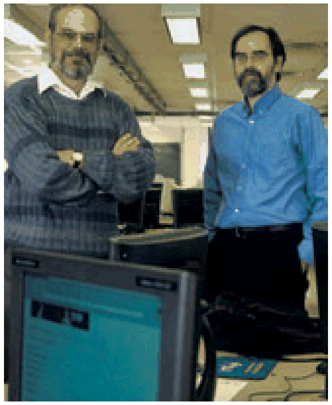
Online Presentations Show Academic Performance Advantages Over Auditorium Lectures

John H. Newman
Department of Psychology
Kutztown University of Pennsylvania

Tomas Lozano-Perez & W. Eric Grimson
Department of Electrical Engineering & Computer Science
Massachusetts Institute of Technology

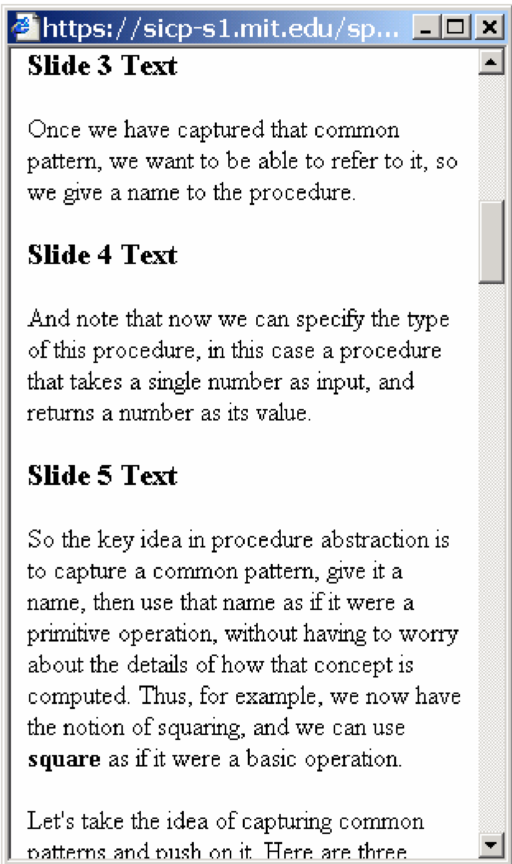


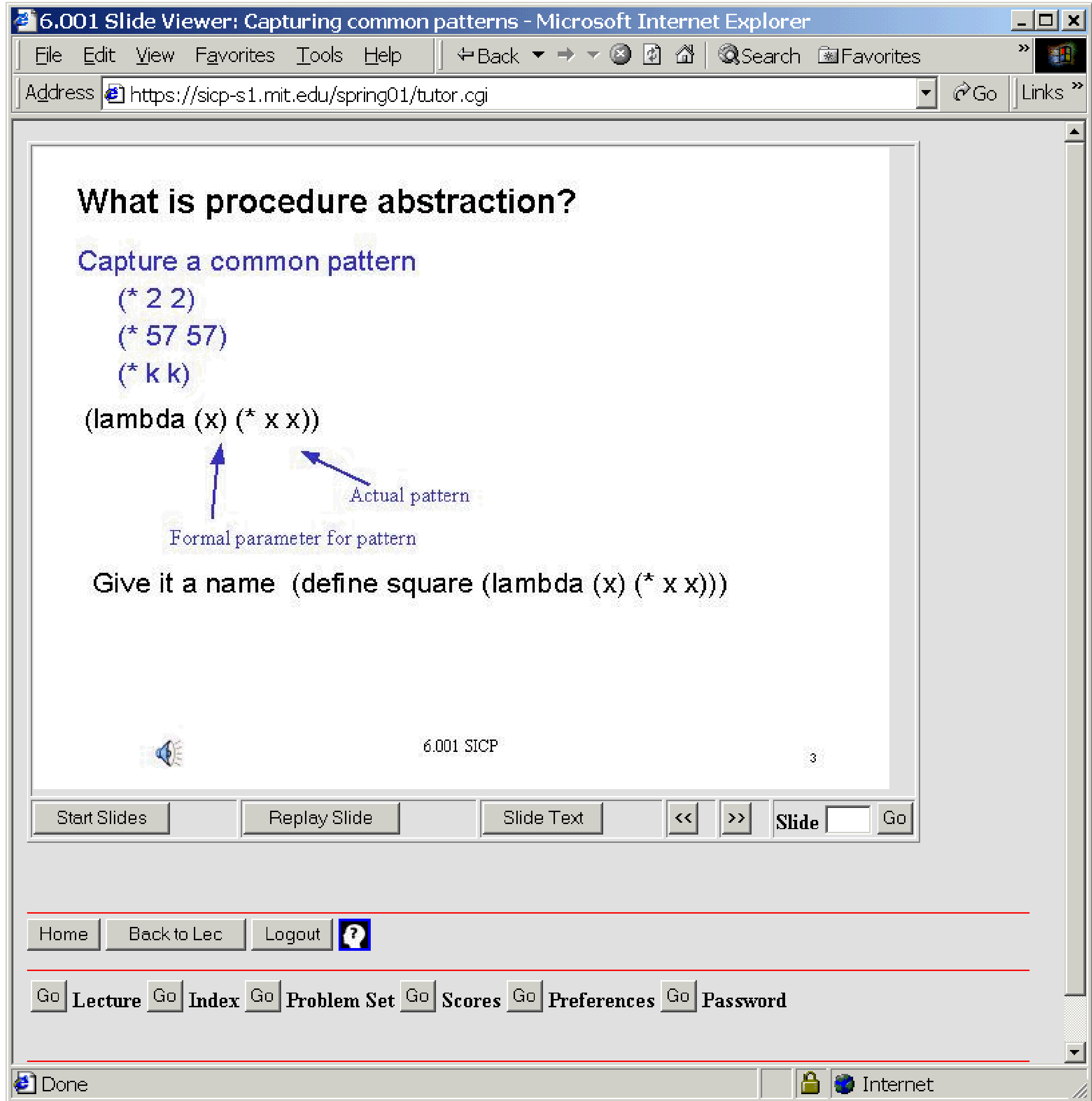
Prof's Grimson & Lozano-Perez

Massachusetts Institute of Technology

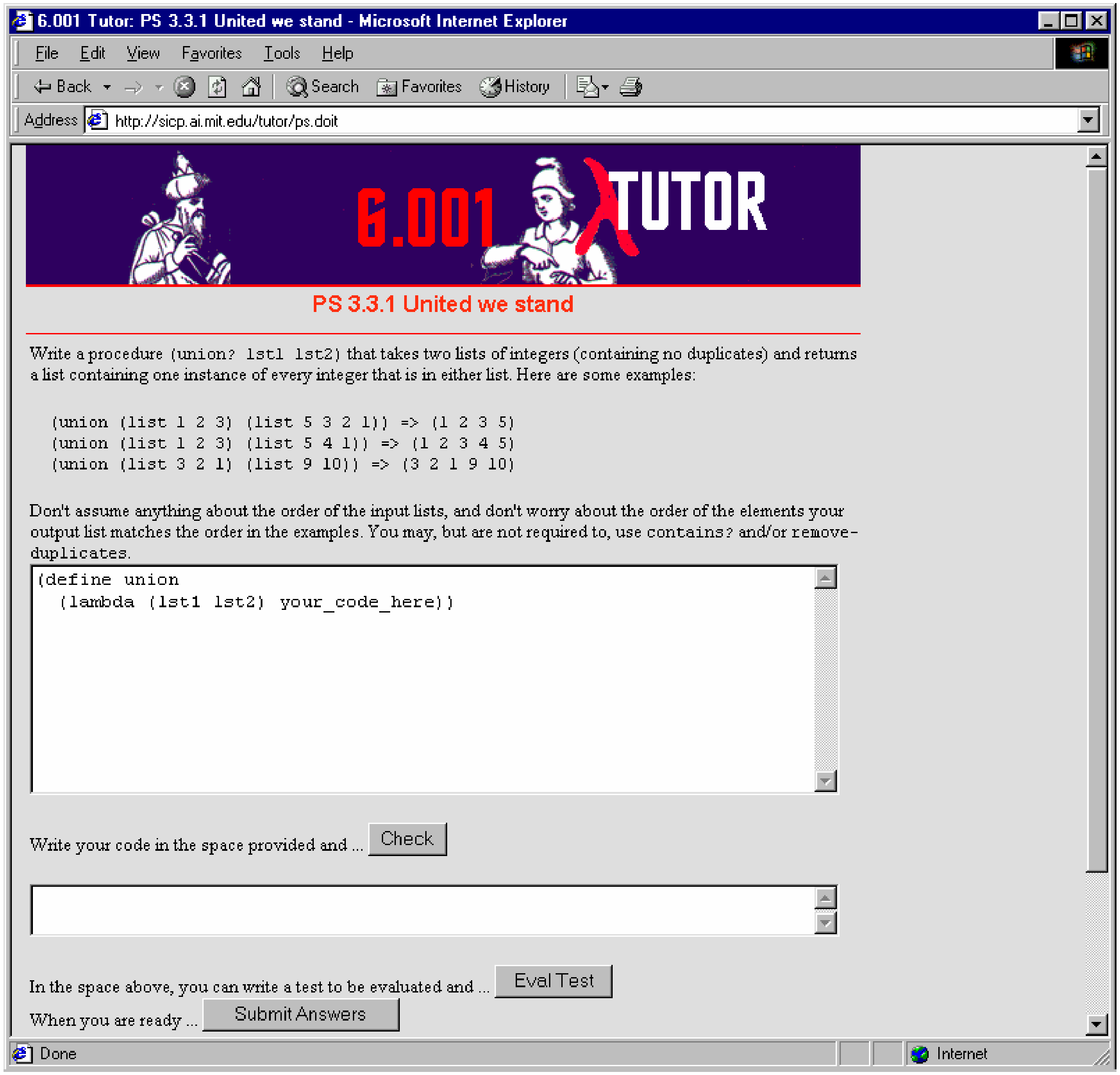
KUTZTOWN UNIVERSITY

Online Presentation Windows with Audio Stream





Online Problem Sets



Abstract

Is it possible for students learning from online presentations to achieve the same level of, or even improve on, academic performance in basic computer programming compared to traditional lecture-based education? A multi-treatment, repeated-measures, experimental design was used to compare differences in learning **broad concepts** and **specific details** (content type) during the course "Structure and Interpretation of Computer Programs." 168 students completed all the research instruments and course requirements. Six **live lectures** were matched with six **online presentations** (delivery method) crossed with content type, resulting in a 2x2 matrix: online broad, online detail, live broad and live detail. Online, web-based presentations contained windows for