Learning Styles Across the Curriculum

A.T. Chamillard
Computer Science Department
University of Colorado at Colorado Springs
Colorado Springs, CO 80933-7150
719-262-3150
chamillard@cs.uccs.edu

Ricky E. Sward
Department of Computer Science
U.S. Air Force Academy, CO 80840
719-333-7664
ricky.sward@usafa.af.mil

ABSTRACT

Recent research has shown that a student's learning style – essentially, the way a student approaches and masters new material – can affect student performance in introductory computer science courses. We show here that a student's learning style can also affect student performance across the courses in the computer science curriculum.

This paper presents the results of a case study in which we collected learning style data for students completing the required courses in a typical computer science curriculum. We then used a wide range of statistical analyses to check for bias in the dataset and to examine the relationships between student learning style and student performance in those courses.

Our analysis identified a number of statistically significant relationships between student learning style and performance. We examine potential explanations for those relationships and discuss ways in which the results can be used to enhance student learning.

Categories and Subject Descriptors

K.3.2 [**Computers and Education**]: Computer and Information Science Education – *computer science education, curriculum.*

General Terms

Measurement, Performance, Experimentation.

Keywords

Learning styles, student performance.

1. INTRODUCTION

A student's learning style indicates how that student responds to a wide range of intellectual and perceptual stimuli and how they prefer to approach new material. For example, some students may prefer to discuss new concepts in small groups, while others may prefer solitary study of those concepts. Some students may learn better by participating in active classroom activities, while others may learn better through reflection on the material.

Research has shown that there is a relationship between learning style and performance in specific subject areas. Specifically, researchers have found that learning style can affect an individual's skill in information processing [8] and student performance in introductory computer science courses [3, 11]. It has also been suggested that student learning style information can be used to help guide instructional delivery approaches and student study habits in introductory computer science courses [3].

Because it has been shown that learning style can have an effect on performance in introductory computer science courses, we hypothesized that learning style may affect student performance across the entire computer science curriculum. Although we believe that the best use of any insights gained about the relationships between student learning style and student performance would be to develop ways in which to reach students of all learning styles, the first step is identifying where performance differences based on learning style are found.

This paper describes the results of a case study in which learning style data was collected for students in four class years completing the computer science curriculum at the U.S. Air Force Academy (USAFA). We performed a wide range of statistical analyses to examine the relationships between student learning style and student performance in the required courses in that curriculum; the analysis results are described below.

The paper makes two contributions to the body of knowledge related to student learning style and computer science performance. The first contribution is the description of a sound statistical analysis process that can be used to explore the relationships between student learning style and student performance in computer science courses. The second contribution is the presentation and discussion of the results obtained by applying the process across the required courses in a typical computer science curriculum. Our analysis yielded numerous statistically significant relationships between student learning style and student performance. While we present and discuss these results and the ways in which they can be used to enhance our teaching, we also recognize that others could certainly achieve other results using different datasets. We therefore focus much of our attention on describing the analysis process itself so that others can also explore these relationships in their own environments.

2. LEARNING STYLES AND PERSONALITY TYPES

Three instruments were used to collect the learning style data included in the analysis reported here: the Felder Index of Learning Styles, the Kolb Learning Styles Inventory II '85, and the Keirsey Temperament Sorter.

Felder's Index of Learning Styles (ILS) measures four different dimensions of an individual's learning style [4]. The four dimensions are active/reflective, sensing/intuitive, visual/verbal, and

ITiCSE'05, June 27-29, 2005, Monte de Caparica, Portugal.

sequential/global. Active learners learn better by doing something active – discussing the material, explaining it to someone, or using it to solve problems. Reflective learners learn better by thinking about the material before trying to explain or use it. Sensing learners like to memorize facts and solve problems using well-established methods, while intuitive learners prefer discovering relationships and using innovative problem-solving approaches. Visual learners retain more from things they see – pictures, diagrams, flow charts, etc. – while verbal learners get more out of words (i.e., written and spoken explanations). Finally, sequential learners gain understanding in linear, logical steps, while global learners tend to learn almost random pieces of material, then suddenly "get it".

