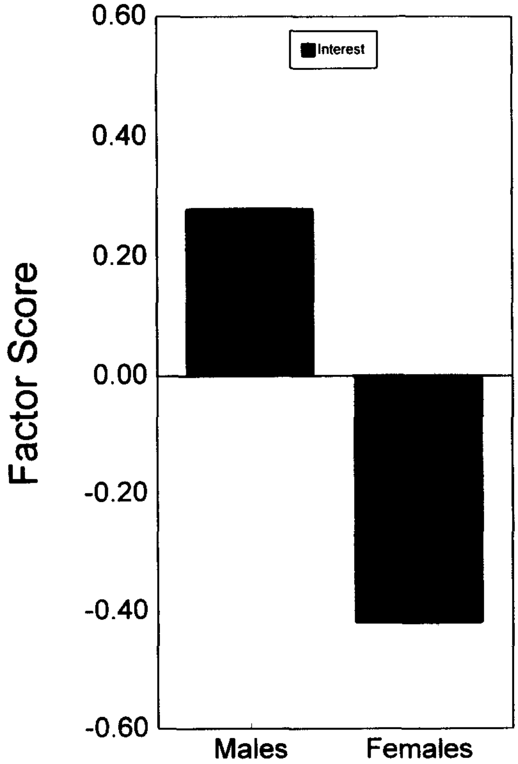

FIGURE 5 Relations among aptitude and interest by gender.

all the independent and dependent variables that were relevant from earlier analyses (i.e., gender, interest, aptitude, procedural tool-usage pattern, and learning outcome). A graphical representation of this model is shown in Figure 6. The numbers along the arrows represent regression coefficients, denoting the strength and directionality of relations and the unique influence of one element on another. These data were tested with a structural equations modeling program, EQS (Bentler, 1989, 1990), and both fit indexes (i.e., Bentler–Bonett Nonnormed Fit Index [NNFI] and the Comparative Fit Index [CFI]) show that the data fit the model very well (CFI = .99; NNFI = .98).

The amount of outcome variance accounted for by this model was $R^2 = .81$, and the amount of procedural (tool-use) factor variance accounted for was $R^2 = .79$. The model suggests that a person's initial interest in the subject matter (electronics) is predicted by gender (i.e., men tend to be more interested than do women). And both aptitude and interest predict an individual's procedural tool-usage pattern evidenced during the learning process. Furthermore, early tool-usage behaviors are positively related to the general tool-use factor, whereas later tool-usage behaviors are inversely related. Finally, the general tool-use factor strongly influences one's learning outcome.

SUMMARY AND CONCLUSIONS

In general, individuals learned a lot from the tutor, as shown by the large pretest to posttest gains. However, we did not find any differences between the inductive and the applied learning environments with regard to outcome performance. Learners in each environment performed comparably on their posttests and required about the same length of time to complete the tutor. This was surprising because we had expected to see longer learning times and greater learning outcomes from those in the more cognitively challenging environment (i.e., RI). In addition, neither environment showed a significant difference in relation to overall tool-use behavior. This, too, was counter to our original belief that the inductive environment would promote more explorations (tool usage) compared to the applied (straightforward) environment. Finally, electing to use the online tools did not add any extra time to the learning process.

Learning Environment, Outcome, and Efficiency

Our finding of no main effect of learning environment on either posttest outcome score or learning efficiency seems inconsistent with reports by others in the field who found such effects with environments that, like ours, differed only in the nature of the feedback that learners received. For instance, Corbett and Anderson (1991) conducted a study comparing outcome and efficiency measures across three feed-