scrolling the presentation text, an audio narration of the presentation text by Professor Grimson, and an animated PowerPoint chalk-talk that built with the audio. Presentations consisted of five (approximately) ten-minute segments, after which a few formative assessment questions were completed by the students for immediate automated feedback. Composite performance instruments were constructed from mid-term and final exam questions related to the three presentations/lectures in each condition (Cronbach's alpha = .67, .73, .81, and .59 respectively.) In a 2x2 repeated measures analysis, three of the four performance instruments indicated significant differences from the null hypothesis of average performance: higher online-broad scores ($F[1, 167] = 24.01$, $p < .0001$); higher online detail scores ($F[1,167] = 15.76$, $p < .0001$); and lower live broad scores ($F[1,167] = 10.22$, $p < .002$). Riding and Rayner's (1998) reaction time assessment of cognitive style was revised for, and collected via, web-based testing. The dimension of cognitive style (categorizing students as imaging, intermediate, or verbalizing) helps explain the lower performance in the live-broad condition. Intermediate students perform at the mean on the live-broad instrument while students with clear preferences for imaging or verbalizing are significantly below the mean ($F[2,165] = 3.05$, $p < .05$). ***Students with a strict preference for a particular cognitive style may be at a disadvantage in large lectures. Self-paced and reviewable online presentations with both auditory and visual components may permit students to compensate for their cognitive style differences.***

Online Presentations
Show Academic Performance Advantages
Over Auditorium Lectures

The first subject in computer programming offered by the MIT Department of Electrical Engineering and Computer Science is 6.001: Structure and Interpretation of Computer Programs. Professor Eric Grimson taught the 6.001 subject with an innovative structure, relying heavily on web-based materials, consistently for four semesters. Faculty members from other departments also expressed interest in the learning outcomes of this widely subscribed subject and in whether certain kinds of students achieved higher levels of performance. Further, did motivation, academic ability, or learning styles influence achievement? This study was conducted to answer these questions. The investigation of performance and cognitive style are reported here.

Partially supported by both the d'Arbeloff Fund for Excellence in Education and the iCampus Project, a collaboration of MIT and Microsoft Research, the study was charged to be more than an internal program evaluation and to make contributions in the fields of learning science and advanced educational technology.