Kolb's Learning Styles Inventory II '85 measures an individual's intrinsic learning style or predisposition in any given learning situation [7]. Kolb describes a learning cycle of involvement in concrete experiences (Concrete Experience), followed by observation of and reflection on those experiences (Reflective Observation), followed by integration of those observations into a sound theory (Abstract Conceptualization), followed by use of those theories to make decisions and solve problems (Active Experimentation), leading back to more concrete experiences.

The Keirsey Temperament Sorter is not strictly a learning style instrument; rather, it was designed to identify different personality types [6]. The model used by Keirsey is very similar to Myers-Briggs and other personality models. The four dimensions used by Keirsey are extravert/introvert, intuitor/sensor, thinker/feeler, and judger/perceiver. Extraverts tend to try things out and focus on others, while introverts tend to think things through and focus on ideas. Sensors tend to be practical, detail-oriented, and focus on facts and procedures. Intuitors tend to be imaginative, concept-oriented, and focus on meanings. Thinkers tend to be skeptical and make decisions based on logic and rules, while feelers tend to make decisions based on personal considerations. Judgers tend to set and follow agendas, and seek closure even with incomplete data. Perceivers tend to be more adaptive, and resist closure in the hopes of procuring more data.

Although Keirsey's instrument identifies different personality types rather than explicitly trying to measure learning style, a student's personality type affects the ways in which they learn. For ease of reference, we refer to Felder, Kolb, and Keirsey data as learning style data throughout the rest of the paper.

3. DATASET

In this section we discuss the characteristics of our analysis dataset. Although the students in our dataset took both required and optional courses to complete the Computer Science major, we only included required computer science courses in our analysis. The computer science program at USAFA was CSAB-accredited during the period under analysis (and continues to be an accredited program). Topic coverage closely follows ACM curriculum guidelines [1].

The content and sequence of major's courses at USAFA is continually examined and modified as necessary. While space precludes including a discussion of all the required courses, we note that 12 required computer science courses at the sophomore through senior levels were included in our analysis. Additional course details will be provided as appropriate in the following sections.

In a previous paper, we reported on the use of learning style data in an introductory computer science course [3]. The dataset for the analysis reported here includes those students from the Class of 2001 (the class comprising the majority of the previous dataset) who declared computer science as their major, as well as the computer science majors from the Classes of 2002, 2003, and 2004.

Only students who graduated with a computer science degree are included in our dataset; for the four class years, this yields a total of 80 students. We note that learning style data is available for 53 of those students. The other students either entered USAFA with validation credit for the introductory course in which the learning style data was collected or attended that course in the Spring 2001 semester (when the primary author was on sabbatical and the learning style data was not collected). We recognize that excluding these students from our analysis could introduce some bias into our results, and we test for this bias in Section 4.2.

We also note that a complete dataset would be comprised of 51 total offerings of the 12 courses included in the dataset. We were unable to obtain student performance data for seven of those course offerings for a variety of reasons. We were unable to obtain data for a single offering of five different courses, as well as two offerings from the seven offerings of our CS1 course. Although we would certainly prefer having a complete dataset, the distribution of the missing course data across six different courses still lets us perform valid statistical analysis on all the required courses in the curriculum.

Another potential limitation of our dataset is the possibility that the students who choose to attend USAFA and major in computer science are not a representative sample of computer science students at other universities. We discuss the ways in which we plan to expand our research to address these limitations in our conclusions.