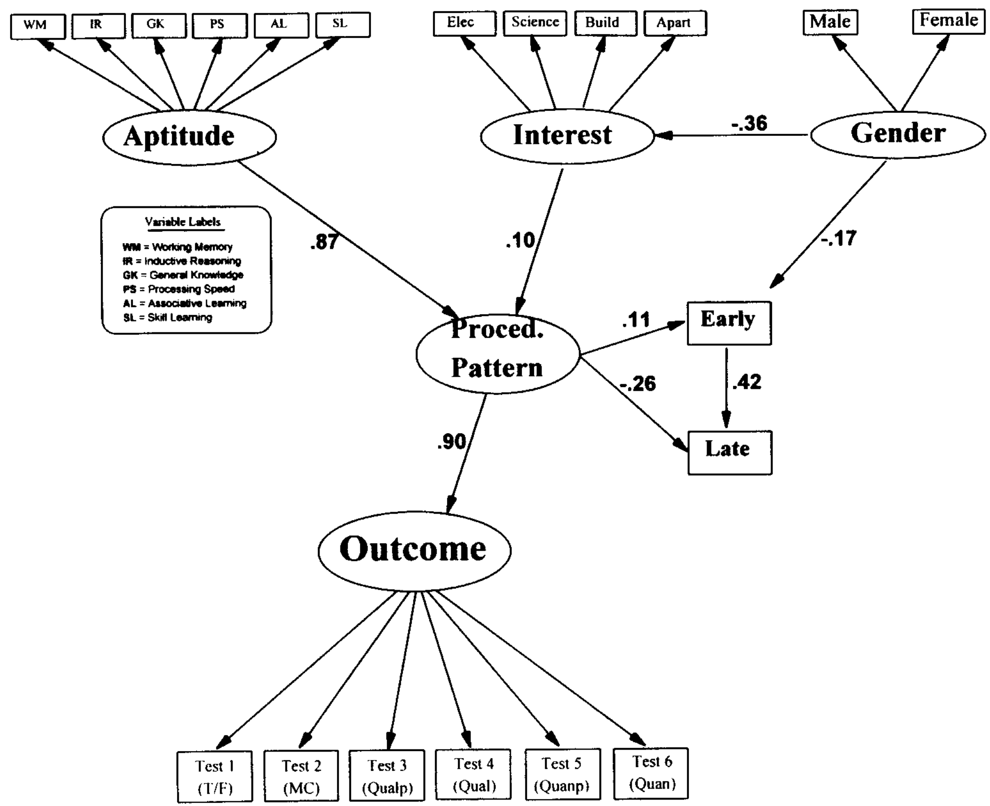

FIGURE 6 EQS solution on model of individual differences in learning. Note that certain representational features of this graph may seem nonintuitive. These are dictated by EQS conventions. For instance, it is standard to represent a factor with arrows pointing away from it to the variables comprising it.

back conditions (immediate, error flagging, and demand feedback) and a control group that received no feedback at all during learning. These were all versions of the Carnegie Mellon University (CMU) Lisp tutor, which varied only in the type (or existence, in the case of the control group) of feedback students received (receiving no feedback proved to be a serious handicap, both in terms of outcome and efficiency). Participants in the feedback conditions achieved virtually identical outcome scores, but they showed substantial efficiency differences, with immediate feedback being the most efficient, followed by error flagging. Demand feedback was the least efficient of the three.

McKendree (1990) investigated the effects of four feedback types on both outcome and efficiency in learning from the CMU Geometry tutor.[3] She found that elaborate, goal-oriented feedback improved both learning efficiency and posttest performance, compared to minimal feedback that merely pointed out a rule violation. She concluded that, in addition to the learning benefits derived from informative feedback, "strategic information in the goal feedback makes it possible to continue successfully" (p. 408), and that is where the efficiency gain is to be found. This particular statement can reasonably explain why we did not see differences in the Ohm tutor's two feedback conditions. A closer look at the conditions in the McKendree study helps clarify this argument. The feedback fell along a continuum ranging from *minimal* (previously described), to *condition violation* feedback (which stated the condition for application of the rule that was violated in the attempted step), to *goal* feedback (which stated the current correct subgoal), and ending at the *maximally informative combined* feedback (a combination of condition violation and goal feedback). The two feedback conditions in the Ohm tutor were most like the goal and combined types from McKendree's study. Individuals in those two conditions were almost indistinguishable on dependent measures, and this would actually be consistent, therefore, with our null difference. Such an interpretation also helps explain why we saw no difference between the induction and application environments in terms of tool usage. It seems reasonable that simply varying the specificity of goal statements would not necessarily influence one's propensity for tool use.

