Supporting Pen-Based Classroom Interaction:
New Findings and Functionality for Classroom Learning Partner

Kimberle Koile, Kevin Chevalier, Capen Low, Sanjutka Pal, Adam Rogal,
David Singer, Jordan Sorensen, Kah Seng Tay, Kenneth Wu

MIT Computer Science and Artificial Intelligence Laboratory
32 Vassar Street, 32-221
Cambridge, MA 02139 USA
kkoile@csail.mit.edu
{kcheval, capen, spal, arogal, singerd, jsorense, kahseng, zstrif}@mit.edu}

Abstract

This paper reports the latest developments for a wireless pen-based classroom interaction system, Classroom Learning Partner (CLP), and two new variations—Group Learning Partner (GLP) and Elementary Learning Partner (ELP). CLP consists of a network of Tablet PCs and software for posing questions to students, interpreting their handwritten answers, and aggregating the answers into equivalence classes. Through controlled classroom experiments, we have demonstrated that use of CLP improves student learning, especially among students who might otherwise be left behind. GLP extends CLP to support student group work in the classroom by supporting the transmission of ink between students. ELP is a version of CLP developed for use in elementary school classrooms.

1. Introduction

Many schools and universities today face the problem of maintaining a personal level of interaction between instructors and students in class. Due to the increasing number of students in the classroom, it is becoming harder for instructors to interact one-on-one with each student. The Classroom Learning Partner (CLP) system was developed to provide instructors with a new teaching method for large classes, a method that incorporates techniques known to improve student interaction and learning: active student engagement; processing, e.g., working problems to put new ideas into practice; and immediate feedback at the time a new concept is being learned, e.g., [3, 4, 14, 21]. Employing all these ideas simultaneously in the classroom can lead to a rich learning environment in which students can learn not only the subject matter, but also how to learn. The approach is similar to that of wireless polling devices, but with support for open-ended questions, which handwritten answers allow. (See [13] for an overview of polling systems.)

We have built a Tablet-PC-based system called Classroom Learning Partner (CLP) that supports this pedagogy, and we have deployed it in introductory computer science classes at MIT over the past two years. We have demonstrated through controlled classroom studies that use of the system improves student learning, particularly among students who might otherwise have performed poorly [8].

In this paper, we briefly describe the system, then present new findings from our recent learning study and describe new CLP features. We then discuss two new variations—Group Learning Partner (GLP), which facilitates peer learning in the classroom; and Elementary Learning Partner (ELP), which is designed for use with elementary school students.

2. Classroom Learning Partner

CLP is a wireless pen-based interaction system built on top of Classroom Presenter [2]. It consists of a network of Tablet PCs and supports the wireless submission of digital ink answers to in-class exercises. The instructor's slides are displayed both on a large screen as well as on the students' Tablet PCs. When the instructor displays a slide containing an exercise, the students write answers in digital ink on their Tablet PCs, and then wirelessly, and anonymously, submit their digital ink answers to the instructor. CLP aims to solve the problem of an instructor being overwhelmed with many student submissions. The key idea is to interpret the student digital ink answers and aggregate them, showing the instructor a histogram and representative answers. With such information the instructor can maximize the number of students to whom she gives feedback, by discussing a variety of submitted answers. Such a discussion also serves to engage students in the process of evaluating different

answers, and learning how to learn. The current version of CLP interprets and aggregates student answers that are numbers, strings, and sets or sequences of numbers or strings [7].

Below is an example of a student submission in an introductory computer science course at MIT (6.001). CLP has been deployed in the first author's 6.001 recitations in controlled studies over the past two years. (Students meet in lecture of 150-250 students twice a week; in recitations of 20-25 twice a week.)
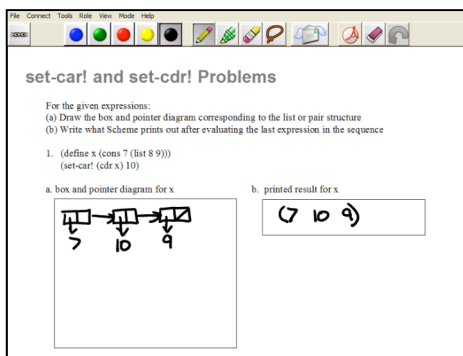

Figure 1. Example student submission using CLP

The rest of this section discusses the findings from our most recent learning study, and CLP's new features.