4. ANALYSIS APPROACH AND RESULTS

4.1 Independent and Response Variables

The independent variables for our analysis were the measures of student learning style discussed above. Specifically, the set of independent variables was comprised of: Felder scores for the four dimensions (Active/Reflective, Sensing/Intuitive, Visual/Verbal, and Sequential/Global), Kolb scores for the learning cycle (Concrete Experience, Reflective Observation, Abstract Conceptualization, and Active Experimentation), and Keirsey classifications for the 4 dimensions (Extravert/Introvert, Intuitor/Sensor, Thinker/Feeler, and Judger/Perceiver).

Measurements fall into five major scales: nominal, ordinal, interval, ratio, and absolute [5]. The Felder scores are measured in 12 ordered, even increments ranging from 11A to 11B. Because the increments are ordered and the distances between adjacent scores (i.e., 11A and 9A) are constant, the Felder scores are measured on the interval scale. The Kolb scores are measured as contiguous integers, but because each score ranges from 12 to 48 (rather than starting at 0), the Kolb scores are also measured on the interval scale rather than the ratio scale. The Keirsey classifications for the different dimensions are bivariate and unranked; they are therefore measured on the nominal scale.

Course performance data for the required computer science courses provided the response variables for our analysis. For each course, we included the student percentages on the assessments in the course as well as the student's overall percentage in the course and their grade in the course. We encoded the course grade using standard GPA values for letter grades (A = 4.0, A- = 3.7, B+ = 3.3, etc.). The resulting set of response variables was comprised of 89 variables for the 12 courses.

4.2 Checking Potential Bias

One of the potential dataset limitations discussed above involved our concern that excluding students who entered USAFA with advanced programming skills or who took the course in which learning style data was collected in Spring 2001 could bias our analysis results. To check for this potential bias we partitioned our dataset into students for whom we have learning style data and students for whom we do not have this data. We then compared the means for each of the 89 response variable distributions for these two partitions under the null hypothesis that the distributions could have been drawn from populations with equal means.

One common method for comparing the means of two distributions is the independent samples t test. One of the underlying assumptions of the t test, however, is that the distributions being compared represent samples from normal populations; this assumption was regularly violated in our dataset.

There are other, non-parametric tests that we can apply instead of the t test. For example, the Mann-Whitney test does not assume normality of the distributions being compared [9]. It does, however, assume that the distributions are the same shape (e.g., they have the same variance); inspection of the descriptive statistics for these distributions indicates that the variances are unequal for many of the response variables. Alternatively, we can use the Kolmogorov-Smirnov Z test, which detects differences in both the locations and shapes of two distributions [10].

Because the Kolmogorov-Smirnov Z test is less powerful than the Mann-Whitney test, we conducted both tests for the response variable distributions. In the cases where the Mann-Whitney test identified a statistically significant difference (at the standard p=0.05 level) but the Kolmogorov-Smirnov Z test did not, we checked the homogeneity of variance for the two distributions of that response variable using the Levene test. Where the variances were equal, we accepted the more powerful Mann-Whitney test results; where they were not equal, we accepted the Kolmogorov-Smirnov Z test results.

For the 89 response variables in our analysis, both the Mann-Whitney and Kolmogorov-Smirnov Z tests identified 8 variables for which the distributions were different with statistical significance. In 7 of the 8 cases the students for which learning style data is not available had higher percentages than those for which we have learning style data. For 7 additional response variables the Mann-Whitney test identified a statistically significant difference but the Kolmogorov-Smirnov Z test did not. Results of the Levene test for homogeneity of variance indicate that we cannot reject the null hypothesis that the variances are the same for all 7 of those variables. This does not prove that the variances are the same, of course, but it does show that the equal variance assumption made by the Mann-Whitney test is not violated in a significant way. We therefore accept the Mann-Whitney test results for the additional 7 response variables.

Although our results indicate that there is some evidence that our dataset may be somewhat biased through exclusion of the students for whom learning style data is unavailable, we view the lack of statistically significant results for 74 out of the 89 response variables as sufficient evidence that such bias is not systemic throughout the response variables.