We feel that differences in curriculum objectives could go the rest of the way toward reconciling our results with these others. The Lisp tutor and the Geometry tutor, as products of Anderson's Advanced Computer Tutoring Project (Anderson, Corbett, Koedinger, & Pelletier, 1995) were designed to guide the acquisition of procedural skills. The Ohm tutor, however, was designed primarily to teach declarative knowledge, as is reflected in the curriculum and composition of the pretest and posttest. The empirical question then becomes, What type of feedback is right for what type of knowledge? Here it seems that we discovered a case in

[3]McKendree (1990) actually used a revised portion of the Geometry tutor that focused on proofs from the topic of congruent triangles.

which it simply did not make any difference, perhaps because varying levels of elaboration in goal directedness usually will not affect declarative knowledge acquisition. Only continued empirical research in this direction will demonstrate whether this result is anomalous to the Ohm tutor or characteristic of a more general phenomenon.

Tool-Usage Behaviors and Learning Outcome

In terms of outcome performance, we found that (a) in general, there was a main effect of tool use on learning outcome (i.e., more was better, overall); and (b) specifically, it was much more effective to engage the tools earlier in the curriculum rather than later. This finding resulted in a categorization of learners into four different styles of tool-usage behavior (High–Low, Low–High, Low–Low, and High–High) based on both declarative and procedural tool-use patterns.

These patterns differed significantly in terms of learning outcome. By far, the most successful pattern (High–Low) was characterized by those who started out actively using their online tools to examine their environment, then decreased activities during the second half of the tutor. The lowest outcome scores were associated with the other mixed group (Low–High), who minimally utilized the tools at the outset but later engaged in a flurry of tool-related activities. The two stable groups (Low–Low and High–High) performed about average in terms of posttest scores (with a slight benefit for the consistently High–High group).

Learner Characteristics and Tool-Use Pattern

We believe there are at least three reasons why a learner would engage online tools: (a) hypothesis testing, (b) floundering, and (c) curiosity. That is, a person may use a particular tool after designing an experiment involving the systematic manipulation of a circuit and taking controlled meter readings (i.e., testing a hypothesis). This use of tools may ultimately supplement a student's current understanding and yield other valuable insights. Alternatively, a learner who is unable to solve the current problem may grope for something that he or she can do instead. In this case, tool use would reflect floundering on the learner's part. Finally, a person may simply use the online tools for fun and diversion (i.e., a playful curiosity). Hypothesis testing may be associated with cognitive surfeits, whereas floundering may be associated with cognitive deficits. Playful curiosity may be associated with cognitive ability, but it is probably more associated with an interest in the subject matter (or computers in general). Although specifying the exact underlying reasons for tool usage is not the focus of this article, it is important to note that persons demonstrating the same level of tool usage may have entirely different reasons for doing so.

Figure 4 shows each of the learner variables separated into the four procedural tool-usage patterns, and each pattern differentially impacts learning outcome (Figure 3). We show how it is possible to characterize each pattern by combining these data. To simplify the story, we averaged the declarative and procedural outcome scores shown in Figure 3: High–Low = .28, Low–High = –.32, Low–Low = –.02, and High–High = .07.

Pattern 1: High-early, Low-late. Learners who demonstrate this pattern had the highest aptitude scores of all, but only average interest in the topic. Furthermore, there were more men than women in this category. These learners showed the highest outcome scores (.28) across the four categories of behavioral patterns. Thus, early engagement of the tutor's online features suggests that these persons entered the instructional environment intellectually curious, but only moderately interested in the specific domain. Moreover, they had the mental capacity to adjust quickly to the interface and pedagogical style, thereby reducing the need to employ the tools later on. The decline in tool usage, therefore, was not a function of attenuated interest, but a result of the fact that these learners had the cognitive capacity to acquire the early material quickly and assimilated later material with ease. They did not use the tools later in the tutor because they simply had no need to.

Pattern 2: Low-early, High-late. These individuals shifted from low to high tool usage, a change that could reflect either (a) some floundering in the later stages of the tutor or (b) increasing interest during the course of instruction. Figure 4 shows that participants in this category (about equally split between men and women) exhibited a fairly severe deficiency in aptitude level and a slightly below-average interest in learning about electricity. Given the low aptitude distinction of these individuals, as well as the data that this tool-use pattern resulted in the lowest outcome score (–.32) among the four groups, we concluded that the first explanation (a) is the more likely. These learners arrived at the tutor fairly inactive (Low-early), perhaps as a result of being cognitively challenged by the novel material and too busy attempting to assimilate what information they were able to acquire. In other words, at the outset of instruction, their cognitive resources were fully engaged, and there were no resources available to engage the elective online tools. Later, upon encountering procedural or conceptual impasses, these learners had great difficulty and began using the tools (i.e., floundering), albeit somewhat aimlessly.