The 6.001 students met in large (enrollment usually > 300), auditorium-style lectures only six times per semester; they met in a recitation section of 20 to 30 students and a faculty member twice a week; and they met in a tutorial group of 5 to 8 students and a graduate teaching assistant once a week. They also had access to one of the graduate laboratory assistants for help with their three assigned programming projects. Two 100-point quizzes were taken mid-term and a 200-point final exam was taken at the end of the semester.

Historically, the students enrolled in 6.001 would have met in an additional twenty-two auditorium-style lectures and been given weekly problem sets to be returned for grading. These meetings and assignments have been replaced with web-based online materials. Each of the twenty-two online presentations contains approximately fifty-minutes of audio/text/graphic material delivered by Professor Grimson in four or five sequential parts. Between each part of an online presentation, students respond and receive automated feedback to a few questions as non-evaluated formative assessment.

The online lectures may be accessed at any time and any part may be repeated as often as desired. Eleven problem sets are also completed and submitted online and have the additional web-enabled features of check and hint buttons. With the check button, solutions to the programming problems may be run against models designed to find bugs and errors resulting from common misconceptions or ordinary carelessness. The hint buttons were not yet fully implemented at the time of this study but provided suggestions at a few known choke points. The check and hint buttons could be used by students for immediate feedback, without penalty, before electronically submitting their solutions for evaluative grading with a deadline at the end of each week.

Method

Participants

Consisting of required academic activities, participation in this study was not voluntary but completion of instruments introduced specifically for the purpose of research was voluntary. During spring semester 2002, 347 students enrolled in 6.001; 168 (48.4%) of the enrolled students completed all the research instruments and subject requirements. Of these students, 37.3% were female and 61.9% were male (0.8% unknown); 37.3% were Caucasian American, 27.6% Asian American, 3.7% African American, 10.5% Hispanic/Latino American, 0.6% Native American, and 9.7% International (10.6% unknown). Age ranged from 16 to 30-years with an average of 18.8 years. Their 1st-year, fall semester GPA ranged from 1.25 to 5.00, with a 4.20 average GPA and a 0.62 standard deviation. The points earned in 6.001 ranged from 54.0% to 95.0% with a 76.89% average and a 10.22% standard deviation.

Procedure

Experimental content to be learned by students was presented in twelve subject units, filling out a 2x2 matrix with three subject units in each of the four cells: online broad, online detail, live broad and live detail. Questions included on the regular midterm and final examinations had been prepared to assess the information presented in the twelve subject units. Those questions pertaining to the same presentation method and content type were combined into composite performance measures, evaluating learning under that condition. Additional data on student satisfaction, academic motivation, and cognitive style were collected through the web at the end of the semester with the informed consent of participants. (Only cognitive style is considered here.) It was explained to students that an outside evaluator (the first author) would insulate the instructor from the research-based information provided. The information from all instruments was clearly explained to be "not" anonymous so it could be related to performance measures but that confidentiality would be protected. Information regarding the sex, ethnic identity, and 1st-year fall-semester GPA was obtained from the Office of the Registrar.

Instruments

Performance. The overall 6.001 grade was determined by the total points earned in the semester from quizzes, exams, projects, and problem sets. Four specific research instruments were constructed to assess learning from each of the experimental conditions using exam questions that were "natural" to the usual pedagogical procedures of the instructor. To establish consistency in unit of measurement for the four instruments, raw scores were converted by computing the percentage of possible points achieved for each instrument. The percent score (PS) was then transformed into a standard score (z-score) by subtracting the mean and dividing by the standard deviation of the percent scores aggregated across all four instruments. [(PS – 72.6652) / 18.82654 = Z] This procedure placed all four instruments on the same scale of measurement.

- *Online Broad.* Constructed of one question from each of quiz1, quiz2, and the final exam for a total of 56 possible points. A Cronbach's alpha = 0.67 indicates modest consistency among the items.

- *Online Detail.* Constructed of two questions from quiz1, one from quiz2, and two from the final exam for a total of 88 possible points. A Cronbach's alpha = 0.73 indicates adequate consistency among the items.

- *Live Broad.* Constructed of one question from quiz1, two from quiz2, and one from the final exam for a total of 82 possible points. A Cronbach's alpha = 0.81 indicates good consistency among the items.

- *Live Detail.* Constructed of one question from quiz2 and two from the final exam for a total of 55 total possible points. A Cronbach's alpha = 0.59 indicates low (but not terrible) consistency among the items.