4.3 Comparing Means

We also applied the techniques for comparing distribution means described in the preceding section to analyze the relationships between the response variables and those independent variables that are measured on the nominal scale. Each of the four Keirsey dimensions are bivariate, so we used the same partitioning approach described above to compare the distribution means in each of these dimensions. Our analysis yielded 24 statistically significant results on 21 distinct response variables.

For 9 of these results, both the Mann-Whitney and Kolmogorov-Smirnov Z tests yielded statistically significant results. The Mann-Whitney test yielded statistically significant results for an additional 14 comparisons. Results of the Levene test for homogeneity of variance indicate that we cannot reject the null hypothesis that the variances are the same for 12 of those 14 comparisons, so we accept those as significant results as well. Finally, in one comparison only the Kolmogorov-Smirnov Z test yielded a statistically significant result. We therefore include 22 (9+12+1) statistically significant results in the following discussion. These results are summarized in Table 1.

Table 1. Significant mean differences

Judger/Perceiver Dimension	**11**
Judger Higher	8
Perceiver Higher	3
Intuitor/Sensor Dimension	**7**
Intuitor Higher	0
Sensor Higher	7
Thinker/Feeler Dimension	**3**
Thinker Higher	1
Feeler Higher	2
Extravert/Introvert Dimension	**1**
Extravert Higher	1
Introvert Higher	0
Total	**22**

The Judger/Perceiver dimension yielded 11 of the 22 results; note that this represents statistically significant results in this dimension for over 12% of the response variables included in the analysis. Students who were classified as judgers performed better than students who were classified as perceivers for 8 of the response variables. These results are consistent with those reported in [3], where judgers exhibited better performance than perceivers on the programming assignments, course percentage, and grade (as well as other variables) in an introductory course. It is interesting to note, however, that for 3 of the 4 significant results related to tests and final exams, students who were classified as perceivers performed better than students who were classified as judgers.

This result is particularly interesting given the typical expectation that students who do well in other components of a course are also likely to do well on the formal course examinations. Although many of us have heard the argument that "I know the material but I just don't test well", our results indicate that, in some cases, student learning style may well be driving their performance on tests.

Another explanation, of course, is that the way that the professors structure the formal examinations in these courses may provide some advantage to students classified as perceivers.

The Intuitor/Sensor dimension yielded 7 of the 22 results. In all 7 cases, students who were classified as sensors performed better than students who were classified as intuitors. The majority of these results relate to programming assignments, including the large software engineering projects required in the software engineering capstone courses [2]. Given that sensors tend to be more detail-oriented and focused on procedures than intuitors, these results are consistent with our intuition that sensors would be expected to perform better on such assessments.

We had a very limited number of statistically significant results in a Thinker/Feeler and Extravert/Introvert dimensions. Students who were classified as thinkers performed better than students who were classified as feelers in overall course percentage in our junior-level architecture course, but performed worse than feelers on two of the assessments in our compilers course. Students who were classified as extraverts performed better on the programming assignments in our networks class than students classified as introverts. Given the limited number of significant results from the Thinker/Feeler and Extravert/Introvert dimensions, we do not believe we can draw any general conclusions about a student's classification in these dimensions and their performance in computer science courses.

4.4 Correlations

Unlike the Keirsey classifications, the Felder and Kolb independent variables are measured on the interval scale. To examine the relationship between these independent variables and student performance, we correlated each of these independent variables with each of the response variables. For each such correlation we calculated Pearson's Correlation Coefficient (r) and a measure of statistical significance (p). The coefficient can range from -1.0 to 1.0, with a coefficient magnitude close to 1.0 indicating a strong linear relationship and a magnitude close to 0.0 indicating no linear relationship.

We found 53 statistically significant correlations between the Felder and Kolb variables and the response variables. The full results are too voluminous to present here given space constraints, but are available from the author. We limit our discussion below to summary observations about those results.