Pattern 3: Low-early, Low-late. These individuals represent stable, low tool use learners. Figure 4 shows that these learners displayed slightly lower aptitude levels compared to the other groups, reported a fairly low interest in the subject matter, and consisted of more women than men. These Low–Low learners already had one strike against them upon entering the instructional setting—their disinterest in electricity did not motivate them to actively seek out new information on a regular basis. Additionally, a good portion of whatever minimal tool usage they did engage in was probably not very beneficial because their working memory and associative learning capacities were likely strained in the novel environment. This group also had slightly below-average outcome scores (–.02), which we expected given their composite of attributes. These "minimalist" learners did what it took to get through the program, but nothing more.

Pattern 4: High-early, High-late. Participants in this predominantly male group demonstrated a consistent tendency to explore their environment via the online tools. Without so much as a glance at Figure 4, one could easily predict that these individuals would report a higher interest in the domain than would the members of any other group. Indeed, this was the case. High–High learners also tended to show slightly better-than-average cognitive ability compared to those in other groups. As a result of this combination of characteristics, it was not surprising that this group showed the second highest outcome score (.07) across the four categories, consistent with the finding of a positive main effect of tool-usage behavior on outcome. These stable tool users were presumably seeking out more information from the tutor by taking advantage of the optional activities. These learners were bright enough to appreciate that they were getting something out of the explorations and stuck with it. Judging from their outcome scores, this approach eventually paid a valuable dividend.

Looking back on the results and interpretations from these analyses, we contend that elective tool-usage behavior is more likely a result of cognitive and conative influences on learning than a result of environmental impact. Thus, people tend not to spontaneously change behavior in relation to the two kinds of instructional environments used in this study. Rather, these behaviors are indicative of characteristics of the learner. The graphs in Figure 4 and the model shown in Figure 6 form a composite that represents an intriguing correlational description. It appears that one's pattern of tool-usage behavior is directly related to a combination of general cognitive ability, domain-related interests, and gender, and should be interpreted in relation to those learner characteristics. That is not to say that one learning style is not more advantageous in one environment than it is in the other. For instance, individuals who elected not to employ the online tools at the beginning of instruction, but started using them a lot later on (Low–High), performed poorly on the posttests in both conditions, but even worse if they learned from the inductive environment (Postfac $M = -.49$) than from the applied environment (Postfac $M =$

–.26). Alternatively, individuals who consistently employed the online tools (High–High) performed relatively well on their posttests, but these high tool users in the inductive environment performed even better (Postfac $M = .14$) than did those in the applied environment (Postfac $M = .09$). Our point is merely that the environment does not mediate certain behaviors; characteristics of the learner do. Thus, any given behavior should be interpreted in relation to the characteristics of the student evidencing it and its impact on posttest performance, rather than solely in relation to instructional environment.

Applications of This Research

Findings from this study implicate a variety of methods for enhancing learning outcome and efficiency. Different types of pretraining regimes could be developed and used for certain kinds of learners; and although we cannot readily change a person's aptitude level or gender, it is possible to modify (enhance) one's interest in the subject matter. For example, students who have expressed a low interest in the subject matter (a priori) may greatly benefit by being explicitly informed about how learning the current topic can help them in the real world and why they should be interested in the topic. An increase in interest may alter the student's behavior and subsequently influence learning outcome.

We feel it is important for those involved with the design and development of computer-based learning environments to take note of the consistently positive effect of elective tool use, in general. It seems that every extra chance that learners have to take the initiative to augment their learning during instruction can only have a positive impact. In order for those behaviors to have their effect, it must first, of course, be possible to engage in such activities, and so we encourage the engineers of computerized tutors to include those opportunities in their design whenever possible. These could take any number of forms, such as online dictionaries, formula banks, databases on related topics, or opportunities to go beyond the minimum requirements of the instruction and modify variables within a simulation or microworld. The important thing is simply to have them available.