2.1 Fall '06 Study

Over the past two years we have seen statistically significant improvements in student learning in the first author's classes that have used Tablet PCs and wireless submission of digital ink answers to in-class exercises. In particular, in controlled studies in which the first author taught one class with Tablet PCs and one without, we consistently have seen learning gains in the bottom third of the Tablet PC class, i.e., among students who would otherwise have done poorly. This past term, Fall '06, we investigated performance and learning styles using the same methodology as in our previous two studies [8].

Performance. We saw no difference in scores in experimental (Tablet-PC) vs control (non-Tablet-PC) groups performing in the top half of the class. In the bottom half of the class, however, there was an increasing difference in performance scores as the scores decreased, especially in the bottom third of the class. (See Figure 2.) Statistical significance could not be determined because there were no students using Tablet PCs who scored in the bottom third of the class. The results do suggest, however, that the Tablet PC users in the bottom half of the class consistently performed better, especially the poorer performing students. This finding is consistent with our previous studies, which had statistically significant results [8].
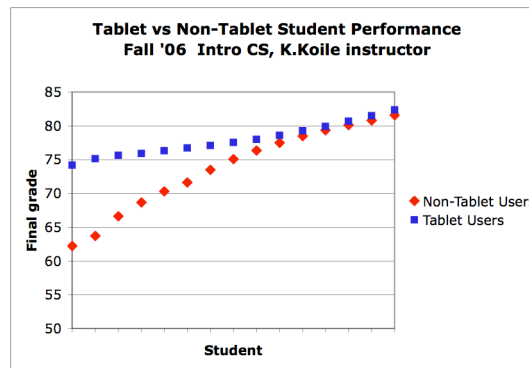

Figure 2. Student performance scores [7]

Learning styles. We assessed the learning styles of students in both the experimental and control groups through two surveys and classroom observations. Nine learning styles were assessed, including motor learning, motor processing, motor spatial learning, visual textural learning, visual spatial learning, visual color learning, aural learning, aural processing, and social learning. Two learning style preferences stood out when quiz 1 and quiz 2 scores were compared for Tablet and Non-Tablet users. First, students whose learning style emphasized solving written problems (motor processing) benefited when they had use of the Tablet PC and CLP, i.e., their increase in quiz 2 scores was statistically significant when compared with students who preferred other learning styles ($p < .02$). Second, students who used color when taking notes or solving problems (visual color learning) benefited from using the Tablet PC and CLP ($p < .015$). The Tablet PCs were introduced after quiz 1, so comparison of quiz 2 grades can be considered a direct measure of the impact of the Tablet PC and CLP. We speculate that both of these groups benefited from the pen-based input and feedback afforded by the combination of the Tablet PC and CLP.

2.2 Ink Interpretation

We are in the process of adding three important new features to CLP's ink interpreter: a new chunking algorithm and algorithms for interpreting two new answer types—sketched box-and pointer-diagrams and marked answers.

Chunking for Sequences. An important aspect of ink interpretation is the chunking of the collection of ink strokes into words and symbols before applying any interpretation algorithm on the chunks. Good chunking of ink strokes allows for interpretation accuracy to be improved by reducing the problem into smaller manageable chunks, and allowing us to apply different biasing factors on the appropriate chunks, e.g., preferring a number over a character. If an expected answer type is a number and CLP receives an answer such as S, for example, the ink interpreter can

Ink	CLPv2		CLPv1.5		CLPv1		Microsoft	
	Interpreted	%	Interpreted	%	Interpreted	%	Interpreted	%
[1 2 3]	[1,2,3]	100.00	TI,2,3]	71.43	->,23]	57.14	[I,23]	71.43
1, 3, 6, 10, 15	[1,3,6,10,15]	100.00	[1,3,6,10I15]	92.31	[li3,6,10,15]	84.62	[1,3,6,10115]	92.31
defg abc	[d,e,f,g,a,b,C]	100.00	[defy,abc]	60.00	[defy,abc]	60.00	[defog,abc]	66.67
ABEFGKLHCIJD	[A,B,E,F,G,k,L,H,C,I,J,D]	100.00	[A,B,E,F,G,k,L,H,C,IJD]	96.00	[ABE,F,Gk,H,->,JD]	64.00	[ABE,Fatal,H,CI,JD]	64.00