The most compelling correlation results were between Kolb's measure of predilection toward Concrete Experience and the response variables. We found 19 such correlations, with magnitudes ranging from 0.298 to 0.737. All of the correlations were negative, indicating that students with a stronger predilection toward Concrete Experience were likely to perform more poorly on a wide variety of assessments in 5 of the 12 courses included in the dataset. While we certainly would not generalize those results to conclude that such students are less suited for computer science, we do believe these results are sufficiently compelling to examine those courses to identify whether or not the course structure and delivery approach are in some way harming those students.

We also found 10 significant correlations between Felder's Sequential/Global dimension and the response variables, ranging in magnitude from 0.315 to 0.622. All of these correlations were also negative, indicating that students who are classified as more sequential than global tend to perform better on more than 11% of the response variables. This result has intuitive appeal, since many of our courses are structured to proceed in the linear, logical steps preferred by sequential learners, but we do note that previous work in this area has not identified significant results for this dimension.

The correlations with the strongest magnitude are between Kolb's measure of predilection toward Reflective Observation and the response variables. Although we only found 6 such correlations, the highest 4 ranged from 0.735 to 0.848 (all of them positive). It is interesting to note that these 4 correlations correspond to the 4 strongest (negative) correlations for Concrete Experience. These 4 response variables are for the programming assignments, tests, overall percentage, and grade in our theory of automata course. Given that context, this result is not particularly surprising, since it would be reasonable to expect theory courses to be more "conceptual" than "concrete."

We found 6 or fewer statistically significant (and relatively low) correlations for Felder's Active/Reflective, Sensing/Intuitive, and Visual/Verbal dimensions and for Kolb's measures of Abstract Conceptualization and Active Experimentation. These correlations therefore to do not seem to provide significant additional insight into the relationships between learning style and student performance.

5. USING THE RESULTS

We are faced with an interesting paradox as we consider the results of our analysis. As researchers, we are interested to find numerous statistically significant results. As teachers, however, we would rather find that our teaching techniques foster an environment in which all student learning styles are addressed so that a student's learning style doesn't have any noticeable effect on their course performance.

We would therefore assert that the analysis results presented above should be used to help improve teaching methodology rather than as an indicator that students with some learning styles are better at particular computer science course activities than students with others. We discuss some suggested ways of doing so below.

For example, in the previous section we noted that students classified as judgers tended to perform better than those classified as perceivers, except on formal examinations. These results indicate that we should examine both the assessments favoring judgers and the assessments favoring perceivers to identify possible ways to modify those assessments to be more balanced. Five of the statistically significant results favoring judgers were related to homework and programming assignments. These assignments could provide a way to iteratively acquire more data to support decisions, a valuable experience for both judgers and perceivers. Similarly, the formal examinations could be evaluated for characteristics that punish judgers for making decisions based on incomplete data; if such characteristics are found, the examinations could be modified to make them more balanced.

As another example, we note that students classified as sensors performed better than students classified as intuitors on a variety of programming assignments, including the large team projects in the software engineering capstone courses. We could therefore consider modifying those activities to strengthen intuitor performance. One possible technique would be to require that individuals or project teams pursue and document various alternate problem solutions, an activity that seems to be more suited to the preferences of intuitors (and that is useful in its own right). Because providing such focused attention based on learning style is time consuming for teachers, it is reasonable to conduct the statistical analyses described above to identify areas of differing performance before pursuing such courses of action.

The approach described here for examining the relationships between learning style and student performance and using the results to modify course delivery and assessment techniques could also be used as part of a continuous improvement program. If, for example, a particular course yielded a large number of statistically significant results, opportunities to more effectively reach students with diverse learning styles could be explored in the context of that course. For example, we found statistically significant (and relatively large) results for five out of seven response variables in the theory of automata course for both Felder's Sequential/Global dimension and Kolb's Concrete Experience and Reflective Observation measures. These results indicate that we should explore additional delivery techniques in that course to try to achieve more balanced student performance that is independent of particular learning styles.

