Multiple Mouse Text Entry for Single-Display Groupware

Saleema Amershi[1], Meredith Ringel Morris[2], Neema Moraveji[3], Ravin Balakrishnan[4], Kentaro Toyama[5]

[1]samershi@cs.washington.edu, [2]merrie@microsoft.com, [3]neema@moraveji.com, [4]ravin@dgp.toronto.edu, [5]kentoy@micrsoft.com

[1]Computer Science & Engineering, DUB Group, University of Washington, Seattle, WA, USA
[2]Adaptive Systems and Interaction, Microsoft Research, Redmond, WA, USA
[3]School of Education, Stanford University, Stanford, CA, USA
[4]Department of Computer Science, University of Toronto, Ontario, Canada
[5]Technologies for Emerging Markets, Microsoft Research India, Bangalore, India

ABSTRACT

A recent trend in interface design for classrooms in developing regions has many students interacting on the same display using mice. Text entry has emerged as an important problem preventing such mouse-based single-display groupware systems from offering compelling interactive activities. We explore the design space of mouse-based text entry and develop 13 techniques with novel characteristics suited to the multiple mouse scenario. We evaluated these in a 3-phase study over 14 days with 40 students in 2 developing region schools. The results show that one technique effectively balanced all of our design dimensions, another was most preferred by students, and both could benefit from augmentation to support collaborative interaction. Our results also provide insights into the factors that create an optimal text entry technique for single-display groupware systems.

Author Keywords

Multiple mice, text-entry, education, children, ICTD, SDG.

ACM Classification Keywords

H5.3. Information interfaces and presentation (*e.g.,* HCI): Group and Organization Interfaces – *CSCW*.

General Terms

Design, Human Factors, Experimentation

INTRODUCTION

Single-display groupware [26] systems enable students to concurrently share and interact with a computer via mice and on-screen cursors. Such setups have recently received attention from researchers and educators in developing region classrooms because of their potential to increase student motivation, engagement, and social interaction during learning while dramatically reducing the per-student cost of computing [11, 12, 16, 17, 18, 21]. Even compared to low-cost computers (e.g., the One Laptop Per Child's XO laptop costs US$199 [14]), single-display, multiple mouse systems are less expensive in overall cost of ownership, maintenance, and administration.

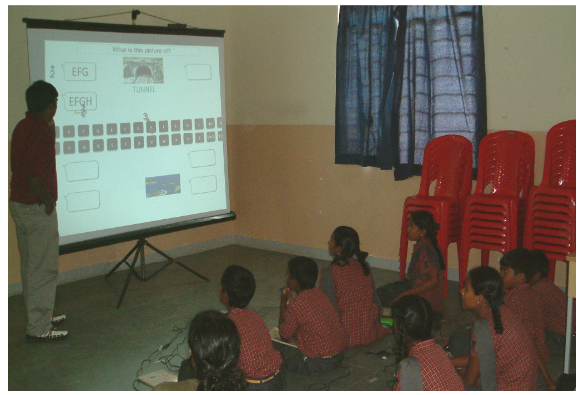

Figure 1. Students using mice for a text entry activity.

To date, most classroom-based single-display groupware systems involving mice have remained limited to point-and-click tasks. This suffices for simple true-false or multiple-choice based activities, but a natural question that follows is how to best allow students to enter text – a crucial requirement in an educational setting. While multiple-choice type activities are useful for factual recall, text-based short-answer questions require greater conceptual understanding [24] and result in greater long-term retention of content [5]. When we asked teachers in rural developing region schools how they might use text-entry enabled multiple mouse single-display groupware systems, they proposed several activities including labeling diagrams and processes, filling in the blanks in sentences, identifying images or spelling out words spoken by the teacher, collaboratively writing grammatically correct sentences,

CSCW 2010, February 6–10, 2010, Savannah, Georgia, USA.

collaboratively completing crossword puzzles, and free-form question answering.

Multiple keyboards could be deployed with mice to enable text entry in these single-display groupware systems, but this would incur a significant additional cost (at ~$10/keyboard compared to ~$2/mouse) as well as a more cumbersome hardware setup; in developing regions where multiple mouse systems are used, these costs and technical factors are prohibitive. On-screen 'soft' keyboards are also not wholly desirable due to their considerable screen footprint, an important consideration in single-display groupware environments.