Figure 3. Interpretation results for four individual ink samples

identify the answer as a 5 rather than an S. After analyzing many chunking examples from interpreters that we have tried with CLP, namely Microsoft's built-in Ink Analyzer and our interpreters [11], we found that "incomplete chunking" contributes to poor interpretation of sequences. Incomplete chunking refers to cases in which an ink stroke collection could have been decomposed further into smaller collections before applying interpretation. Our latest interpreter employs new ideas that improve chunking and, hence, accuracy rates for interpretation, especially for sequences; e.g., analyzing individual strokes of ink and recursively combining strokes that intersect each other in order to identify chunks. Examples of chunking and accuracy measurements for each of the tested ink interpreters are shown in Figure 3. We measure accuracy as a percentage computed using the Levenshtein string edit distance [1] between the interpreted semantic representation and the expected representation. Figure 4 shows more results.

Example Sequence	CLPv2	Microsoft
[1 2 3]	85.71 %	63.10 %
1, 3, 6, 10, 15	98.46 %	92.31 %
defg abc	88.67 %	62.67 %
A B E F G K L H C I J D	90.67 %	52.89 %
a, b, c, d, e, f, g, h, i, j, k, l	81.33 %	58.67 %
#, #, # -> #	69.79 %	38.54 %
Total	**86.48 %**	**60.99 %**

Figure 4. Interpretation accuracy results computed using 10 ink samples per sequence

Box-and-Pointer Sketches. We are adding support in CLP for sketched answers, in particular, sketches of what is referred to as a box-and-pointer diagram. Such diagrams often are used to represent data structures in computer science courses. In 6.001, they represent a pair data structure in the Scheme programming language. The diagrams contain double boxes, with each box representing an element of the pair. Arrows point to values stored in the pair. By combining multiple boxes more complicated structures such as lists can be created. Examples of diagrams that represent lists are shown in Figures 1 and 5.

Box-and-pointer diagrams differ from other types in CLP in that they contain both handwritten text and sketched parts. We exploit this distinction and interpret the two parts separately, then combine the results to compute a diagram's semantic representation. Our text interpreter is used on the text parts; LADDER [6] is used on the sketched parts.
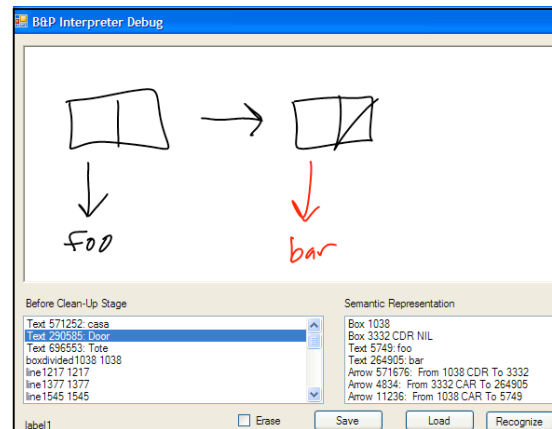

Figure 5. Interpreted box-and-pointer sketch

The output from the text and sketch interpreters is often not correct, often because text and shapes may share the same strokes. We are able to correct the errors, however, with multiple interpretation passes in which we repeatedly move ink strokes from the text object to the shape object then reinterpret the two objects. What remains after these passes is combined into a semantic representation such as the one shown at the lower right in Figure 5. Interpretation errors are visible in the lower left, labeled "Before Clean-Up Stage." Here, the text interpreter has recognized the arrow and the word "bar" as "Door". The clean-up stage removes the arrow from the ink for the word "Door". The text, shown in the lower right of the figure, then is interpreted correctly as "bar".

Marks. In addition to sketches, we are adding support in CLP for what we are calling marks, ink strokes that indicate a region. Marks differ from sketches in that their semantics is dependent on an associated background image. In Figure 6, students were asked to identify a particular programming construct in a diagram provided by the instructor.