There is quite a bit of empirical support for the idea of teaching students how to enhance learning in a given instructional environment (e.g., King, 1989; Shute, Lajoie, & Gluck, in press; Swing & Peterson, 1982). Thus, another approach for maximizing the learning experience may be to teach students how to engage in systematic and productive explorations, particularly early in the learning process. For instance, if you had reliable aptitude data suggesting that a student was relatively low in ability, a brief remedial course in effective tool usage and hypothesis testing could be offered. This is less necessary for high-ability learners because they are more likely to participate in such explorations spontaneously.

One more way to increase interest and learning outcome involves the application of aptitude–treatment interaction (ATI) findings. Shute (1993) reported the results from a study employing the same kind of learning environments as discussed in this article (RA and RI). Similar to the findings reported here, results from that study showed that for low-exploratory individuals, learning was enhanced in the more structured RA environment, as compared to the RI environment. For high-exploratory individuals, learning was enhanced in the RI environment, as compared to the RA environment. This study, then, serves as a replication of certain previous results. The lesson is that it may be educationally valuable to take an active ATI approach in assigning students to learning environments. To employ ATI methods in an instructional setting, one needs to make certain critical decisions: What aptitude(s) should be measured before (or during) instruction? Which treatment effects should be manipulated, and what learning outcome and efficiency measures should be used? A taxonomy of learning skills (Kyllonen & Shute, 1989) can assist in rendering principled decisions when answering these questions.

Future Research

The goal underlying this and similar research projects is to find new and creative ways to increase the breadth and depth of knowledge obtained from a computer-based tutor while maximizing learning efficiency. Research on the impact of domain-related interests on learning needs to continue. Instructional designers should pay close attention to the results of that research effort as, in an effort to raise the interest level of those they are trying to educate, they begin to place a stronger emphasis on motivational issues.

Second, applied research should push forward in seeking out innovative ways to increase the flexibility and freedom built into computer tutors so that they encourage the student in the active pursuit of extra knowledge and skill. It is imperative that a learning environment supports the natural intellectual curiosities of any and all students, and we should be looking for ways to facilitate that.

Finally, the learning research community needs to continue to allocate a significant portion of its resources to research on methodologies that would eliminate performance differences resulting from lower initial knowledge and skill. A tutoring system could assess domain-relevant knowledge and ability, then offer the learner an appropriately matched tutorial environment. For some students with lower requisite abilities, the program might offer additional practice opportunities, more elaborative feedback, and so forth to elevate outcome performance and concurrently decrease frustration. Engendering knowledge and skill at this level, prior to more advanced instruction, would almost certainly boost subsequent learning.

ACKNOWLEDGMENTS

Valerie J. Shute is employed by the Intelligent Training Branch, Human Resources Directorate, Brooks Air Force Base, Texas. Kevin A. Gluck is currently a graduate student in the Cognitive Psychology program at Carnegie Mellon University, Pittsburgh, Pennsylvania.

Funds for this research were provided by the U.S. Air Force Office of Scientific Research and the Armstrong Laboratory TRAIN Project, AL/HRTI, Brooks Air Force Base, Texas (J. Wesley Regian, Director). We thank Barry Goettl, Lisa A. Gawlick, Malcolm Ree, and Pat Kyllonen, who provided valuable suggestions about various aspects of this article; Trace Cribbs and April Bremner, who handled all of the programming tasks and associated support; and Linda Robertson-Schüle and Wayne Crone, who provided assistance in data collection.

REFERENCES

Ackerman, P. A., & Kyllonen, P. C. (1991). Trainee characteristics. In J. E. Morrison (Ed.), *Training for performance: Principles of applied human learning* (pp. 193–230). New York: Wiley.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences, 4,* 167–207.

Bentler, P. M. (1989). *EQS structural equations program manual.* Los Angeles: BMDP Statistical Software.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145–182.

Corbett, A. T., & Anderson, J. R. (1991, April). *Feedback control and learning to program with the CMU Lisp Tutor.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Kanfer, R. (1989). Conative processes, dispositions, and behavior: Connecting the dots within and across paradigms. In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology* (pp. 375–388). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

King, A. (1989). Effects of self-questioning training on college students' comprehension of lectures. *Contemporary Educational Psychology, 14,* 366–381.