Our ultimate goal is to improve the educational experience of students using such systems in their classrooms. However, before the educational value of text-entry enabled mouse-based single-display groupware can be evaluated, it must be effectively enabled. In this paper, we explore text entry methods for the multiple mouse single-display groupware classroom scenario (Figure 1). Soft keyboards and other methods for text entry have been explored for single-person use (e.g., [8, 10, 20, 22, 23, 30]), but the multiple mouse scenario imposes additional challenges that come with multiple users, cursors, and shared screen real estate. Our contributions include an exploration of the design space for this problem, 13 text-entry techniques specifically designed for our target scenario, and the results and analysis from a three-phase evaluation of the 13 techniques with over 40 students in a 14 day period in two resource-constrained educational institutions in South India.

RELATED WORK

Several technologies for classroom-wide, simultaneous student participation have been targeted toward students in developed world classrooms [1, 27, 29]. These systems are often prohibitively expensive for schools in developing regions, or only support limited student interaction. HubNet [29] supports participatory simulations using networked computers or graphing calculators. Individual computers and graphing calculators (common in U.S. schools) allow for rich forms of input and output, but are relatively costly (~$80-$100 per graphing calculator). Audience Response Systems (ARS), often referred to as "clickers" (~$25 per clicker, ~$100 for the hub), allow individual student input to be aggregated and shown on a projected display to foster classroom discussion. Clickers have had penetration in developed world schools [1], and have been shown to increase student engagement and motivation (e.g., [2]), but are still economically infeasible in the developing region context, and are limited to question-answer based activities.

The One Laptop per Child (OLPC) initiative [14] specifically targets children in developing regions. The laptops' mesh networking has the potential to connect teachers and students in classrooms for collaborative action. However, Patra et al. [19] argue that providing students with their own input device (e.g., a mouse) connected to a shared computer may be as pedagogically effective as individual laptops for some learning outcomes, while other researchers point out the social and organizational value of shared-display systems in the classroom [13]. The low cost of single-display groupware systems (approximately $2-3 per mouse, $20 per USB hub, plus $400 for a computer and $300 for a projector compared to a $199 laptop for each student) make them promising for further research.

The multiple mouse and cursor model has been evaluated in both developed and developing countries, but mostly in small group settings. These studies have shown that providing groups of two [6] and five [21] students with their own mice and cursors positively impacts their motivation and engagement compared to a single mouse shared amongst a group. The Mischief system [11] was designed for classroom-wide use (i.e., 10-30 students, each with a mouse and cursor, and a shared, projected display). Mischief currently includes both point-and-click based activities (e.g., clicking on answers to questions on-screen) as well as text entry, but only with a standard on-screen keyboard laid out alphabetically, which consumes much of the shared display's valuable screen real estate. While there exists a reasonable body of research on text entry techniques, these have focused on the single user scenario (e.g., [8, 10, 20, 22, 23, 30]). Building upon this prior work, our research explores alternative text entry techniques for the multiple mouse and cursor classroom scenario, which may better balance the tradeoffs between speed and accuracy, screen footprint, and other design factors inherent in a multi-user scenario.

MULTIPLE MOUSE TEXT-ENTRY TECHNIQUES

There are several design factors to consider for mouse-based text entry techniques for use by multiple students simultaneously in a classroom setting:

Cost. Given the resource constraints of our target schools, minimizing cost is crucial. Conventional wired mice with two buttons and a scroll wheel cost about ~$2. We also considered wired mice with five buttons (~$6). Wireless mice would alleviate the mess of multiple wires, but are significantly more expensive than wired mice.

Screen Footprint. Screen real estate is precious, particularly given the mix of educational content, multiple concurrent users, and limited resolution. We used a 1024x768 pixel display, which is common in developing region schools.

Scalability. Multiple mouse single-display groupware systems require techniques that scale as the number of students increases from 1 to 30 or more.

Leveraging Multiple Users. Investigating systems that take advantage of the potential to enter text collaboratively could offer advantages in both efficiency and pedagogy.

Learning Rate. A rapid learning rate for a text entry technique leaves more time for learning actual content.

Speed, Accuracy, and User Preference. As with any input technique, we seek to maximize these three factors.