Performance Correlations in 6.001

	Online Broad	Online Detail	Live Broad	Live Detail	Subject Points
Online Detail	.68				
Live Broad	.73	.77			
Live Detail	.61	.72	.67		
Subject Points	.70	.84	.81	.79	
1st-year Fall GPA	.36	.45	.44	.40	.58

All correlations significant at the .001 alpha level.

Cognitive Style Analysis (CSA). The CSA was created by Riding and Rayner (1998) and published by Learning & Training Technology, Inc. The CSA had been constructed for computerized administration with a dedicated workstation. With permission, Carter Snowden, of MIT Information Services, wrote a Java script to administer the CSA over the web. The CSA observes reaction time to embedded figure and word comparison tasks. It then describes test-takers along two dimensions of cognitive style: **imager/verbaliser** and **analytic/wholistic**. Assessments are provided in two forms: (1) categories of verbaliser/intermediate/imager and wholistic/intermediate/analytic and (2) ratios for I/V (imager > 1) and A/W (analytic > 1). These dimensions are normed for the general high school and college population of the United Kingdom. In the present study, the two ratios were confirmed to be orthogonal to each other (r = -.04, **p** = .62). Frequencies with which the CSA categories are represented among the 6.001 students are detailed in the following table.

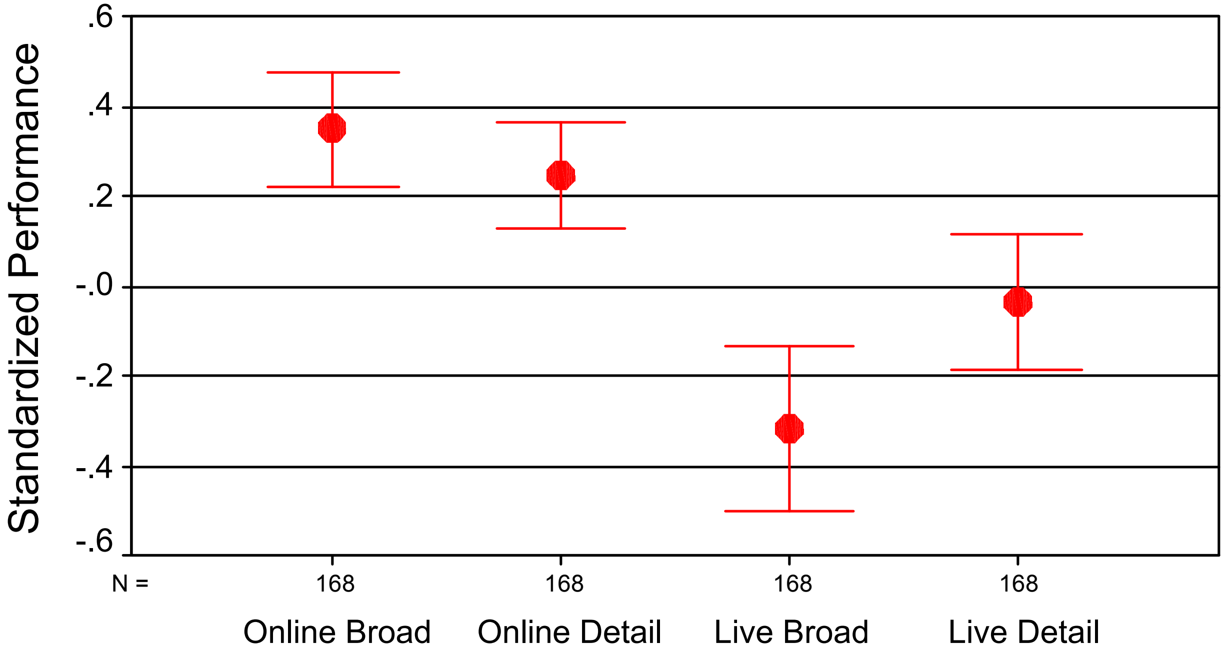
Distribution of Cognitive Style in 6.001

	Imager	Intermediate	Verbaliser
Analytic	12.0%	4.0%	7.4%
Intermediate	12.6%	10.9%	8.0%
Wholistic	20.6%	13.7%	10.9%