Our mark interpreter is flexible enough to understand different ways of indicating a region, e.g., someone may encircle, fill in, cross out, or place check marks in a region to select that region. To increase the efficiency and accuracy of the interpreter, we bias the interpreter for each problem by using instructor-provided information about likely stroke types. In the example above an instructor can indicate that students are likely to encircle or check particular regions.
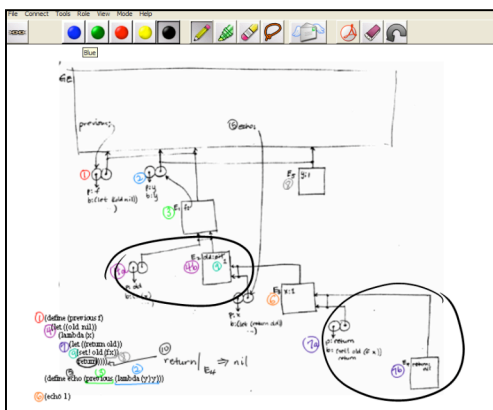

Figure 6. A student answer with circled regions

For the interpreter to understand X's or check marks, we use velocity and angle data to determine whether a stroke is part of an X or a check mark. To decide whether a region is filled in, the interpreter must factor in the width of the pen tip selected by the student when filling in an area, as well as closely packed multiple strokes or squiggly strokes consisting mostly of parallel regions, as shown in Figure 1. Encircled regions are a challenge: Because regions may be encircled incompletely or with overlapping strokes, the interpreter must piece together the outline of each region. Examples of our interpreter working on encircled regions are shown in Figures 7 and 8.
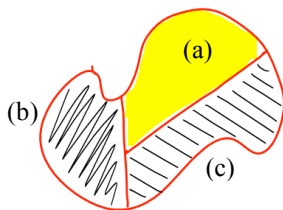

Figure 7. Selecting parts of a region (a) highlighting, (b) one squiggly stroke, (c) multiple parallel strokes
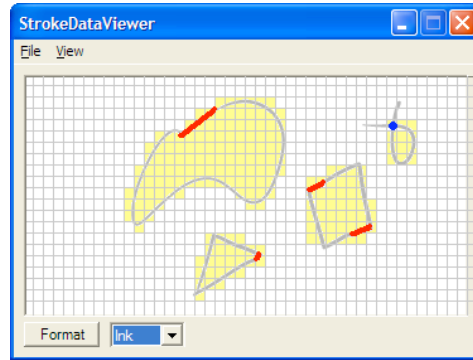

Figure 8. Interpreter results for encircled regions; the interpreter breaks up the area into a grid and identifies the (yellow) grid squares as selected regions

2.3 Aggregation

The answer types in the current CLP system are aggregated using similarity measures related to Levenshtein string edit distance [1], which measures the distance between strings defined as the minimum number of operations needed to transform one string into the other. (See [7, 15] for aggregator details.) Our new aggregator work adds methods for aggregating our sketch and mark answer types.

Sketches. To aggregate box-and-pointer diagrams, we are starting with acyclic diagrams that can be represented by tree structures. This assumption allows us to use tree edit distance algorithms, such as [19], which generalizes string edit distance. The goal is to compare nodes and edges of two trees in search of the largest common substructure as a measure of similarity between the two trees.

We hope to generalize our tree edit distance approach to deal with graph structures, using as a similarity metric the idea of subgraph isomorphism, e.g. [20]. Within the context of CLP, the graph sizes most likely will be small so that we can avoid the exponential growth usually associated with subgraph isomorphism algorithms.

Marks. An approach to aggregating marks on a background image must take into account the context of each mark. The instructor defines context by specifying regions that will correspond to equivalence classes for student marks associated with a particular exercise. The mark aggregator then puts all answers either into an instructor-specified equivalence class or into another miscellaneous group of unmatched student answers. Since regions specified by students will not necessarily coincide pixel-by-pixel with answer regions specified by an instructor, the mark aggregator must not depend on exact matching. Instead, it groups student answers by calculating a similarity measure based on percentage of overlap with instructor answer regions.

3. Group Learning Partner

Research has shown the value of peer learning and small group work, e.g., [4, 9, 21]. Could CLP be used to foster group interaction in a large class? We implemented a version of CLP called Group Learning Partner (GLP) to explore this issue. The key idea: Support wireless ink transmission between students in a group, rather than just between instructor and students. To implement this idea, CLP's interface was augmented to include group creation and group monitoring interfaces for the instructor, a group activity interface for students, and a new wireless ink transmission component. GLP is tightly integrated with CLP so that an instructor can switch easily between individual and group activities.