We developed 13 mouse-based text-entry techniques to investigate. These were either based directly on existing text-entry techniques in the literature (possibly modified to better suit a multiple user scenario), or specifically designed to leverage the design considerations discussed above. They fall roughly into five categories: on-screen keyboards, multi-letter keyboards, scrolling techniques, collaborative techniques, and other more advanced techniques.

We focused on English-language text, but we believe these techniques could apply to other character-based languages. Each technique enabled entry of all 26 English characters, plus a 'space' character and a 'delete' key which functioned as 'backspace'. All 13 techniques were incorporated into the Mischief system [11]. For each technique, each student had their own cursor and a designated space on-screen (called a 'Blank', see Figure 4) in which their entered text was displayed, except for one of the collaborative techniques where output was displayed collectively.

We chose not to consider any auto-completion techniques (e.g., T9) due to concerns over complexity. However, auto-completion could augment any of these techniques, and is best left for future exploration in its own right.

On-Screen Keyboards

We evaluated three on-screen keyboards:

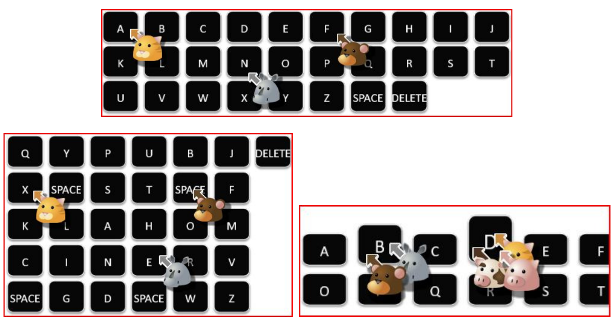

Figure 2. On-screen keyboards: Alphabet (top), GAG II (bottom left), and Fish-Eye (bottom right).

Alphabet Keyboard. The characters are positioned in three rows alphabetically on-screen (Figure 2, top). This was the baseline for our evaluations and was chosen over a QWERTY layout, since the children in our studies were not familiar with QWERTY keyboards and the alphabetical layout has been used for text entry in developing region settings in previous research [11].

GAG II Keyboard. Several on-screen keyboards optimized for speed have been presented in the literature (see Zhai, *et al*. [32] for a review). Of these, we chose the top performing *GAG II Keyboard* [23] (Figure 2, bottom left).

Fish-Eye Keyboard. Multiple cursors can occlude on-screen keys and inhibit visual search. Thus, we created a *Fish-Eye Keyboard* whose keys, arranged in alphabetical order in two rows, increased in height proportionally to the number of cursors hovering over them (Figure 2, bottom right).

Multi-letter Keyboards

On-screen keyboards typically have a large screen footprint. Thus, we considered more compact, multi-letter keyboards, where several consecutive letters are mapped to each key:

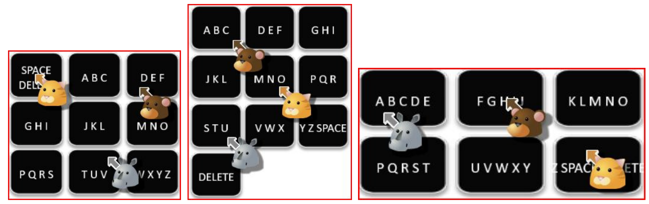

Figure 3. Multi-letter keyboards: Multi-tap (left), Triplet (middle) and Quintuplet (right)

Multi-tap Keyboard. Modeled after text entry on 12-button mobile phone keypads [20], ours was a 9-button keyboard, wherein letters appearing on the same key are selected by clicking repeatedly on that key (Figure 3, left). As a user cycles through the characters on any given key, the current character appears over that user's Blank. This character is selected after a 500 millisecond timeout.

Triplet Keyboard. (Figure 3, middle). This leverages the three buttons (or two buttons plus clickable scroll-wheel) on a standard mouse. Three consecutive letters appear on each key (in contrast to the Multi-tap keyboard which has 4 characters on some keys, as on dialpads). Pointing to a key and pressing the left, middle, or right mouse buttons selects the respective character (e.g., to select 'B' a user clicks the middle button while atop the 'ABC' key).

Quintuplet Keyboard. A variant of the Triplet keyboard, this requires a five button mouse. This keyboard has six keys, with (up to) five characters per key (Figure 3, right).