Kyllonen, P. C. (in press). CAM: A theoretical framework for cognitive abilities measurement. In D. Detterman (Ed.), *Current topics in human intelligence: Vol. 4. Theories of intelligence.* Norwood, NJ: Ablex.

Kyllonen, P. C., & Shute, V. J. (1989). A taxonomy of learning skills. In P. L. Ackerman, R. J. Sternberg, & R. Glaser (Eds.), *Learning and individual differences* (pp. 117–163). New York: Freeman.

Kyllonen, P. C., Woltz, D. J., Christal, R. E., Tirre, W. C., Shute, V. J., & Chaiken, S. (1990). CAM–4: Computerized battery of cognitive ability tests [Unpublished computer software].

McKendree, J. (1990). Effective feedback content for tutoring complex skills. *Human–Computer Interaction, 5,* 381–413.

Newcombe, N. (1982). Sex-related differences in spatial ability. In M. Potegal (Ed.), *Spatial ability: Development and physiological foundations* (pp. 223–250). New York: Academic.

Ohlsson, S., & Rees, E. (1991). The function of conceptual understanding in the learning of arithmetic procedures. *Cognition and Instruction, 8,* 103–179.
Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 183–211). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Shute, V. J. (1993). A comparison of learning environments: All that glitters .... In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 47–74). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Shute, V. J. (1994). Learning processes and learning outcomes. In T. Husen & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (2nd ed., pp. 3315–3325). New York: Pergamon.
Shute, V. J., & Glaser, R. (1990). A large-scale evaluation of an intelligent discovery world: Smithtown. *Interactive Learning Environments, 1,* 51–76.
Shute, V. J., Lajoie, S. P., & Gluck, K. A. (in press). Individualized and group approaches to training. In S. Tobias & D. Fletcher (Eds.), *Training handbook.* NewYork: Macmillan.
Swing, S., & Peterson, P. (1982). The relationship of student ability and small-group interaction to student achievement. *American Educational Research Journal, 19,* 259–274.
Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research, 64,* 37–54.
VanLehn, K. (1990). *Mind bugs: The origins of procedural misconceptions.* Cambridge, MA: MIT Press.

APPENDIX
The 20 Electricity Principles Comprising the Curriculum of the Ohm Tutor

Kirchhoff's Law (Current, Series, Before or After Something)

1. The current at one point in an uninterrupted piece of wire is equal to the current at another point in the uninterrupted piece of wire.
2. The current is the same before and after a voltage source.
3. The current is the same before and after a resistor.

Kirchhoff's Law (Current, Parallel)

4. The current before a resistor is equal to the current after a resistor in a parallel net.
5. The current in the branches of a parallel net sums to the current in the entire net.
6. The current in a component of the net is lower than the current for the entire net.

Kirchhoff's Law (Voltage Drop, Series)

7. The voltage drop across all individual components of a series net sums up to the voltage of the *entire* net.

8. Voltage drop is lower across any single component of a series net than across the whole net.

Kirchhoff's Law (Voltage Drop, Parallel)

9. The voltage drop is the same across components in a parallel net.
10. The voltage drop is the same across any components in a parallel net as across the whole parallel net.

Ohm's Law—Relations Among Voltage, Current, and Resistance

11. Voltage (V) is equal to current (I) multiplied by resistance (R), or $V = I \times R$.
12. When the current (I) goes up or down and the resistance stays the same, this implies that the voltage should also go up or down.
13. Current (I) is equal to voltage (V) divided by resistance (R), or $I = V/R$.
14. If the voltage goes up or down and the resistance stays the same, this implies that the current will go up or down with the voltage.

Kirchhoff's Law (Current, Series, Before or After Something)

15. The current at one point in an uninterrupted piece of wire is equal to the current at another point in the uninterrupted piece of wire.
16. The current is the same before and after a voltage source.
17. The current is the same before and after a resistor.

Kirchhoff's Law (Current, Parallel)

18. The current before a resistor is equal to the current after a resistor in a parallel net.
19. The current in the branches of a parallel net sums to the current in the entire net.
20. The current in a component of the net is lower than the current for the entire net.