Instructor. We implemented four ways in which an instructor could create student groups: by students' geographic locations in the classroom, students' names or machine names, randomly, or by letting students select groups themselves. We also implemented an interface that allows an instructor to view a list of students in a particular group, a group's submitted answers, and if desired, a group's current shared slide.

Student. The student group activity interface is shown in Figure 9. It contains a panel that allows students to switch between viewing the instructor's original deck, and their personal and group slides; a main slide view of the shared group slide; and a "buddy list" panel. The group slide contains the group's answer to an in-class exercise; it is worked on by members of the group then submitted to the instructor in the same manner as non-group answers are submitted in CLP. Group members are listed in the panel; the member with the pen icon next to his or her name has current control of the pen. Only one group member at a time writes on the shared slide so that members don't "clobber" each other's ink by writing in the same place.

New ink transmission. In order to direct ink between students, rather than just between students and the instructor, we implemented the notion of a unique group identifier which was used to identify students in a particular group. That identifier is attached to each ink "scribble" object created by a group member, and only machines with a user having the same identifier as the scribble will display ink on their screens.

Evaluation. We evaluated GLP in one of the first author's 6.001 recitation sessions. Two rounds of group activity were conducted. The students in the class had no difficulties learning to use GLP, since they already had been using CLP for several months. Our observations suggested possible improvements, e.g., students could not erase another group member's ink on the shared slide—originally thought to be a guard against malicious erasure—but found occasion to need to modify others' ink. Students also wanted a group scratch slide because they were reluctant to do scratch work on the group shared slide, thinking it reserved for the final answer to be submitted to the instructor. On the whole students enjoyed using GLP to work on a group activity. The concept of shared and personal slides worked well. The instructor found the group creation mechanism easy to use and enjoyed being able to vary the classroom dynamic while still employing in-class exercises and feedback.
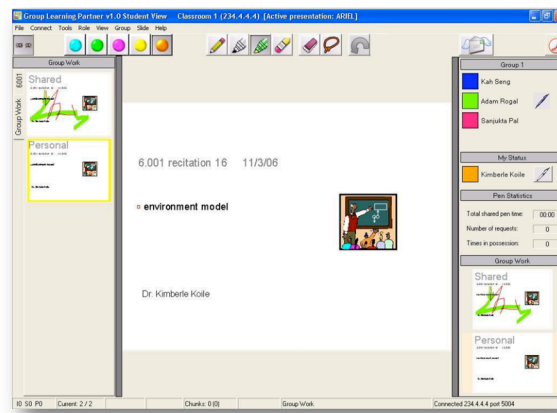

Figure 9. Student group activity slide

4. Elementary Learning Partner

Educators at all levels grapple with many of the same issues—keeping students engaged, increasing instructor-student interaction, improving student learning, dealing with increasingly larger class sizes, meeting the needs of students with different learning styles and skill levels. Could CLP demonstrate success in K-12 as it has in undergraduate classes? In particular, could the pen-based interaction it affords benefit young students, since they don't use keyboards? Elementary Learning Partner (ELP) is a version of CLP that we developed for use in elementary school classrooms.

To create ELP. we added more colors of ink, as requested by beta-testing first graders; and simplified the top menu bar by removing command icons that had proven distracting to adults. We deployed ELP in a first grade class at Fiske Elementary School, Lexington, MA, in June of 2006. (See Figure 10.) The teacher displayed slides containing exercises, the students worked the exercises and submitted their digital ink answers to the teacher—just as students do in the first author's undergraduate computer science classes. (The ink interpreter was not ready in time for us to use it with the first grader's answers.) We note three remarkable observations about the experience: (1) The teacher was able to teach with the system without prior training, having only seen a brief demo

the day before; (2) the students sat focused and attentive, working problems, for 1 hour and 20 minutes, and didn't want to stop; (3) the students' different learning styles were immediately apparent when they were asked to show their work on their slides. (See Figure 11.)

Quotes of note:

Principal: This experience is amazing; this bunch doesn't sit still for more than 20 minutes!

Teacher: Seeing the answers so quickly was fabulous. I could easily see whether they understood the ideas or not.

Student: I loved everything, but especially seeing what everyone wrote.