Scrolling Techniques

Three of our text-entry techniques take advantage of the scroll wheel available on standard mice:

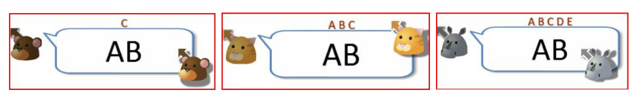

Figure 4. Scrolling techniques: Scroll (left), Triplet Scroll (middle), and Quintuplet Scroll (right). An icon indicating each Blank's owner appears to the left of each Blank, rendered as a "speech bubble." The text within the Blank ("AB") has already been entered and the user is scrolling through the alphabet (appearing above the Blank) to search for the letter "C".

Scroll. The user scrolls through the alphabet using the scroll wheel, one letter at a time (with the current letter displayed above the user's Blank for feedback) (Figure 4, left). With the cursor anywhere on screen, clicking on the left mouse button selects the letter currently displayed above that user's Blank. Successive scrolling begins with the last selected character. This is modeled on a mobile-phone text-entry technique wherein users scroll through letters with left and right arrow keys [8].

Triplet Scroll and *Quintuplet Scroll*. These allow scrolling through sets of three and five letters at a time, respectively (Figure 4, center and right). To select one of the letters appearing above a user's Blank, the user clicks the corresponding button on her three- or five-button mouse.

Collaborative Techniques

We also designed two techniques where students' actions are more closely coupled. Both techniques are based on the Alphabet Keyboard, though they could also be used in conjunction with any of the previous techniques:

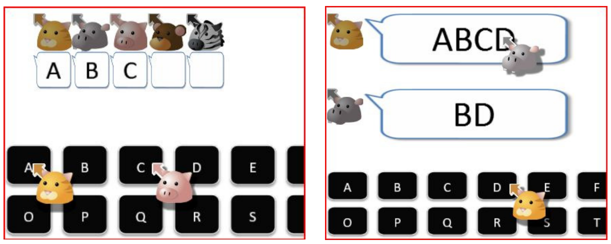

Figure 5. Collaborative text entry techniques: Collaborative Blanks (left) and Reuse (right).

Collaborative Blanks. Each user is assigned to one of a series of single-character-sized Blanks. Users then select letters using the Alphabet Keyboard to fill in their assigned Blank. For example, as shown in Figure 5 (left), to enter the string "ABCDEF", the user with the leftmost cursor would select A, the cursor second from left would select B, and so on. After any user enters a letter, a new single-character Blank appears assigned to that user. This technique could also be used on a word-level (rather than character-level) basis, and is designed to encourage students to communicate and work together to enter phrases.

Reuse. Users enter characters via any input method (we used the Alphabet Keyboard in our evaluation), but may also copy letters already entered by others. Letter reuse is achieved by clicking on any single letter on screen (as in Figure 5, where the "hippo" cursor copies the letter "D" from the "cat" cursor's Blank), or by clicking and dragging the cursor over a series of letters to copy multiple adjacent letters simultaneously. This technique could be used to copy letters from any on-screen text (such as information a teacher may display for lecture purposes).

Advanced Techniques

We also considered two more advanced text-entry techniques that might trade off steep learning curves for reduced screen space and eventual gains in speed. We provided on-screen legends, but once the codes are memorized, these legends would no longer be needed:

Morse Code. We mapped 'dots' to clicks of the left mouse button, and 'dashes' to clicks of the right mouse button, with a 400ms pause indicating completion. Pierpoint [22] suggests that typing speeds of 15-20wpm can be reached in a matter of weeks using Morse code with a telegraph key.

EdgeWrite [28]. EdgeWrite is a gesture-based technique where text is entered by traversing edges and diagonals of squares bounded by physical edges. We modified this to allow students to traverse squares with on-screen edges.

Screen Footprint per Technique

We used 20-point font characters for the on-screen keyboard techniques. Each key was 71x63 pixels. To accommodate the additional characters per key in the multi-letter keyboards, we made each key of the Multi-tap and Triplet Keyboards 1.5 times the size of the single character keys (107x95 pixels) and each key of the Quintuplet Keyboard 2 times the width of the single letter keys and 1.5 times the height (142x94 pixels). The Scrolling techniques used the space above each user's own Blank to display the scrolling characters. Therefore, the screen space used by these techniques depends on the number of users (*N*). Using 20-point letters, this amounts to 48x48pixels per letter. Table 1 summarizes the total screen footprint per technique.