6.001 Performance Differences

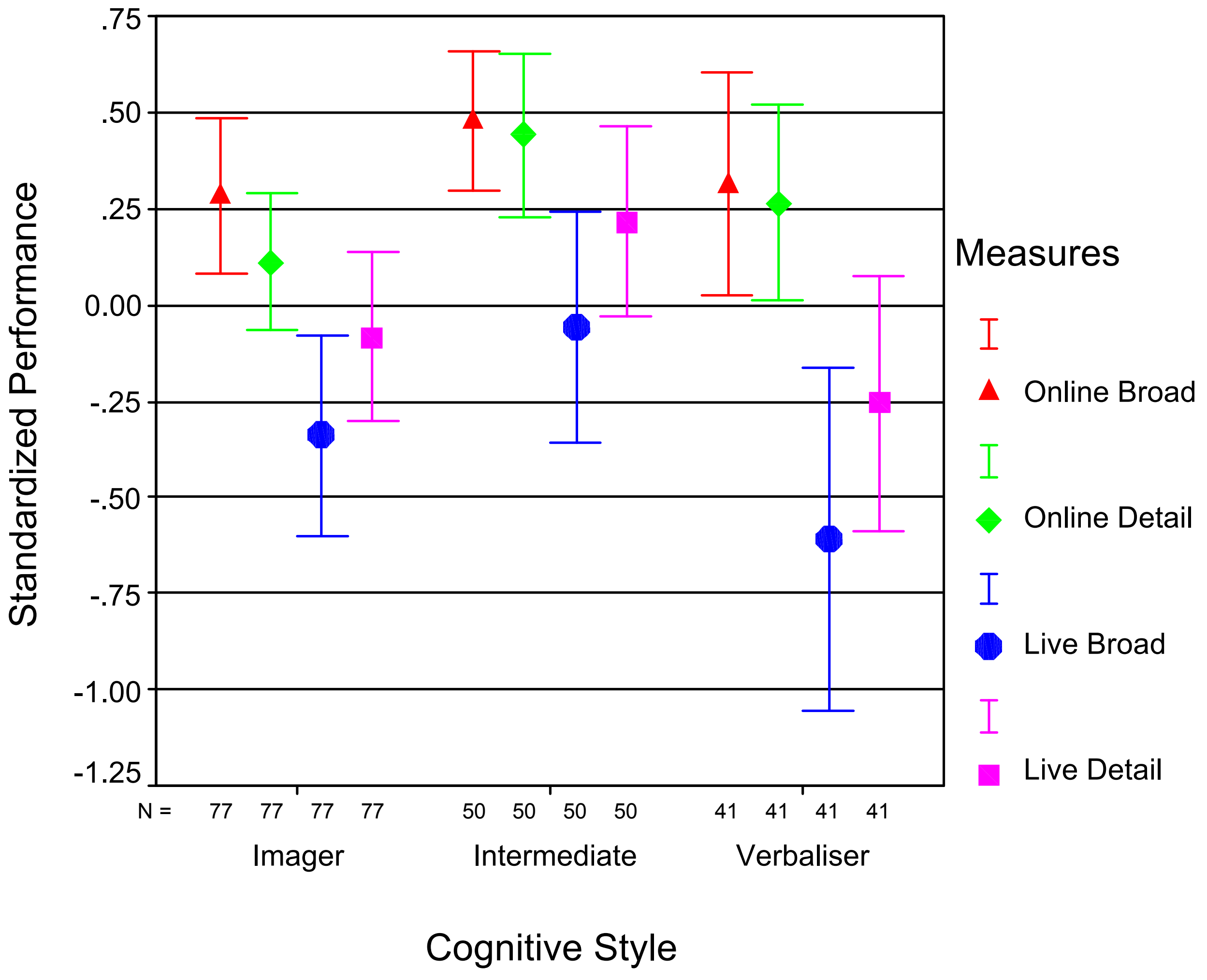
Based on Source and Kind of Information

Means and 95% Confidence Intervals



Scale presented in standard deviations around a mean of zero

6.001 Performance Differences for Imagers and Verbalisers



Results

Three of the four experimental performance measures show significant differences from the null hypothesis of average performance. See Figure 1.

- Online broad **F**[1,167] = 24.01, ***p*** < .0001
- Online detail **F**[1,167] = 15.76, ***p*** < .0001
- Live broad **F**[1,167] = 10.22, ***p*** < .002
- Live detail **F**[1,167] = 0.88, ***p*** = .35

There were also performance differences based on the Imager/Verbaliser (I/V) dimension of cognitive style in students. In the condition of Live Broad, the dimension of Imager/Verbaliser, showed significant performance differences (**F**[2,165] = 3.05, ***p*** < .05) in which Intermediates scored 0.68 higher than Verbalisers (*p* < .04). See Figure 2.