We cannot generalize from one experience, but the use of CLP-like systems in K-12 is promising, as many of the same issues that CLP addresses in undergraduate education are present in K-12 education as well. We note that interpretation of first-graders'

handwriting may pose a challenge, though the practice of training an ink interpreter for a person's particular handwriting may offer a solution. Wireless pen-based interaction, even without ink interpretation and aggregation, is promising for young children, as they too can benefit from a pedagogy that involves in-class exercises and immediate feedback, especially when the use of keyboards is not required. (See [5, 18] for examples of using Tablet PCs throughout a school.)

We note that the use of handheld computers has been successful in K-12 education, e.g., [12, 16, 17, 22]. Such machines do not have the screen space that Tablet PCs offer, but they may be a less expensive platform on which to deploy CLP-like systems. The key features needed are pen-based input, wireless connectivity, a shared public display, and support for the teacher when choosing answers to discuss. (See [9] for an example of using handhelds with a wireless presentation system in undergraduate classes.)

Figure 10. First graders using ELP; the teacher discusses with them their answers to exercises shown in Figure 11
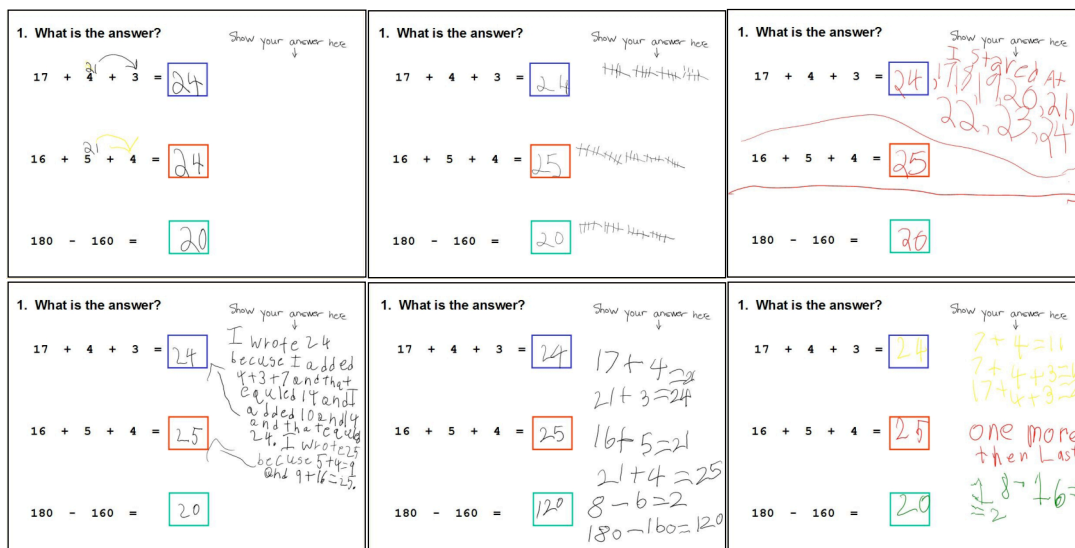

Figure 11. Sample submissions from first graders; note the different learning styles indicated by the different explanations for how the students arrived at their answers

5. Summary and Discussion

The goal of Classroom Learning Partner is to support a pedagogy that turns large classes into more of a two-way conversation between instructor and student than has been possible to date. It does so by actively engaging students in the learning process via pen-based wireless submission of digital ink answers to in-class exercises. It supports the instructor in giving immediate feedback to students by interpreting and aggregating the digital ink answers and providing summary information in the form of a histogram and representative student answers. With such information, both instructor and students can assess level of understanding and misconceptions, and an instructor can adjust course material in real-time.

CLP has demonstrated success in improving student learning. Our most recent learning study is consistent with our previous two studies: Use of a CLP-like system improves student learning among those students who might otherwise perform poorly. Our recent findings also indicate that students with a preference for working written problems, i.e., Motor Processing, benefit from use of CLP-like systems.

Our recent additions to CLP—new interpretation techniques, new answer types, and new versions for student groups and elementary students—hold promise for further improving classroom interaction and learning. We look forward to continuing our investigations and development efforts.