Technique	Screen Space Used
GAG II	17.6%
Alphabet, Fish-Eye, Collaborative Blanks, Reuse	15.9%
Triplet Keyboard	12.8%
Multi-tap	11.5%
Quintuplet Keyboard	10.3%
Quintuplet Scroll	$1.5*N$%
Triplet Scroll	$0.8*N$%
Scroll	$0.3*N$%
Morse Code, EdgeWrite	0%

Table 1. Percentage of screen space used by each technique. *N* represents the number of users.

EVALUATIONS

We conducted a series of evaluations to determine the best multiple mouse text-entry technique with respect to the design considerations described earlier. Given the large number of techniques, we first conducted two preliminary evaluations to cull some before going on to a formal study of the five most promising techniques.

All our studies use the same single-display groupware system. While only some of the techniques are explicitly labeled as collaborative, all of them are impacted by multiple simultaneous users (for reasons such as occlusion) and therefore none could have been appropriately evaluated in a single-user scenario.

Informal Study 1 (CLT Computer Clubhouse)

Our participants were children who attend a "computer clubhouse" at the Children's Lovecastle Trust (CLT), a non-profit organization near Bangalore, India, that provides an after-school educational environment for primary and secondary school students from underserved communities.

Participation was voluntarily, and students were allowed to come and go as they pleased throughout. At any given time, 8 students aged 8-14 years old were participating. Most of them were able to read and write English at a first-grade (USA-system) level, although their spoken English was not as fluent. Their primary language was Kannada.

Our goal was to identify the most promising techniques. Over five days, we evaluated each of the five technique families. Participants used each technique for 15-20 minutes, except when participants wished to withdraw from using a particular technique. Each technique was explained or demonstrated with the help of a Kannada-speaking translator. The techniques were tested using two simple games we created in which students either enter letters presented on screen (the Enter Letter Game), or enter words corresponding to images of objects they had to identify (the Identify Picture Game). The images and words were taken from local first-grade English language workbooks. We allowed students to enter text freely, allowing for errors and error corrections. Only correct entries, however, advanced the game. To motivate students, points were given to the first student to complete each letter or word in a given level.

Each participant used a mouse to control a unique animal-based cursor, as in Figure 5. The room was set up with two rows of four mice each, placed on the floor of the clubhouse, facing a projector screen as in Figure 1.

Results

The Alphabet Keyboard and Reuse techniques were well received. Based on feedback and observations, we eliminated several of the other techniques due to usability problems. We eliminated the EdgeWrite technique because our removal of the physical edges, a key feature of the original EdgeWrite technique, made it unintuitive. We also eliminated both Quintuplet Keyboard and Quintuplet Scroll because these appeared physically harder for the students to use, since they did not use one finger per button as we anticipated, but rather used one to three fingers of one hand while often holding the mouse steady with the other. None of the students had experience with text-entry on mobile phones, and therefore had trouble with the pause-based selection required for both Multi-tap and Morse Code. However, we felt that this behavior could be learned, so we retained the simpler of the two, Multi-tap, and eliminated Morse Code. We also eliminated the GAG II Keyboard as it was the most difficult of the on-screen keyboards to learn.

Based on participant feedback, we modified the designs of some techniques. First, we changed the Alphabet keyboard to be two rows, instead of three, since students said they preferred the two row design of the Fish-Eye Keyboard. Second, we refined Collaborative Blanks such that new single-character sized Blanks would only appear after *every* participant finished entering a character in the first set of Blanks (originally, a new Blank would appear after *any* user entered a character, and that Blank would be assigned to that particular user). In the refined technique, once a user enters a letter, further letter selections simply replace the current letter in that user's Blank (instead of appearing in a new Blank), until all group members finish and a new set of Blanks appears.

Informal Study 2 (Christel House)

One goal of our second informal study was to gather additional data in order to select the best 5 techniques for formal evaluation. Alphabet Keyboard and Reuse were well received in Study 1, and hence we tagged them for inclusion in the formal study. The six remaining candidates (Fish-Eye, Multi-tap, Triplet Keyboard, Scroll, Triplet Scroll, and Collaborative Blanks) were thus re-evaluated to help us pick the best three.