Although we seek to provide balanced delivery and assessment techniques for students of all learning styles, we also recognize that some of the performance differences for students with different learning styles may simply be due to the nature of computer science as a field. While we encourage the use of analysis results to tune our courses appropriately, we also recognize that in some cases students may need to overcome particular learning styles to approach computer science material successfully.

6. CONCLUSIONS AND FUTURE WORK

Although we found a wide variety of statistically significant results in our analysis, we believe it is unreasonable to expect significant results across all course assessments for all dimensions of the various learning style models included in the case study. Such wide-ranging results would be more indicative of an unbalanced learning environment catering to specific learning styles to the exclusion of others rather than providing meaningful insights about learning styles and student performance. As more empirical work is completed in this area, however, we may discover persistent relationships between learning style and student performance on particular computer science activities.

We caution again that there are limitations to the generality of the specific results of our analysis. As for any dataset drawn from a single university, we are concerned that our students may not form a representative sample of computer science students in general. We therefore suggest that others apply similar approaches for their computer science curricula at other schools. Some of the insights to be gained would be course-specific, supporting pedagogical changes to the course as required, while some might also contribute to more general insights about the relationships between learning style and student performance across typical computer science curricula.

We are planning to continue this work to explore these relationships at a different school. We have started collecting learning style and student performance data for students majoring in computer science at the University of Colorado at Colorado Springs (UCCS). UCCS graduates approximately 50 Computer Science majors annually, so the UCCS data will provide a student sample that may be more representative of Computer Science students in general as well as yielding a larger dataset for analysis.

We also note that the dataset analyzed in this paper filtered out those students who began but did not complete the computer science major. We plan to analyze both USAFA and UCCS data to determine whether learning style has a statistically significant effect on successful completion of the major.

7. REFERENCES

[1] ACM/IEEE-CS Joint Curriculum Task Force. *Computing Curricula 1991.* ACM Press, New York, NY, 1991.

[2] Chamillard, A.T. and Braun, K.A. The Software Engineering Capstone: Structure and Tradeoffs. In *Proceedings of the Thirty-Third SIGCSE Technical Symposium on Computer Science Education*, Northern Kentucky, Kentucky, March 2002, pp. 227-231.

[3] Chamillard, A.T. and Karolick, D. Using learning style data in an introductory computer science course. In *Proceedings of the 30th SIGCSE Technical Symposium on Computer Science Education*, New Orleans, Louisiana, March 1999, pp. 291-295.

[4] Felder, R. Reaching the second tier: Learning and teaching styles in college science education. *J. College Science Teaching*, 23(5): 286-290, 1993.

[5] Fenton, N.E. and Pfleeger, S.L.. *Software Metrics: A Rigorous and Practical Approach*. PWS Publishing Co., Boston, MA, 1997.

[6] Keirsey, D. and Bates, M. *Please Understand Me*. Prometheus Nemesis Book Co., Del Mar, CA, 1984.

[7] McBer & Co., *Learning-Style Inventory, 1985: Technical Specifications*, McBer & Co., Boston, MA, 1985.

[8] Moldafsky, N.I. and Kwon, I. Attributes affecting computer-aided decision making – a literature review. *Computers in Human Behavior*, 10(3): 299-323, 1994.

[9] Neter, J., Wasserman, W., and Whitmore, G.A.. *Applied Statistics*, Allyn and Bacon, Inc., Boston, MA, 1978.

[10] SPSS Inc., *SPSS Base 11.0 User's Guide*, SPSS Inc., Chicago, IL, 2001.

[11] Thomas, L., Ratcliffe, M., Woodbury, J., and Jarman, E. Learning styles and performance in an introductory programming sequence. In *Proceedings of the 33rd SIGCSE Technical Symposium on Computer Science Education*, Northern Kentucky, Kentucky, March 2002, pp. 33-37.