When 1st-year fall-semester GPA is included as a covariate, two performance instruments show significant differences based on cognitive style. With inclusion of the covariate, again in the condition of Live Broad, the dimension of Imager/Verbaliser revealed significant performance differences (**F**[2,165] = 4.73, ***p*** < .01) in which Intermediates scored 0.74 higher than Verbalisers (***p*** < .008). Also, the dimension of Imager/Verbaliser showed a newly significant difference (**F**[2,165] = 3.39, ***p*** < .04) for the condition of Online Detail, in which Intermediates scored 0.39 higher than Imagers (***p*** < .03).

Cognitive Style Ratio and Performance Score Correlations

	Online Broad	**Online Detail**	**Live Broad**	**Live Detail**	**Subject Points**	**1st-year Fall GPA**
A/W-ratio	-.18*	-.15	-.10	-.12	-.17*	-.10
I/V-ratio	.01	-.13	.02	.03	-.02	-.06

* *p* < .05;

Discussion

- That the online learning appears to have been so dramatically successful is a strong argument for continuing. The online presentation of information appears to result in better academic performance than auditorium-style lecturing. Online presentation of broad concepts and online presentation of the details of programming tested significantly above the average for all knowledge sources. Live auditorium lectures of broad concepts tested significantly below the average performance and live auditorium lecture of the details of programming was not significantly different than average performance.

- There was clearly an inverted-U shape to performance in all learning conditions as related to the cognitive style of Imager/Verbaliser. However, this I/V-based performance difference only reached statistical significance in the learning condition with lowest overall performance – live auditorium lectures on the broad concepts of programming. A possible interpretation is that students with a dominant cognitive style on the I/V dimension are at a disadvantage compared to students equally comfortable with both verbal and visual information. This appears to be most true of learning in a lecture when broad concepts are being addressed. If a broad concept is portrayed either through imagery or through language, part of the class may be left behind and unable to compensate. On the other hand, the online presentations included auditory, textual, and graphic representations of information. Students are able to select one or more presentation styles to conform to their cognitive style or compensate for misperceptions conveyed through one channel or another. The periodic formative evaluations would likely reinforce such a learning strategy.

- When the academic talent/preparation of students (as measured by 1st-year Fall GPA) is removed statistically, Verbalisers seem to be at a greater disadvantage in testing live broad concepts. In addition, now Imagers seem to be at a disadvantage in testing for programming details learned online. This raises many interesting possibilities: learning details of programming code may be less amenable to visualization and spatial processing; the visual component of the online materials for detailed coding may have been less adequate; or some other of several explanations. However, it is clear that regardless of a relative disadvantage for visual thinkers there is much better academic achievement when all material traditionally presented in large lectures is presented online instead.

- The cognitive measures showed two significant correlations with subject performance measures, the more important being between A/W-ratio and Online-Broad ($r = -.18$, $p < .02$). The more analytic a student the more difficulty with broad concepts presented online.

- The Institute-wide concerns over the adequacy of online content and problem-set presentation in 6.001 has been allayed. However, the lower test performance seen for broad conceptual information learned during auditorium-style teaching has not raised concern over lecturing. ☺

❖ Further analyses including academic achievement motivation (Elliot & Church, 1997) and student perceptions of the course are forthcoming. For instance, although student opinion generally rated the availability of wider internet connectivity during online learning to be of low usefulness, the ability to make connections during online presentations via instant messaging and searches was higher among Verbalisers. Overall, the rated usefulness of these connections became significantly higher in the most recent of four semester studied, which possibly has some link to the increasing familiarity of the online format. Verbalisers rate connectivity 0.68 above Imagers ($F[2,121] = 3.93$, $p < .02$).

References

Elliot, A, & Church, M. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 72*, 218-232.

Riding, R. J., & Rayner, S. (1998). *Cognitive styles and learning strategies.* London: David Fulton Publishers.

Acknowledgments

The authors would like to thank the d'Arbeloff Fund for Excellence in Education, and the iCampus collaborative with Microsoft Research for their support, and Dr. Lori Breslow, Director of the MIT Teaching and Learning Laboratory for her assistance with this research and helpful comments in preparing the presentation of results.

d'Arbeloff Fund for Excellence in Education

Microsoft Research

TLL Teaching & Learning Laboratory