Sixteen 8^{th} Standard students (8 female and 8 male, ranging in age from 13-15 years old) participated in the study. The students were from the Christel House School in Bangalore, India – a non-profit English-language school that caters to students from very poor surrounding communities.

We formed two gender-balanced groups of eight. Each group used three techniques each day, over two days (Multi-tap, Triplet Scroll and Triplet Keyboard for Group 1 and Fish-Eye, Scroll, and Collaborative Blanks for Group 2). The order of the techniques used in each group was reversed on the second day to reduce order effects.

We demonstrated each technique on the first day and then allowed the students to practice each technique for 2 minutes using the Enter Letter Game. They then played the Identify Picture Game for about 10 minutes on each day.

Each group was set up as in the previous study (Figure 1). We used numeric, instead of animal-based, cursor icons to improve identification. This enabled us to number the mice in each group from two to nine to correspond with the numeric cursor icons. We also tied mice to seating positions and had each student sit in the same position on the floor and use the same Blank on the screen each day. For techniques with an on-screen keyboard, the keyboard was always placed in the center of the screen (see Figure 1) so as to minimize differences in how far the students had to move their eyes between their own Blank and the keyboard.

After using each technique, students filled out a survey asking if they understood how to use the technique and how easy it was to use (on a 5-point Likert scale). At the end of each day, we also had the students compare the three techniques that they used by picking which they liked best and least and which were easiest and hardest.

Results

Our survey results showed that students favored Scroll and Triplet Scroll most (50% and 62.5%, respectively). Therefore, we decided to include both of these in our formal study, as well as the Triplet Keyboard which was also well received (37.5%).

We eliminated the Multi-tap Keyboard because several students appeared confused by the pause-to-select behavior, and 7 of 8 Group 1 students (87.5%) reported that this technique was the hardest to use.

As expected, due to the parallelization of input, students appeared to enter text much faster with Collaborative Blanks than with the other text entry techniques. However, this technique also seemed to result in more errors, likely due to multiple students contributing erroneous letters to a single word simultaneously. Collaborative Blanks did achieve the desired effect of instigating communication and coordination amongst students, although most of the communication was from students yelling at each other to enter in their assigned letter or enter it in correctly! Not surprisingly, then, only one student in Group 2 chose Collaborative Blanks as their favorite. Therefore, although this technique was fast, students found entering text this way cumbersome, so we eliminated it. Further research seems necessary to assess the benefits of this technique for group engagement, but is beyond the scope of this paper.

Fish-Eye was relatively fast and accurate to use, but the fluctuating key sizes were distracting to the students. With eight students using the technique simultaneously, many keys would continually increase and decrease in size, doing little to help direct them to any target key. Therefore, we also eliminated this technique.

Formal Study (Christel House)

After our second informal study, we were left with five text-entry techniques (Alphabet, Reuse, Triplet Keyboard, Scroll, and Triplet Scroll). We evaluated these techniques in a study monitoring student performance and preference for each technique over a seven day period.

Sixteen 7^{th} Standard Students (7 male and 9 female, ages 12–14 years) from the Christel House School participated. As in the second informal study, we divided the participants into two roughly gender-balanced groups of eight students each. The groups were set up in the same way as before (Figure 1). Each group used three techniques per day, over seven days (Alphabet Keyboard, Triplet Scroll, and Triplet Keyboard for Group 1, and Alphabet Keyboard, Scroll, and Reuse for Group 2). The Alphabet Keyboard served as the baseline for each group. We assigned each student a position on the floor and a corresponding numeric cursor to use on each day of the study.

On the first day, we demonstrated to each group their set of three text entry techniques, and allowed them to practice the techniques for approximately 10 minutes each using the Identify Picture Game. The next six days of the study were used to evaluate the performance and preference of each student for each technique over time. On each of these six days, we presented the text entry techniques according to one of the six possible permutations of the three techniques, to reduce ordering effects. At the end of each day, we administered a survey in which students picked which of the three techniques were their favorite and least favorite. At the end of the study, we distributed a final survey asking "If you had to use one of these techniques regularly, which would it be?" and "If you had to recommend one of these to other students to use in their classes, which would it be?"