6. Acknowledgments

7. References

[1] Atallah, M. J. (Editor), *Algorithms and Theory of Computation Handbook*, "Levenshtein Distance (13-5)", CRC Press, 1998.

[2] Anderson, R., Anderson, R., Simon, B., Wolfman, S., VanDeGrift, T., and Yasuhara, K. Experiences with a Tablet-PC-based Lecture Presentation System in Computer Science Courses. In *Proc. of SIGCSE '04*.

[3] Black, P, and Wiliam, D. *Inside the Black Box: Raising Standards Through Classroom Assessment*. King's College, London, 1998.

[4] Brown, G. and Atkins, M. *Effective Teaching in Higher Education*, Routledge, Taylor & Francis, London, 1998.

[5] Djanogly City Academy, Notthingham, UK. http://www.djanogly.notts.sch.uk/C11/ICT/

[6] Hammond, T. and Davis. R. Ladder: A Language to Describe Drawing, Display, and Editing in Sketch Recognition. In *Proceedings of the IJCAI, 2003*.

[7] Koile, K., Chevalier, K., Rbeiz., M., Rogal, A., Singer, D., Sorensen, J., Smith, A., Tay, K.S., and Wu, K. Supporting Feedback and Assessment of Digital Ink Answers to In-Class Exercises. To appear in *Proceedings of IAAI 2007*, July, 2007.

[8] Koile, K. and Singer, D.A. Improving Learning in CS1 via Tablet-PC-based In-class Assessment. In *Proceedings of ICER 2006,* September 9-10, 2006, University of Kent, Canterbury, UK.

[9] Mazur, E. *Peer Instruction: A User's Manual*. Prentice Hall, Upper Saddle River, NJ, 1997.

[10] Ratto, M., Shapiro, R.B., Truong, T.M., and Griswold, W. The ActiveClass Project: Experiments in Encouraging Classroom Participation. In *Proceedings of CSCL '03*.

[11] Rbeiz, M. A. Semantic Representation of Digital Ink in the Classroom Learning Partner. MIT EECS MEng thesis. May, 2006.

[12] Roschelle, J. Unlocking the Learning Value of Wireless Mobile Devices. *Journal of Computer Assisted Learning*, 19(3), 2003, 260-27.

[13] Penuel, W.R., Abrahamson, A.L., and Roschelle, J. Theorizing the Transformed Classroom: A Sociocultural Interpretation of the Effects of Audience Response Systems in Higher Education. In D. Banks (Ed.), *Audience Response Systems in Higher Education: Applications and Cases*. Information Science Publishing, PA. 2006, 187-208.

[14] Skinner, E. A. & Belmont, M.J. Motivation in the Classroom: Reciprocal Effects of Teacher Behavior and Student Engagement Across the School Year. *Journal of Educational Psychology*, 85(4): 571-581, 1993.

[15] Smith, A. C. Aggregation of Student Answers in a Classroom Setting. MIT EECS MEng thesis. August, 2006.

[16] Soloway, E. , Norris, D., Blumenfeld, P., Fishman, B., Krajcik, J., and Marx, R. Handheld Devices are Ready-at-Hand. *Communications of the ACM*, 44, 2001, 15-20.

[17] Tatar,. D., Roschelle, J., Vahey, P., and Penuel, W.R. Handhelds Go to School: Lessons Learned. *IEEE Computer*, 36(9), 2003, 30-37.

[18] Villa Duchesne/Oak-Hill School, St. Louis, MO. http://web.vdoh.org/teachers/edtech/tabletPC.htm

[19] Wang J.T.L., Shapiro B.A., Shasha D., Zhang K.: Treediff: Approximate Tree Matching for Ordered (and Unordered)Trees.[http://cs.nyu.edu/cs/faculty/shasha/papers/tree.html]

[20] Wang J.T.L., Shapiro B.A., Shasha D., Zhang K., Currey K.M.: An Algorithm for Finding the Largest Approximately Common Substructures of Two Trees. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (PAMI) 20(8): 889-895, 1998.

[21] Wright, A.W. *Teaching Improvement Practices: Successful Strategies for Higher Education*, Anker Publishing Co, MA, 1995.

[22] Zurita, G. and Nussbaum, M. Computer Supported Collaborative Learning Using Wirelessly Interconnected Handheld Computers. *Comput. Educ.*, 42(3), 2004, 289-314.