Students used each technique for approximately 10 minutes on each day of the 7-day study. For each technique, the students played a game where each participant had to correctly enter phrases of text before the game advanced to the next level. We allowed for multiple spaces between words in each phrase because some students had a hard time seeing how many spaces they had entered between words. The phrases were taken from the MacKenzie and Soukoreff data set [9], excluding phrases that the school principal felt were culturally irrelevant, unfamiliar, or otherwise inappropriate for the students. We randomly divided the remaining 402 phrases into six sets of 67 phrases each, with an average phrase length per set of 28.3 characters. Every technique used each of the six phrase sets, and no two techniques used the same phrase set on any given day. To motivate students to enter phrases quickly, points were awarded to the first student to finish entering a phrase on every level.

Analysis

For all data analyses, we removed the data of one participant from Group 1 who could not attend the last day of the study and one participant from Group 2 who had to leave at various points throughout the study. We were left with data from seven participants per group. Also, although we attempted to give each technique 10 minutes of use each day, due to a power outage on one day and power surges on other days, we were only able to obtain ~48 minutes of data total for one of the techniques and therefore perform all of our analyses on the first ~48 minutes of data obtained from each technique.

To compare speed and accuracy across techniques, we analyzed our data using mixed-model analyses of variance with repeated measures because our experiment was a mixed between- and within-subjects factorial design (with participants in Group 1 using the Alphabet, Triplet Scroll, and Triplet Keyboard, and participants in Group 2 using Alphabet, Scroll, and Reuse). To simplify this analysis, we partitioned the data per technique into three bins (beginning, middle, and end), corresponding to the first, middle, and last two days of the study, respectively. Table 2 shows the amount of time partitioned into each bin for each technique. All of our models include *Technique* (Alphabet, Triplet Keyboard, Triplet Scroll, Scroll, Reuse), *Bin*, and their interaction as fixed effects and *Participant* (nested within *Group*) as a random effect to account for individual differences in performance. Note that mixed-model analyses can appropriately handle the imbalance in our data resulting from having both groups use the Alphabet Keyboard. We also performed Tukey HSD comparisons for post-hoc pairwise analyses.

	Bin			
Technique	**Beginning**	**Middle**	**End**	**Total**
Alphabet (Group 1)	21.54	16.14	10.17	47.85
Triplet Keyboard	17.75	16.36	13.32	47.43
Triplet Scroll	14.70	16.68	15.73	47.11
Alphabet (Group 2)	17.89	15.42	13.75	47.06
Scroll	17.90	16.45	12.57	46.92
Reuse	16.25	16.39	14.08	46.72

Table 2. Each technique's actual usage time (in minutes) partitioned to each 2-day temporal bin.

We measured speed in our mixed-model analyses using keystrokes per second (KSPS) [30] which captures the rate of text entry over the entire input stream (e.g., including deleted characters and deletions themselves) rather than words per minute which only measures the rate of entry over the final transcribed text. We also examined the nuances of some of the techniques and their effect on text entry speed, including the average number of scrolling actions taken to reach a target letter using Scroll and Triplet Scroll, and the number and length of characters reused over time for Reuse.

To predict text entry speed and analyze the learning rates per technique, we fit the time to enter a character (in seconds) per (unbinned) phrase to negatively accelerating power curves according to $y=cx^k$ [3].

To measure accuracy, we first computed the average minimum string distance (MSD) [25] across all the phrases entered throughout the study and across all techniques, in order to determine how much error we introduced into the study by allowing for multiple spaces between words. Next, the error rate was calculated as the percentage of all characters entered that were erroneous [30]. This measure appropriately excludes correct characters which were destructively deleted because we did not allow direct access to positions within the transcribed text.

Results

In total, participants entered 31,621 characters during this study (using all of the techniques). The second column in Table 3 lists the power curve equations that were fit to the time to enter a character per phrase. The third column shows the R^2 values, which represent the amount of variation in the time to enter a character that can be explained by its power relationship with phrase number. The last column reports the ANOVA results, which show how well the model can predict time per character speed. These results show that the text entry speed of participants using each technique improved significantly from the start of the study to the end ($p<.0001$ in all cases).

There was a significant effect of *Technique* ($F_{4,825.6}=31.69$, $p<.0001$), *Bin* ($F_{2,819.6}=106.93$, $p<.0001$), and *Technique*Bin* ($F_{8,819.7}=5.68$, $p<.0001$) on KSPS. Figure 6 shows the least squares means of each of the techniques within each bin, illustrating that KSPS increased over time for each technique. Table 4 and Table 5 report the least squares mean difference between each technique at the beginning and end of the study (statistically significant differences, $p<.05$, are starred). These indicate that Alphabet, Triplet Keyboard, and Reuse were significantly faster than both scrolling techniques (Triplet Scroll and Scroll) at the beginning and end of the study.

Technique	$y=cx^{-k}$	R^2	**ANOVA**
Alphabet	$y=2.6x^{-.15}$	.19	$F_{1,317}=73.19$, $p<.0001$
Triplet Scroll	$y=3.6x^{-.16}$	.31	$F_{1,139}=60.92$, $p<.0001$
Triplet Keyboard	$y=3.0x^{-.22}$	.40	$F_{1,180}=120.45$, $p<.0001$
Scroll	$y=4.5x^{-.10}$	.14	$F_{1,95}=16.07$, $p<.0001$
Reuse	$y=2.8x^{-.17}$	.18	$F_{1,102}=21.86$, $p<.0001$

Table 3. Power curve, R^2 value, and ANOVA results for each text entry technique

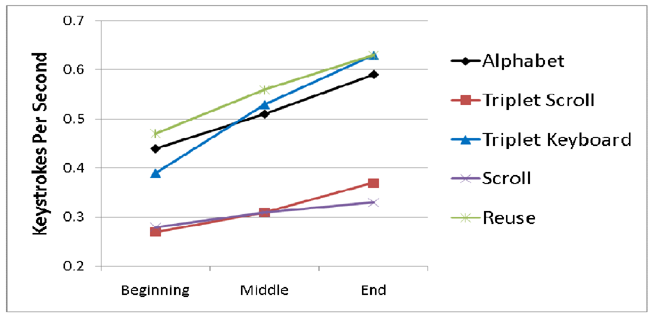

Figure 6. Keystrokes per second (KSPS) over three 2-day bins.

	Triplet Scroll	**Triplet Keyboard**	**Scroll**	**Reuse**
Alphabet	0.18*	0.05	0.16*	-0.02
Triplet Scroll		-0.12*	-0.02	-0.2*
Triplet Keyboard			0.11*	-0.08
Scroll				-0.18*

Table 4. Least squares mean difference in KSPS during the first two days of the study ('beginning'). Statistically significant differences ($p<.05$) are starred.

	Triplet Scroll	**Triplet Keyboard**	**Scroll**	**Reuse**
Alphabet	0.22*	-0.04	0.26*	-0.03
Triplet Scroll		-0.26*	0.04	-0.25*
Triplet Keyboard			0.3*	0.001
Scroll				-0.3*

Table 5. Least squares mean difference in KSPS during the last two days of the study ('end'). Statistically significant differences ($p<.05$) are starred.

The relatively slow text entry speeds for both Scroll and Triplet Scroll is likely because the number of scrolls taken to reach a target letter remained relatively constant for each of these from the beginning to end of the study (Table 6).

The speed of text entry with Reuse can be partially attributed to the fact that it was used in combination with the Alphabet Keyboard. Table 7 shows the average number of letters reused by participants from the beginning of the study to the end as well as the length of the text that they copied at one time. The large standard deviations here are likely because three participants reused letters very little or not at all, only relying on the Alphabet Keyboard to enter text, while some reused letters throughout the study. Interestingly, participants who reused letters less tended to reuse longer strings of text at time, while those who reused letters more tended to reuse shorter strings.

	Beginning	Middle	End
Scroll	24.36/3.59	24.26/7.42	25.84/6.79
Triplet Scroll	9.02/2.32	8.75/3.16	8.59/2.35

Table 6. Mean/Std number of scrolls taken to reach a target letter for each 2-day temporal bin.

	Beginning	Middle	End
Number of Letters Reused	5.77/12.29	9.31/13.00	5.02/8.06
Length of Text Reused	1.02/1.40	0.51/0.62	0.94/1.84

Table 7. Mean/Std number of letters reused and length of the text reused at one time, per 2-day temporal bin.

Our analysis of accuracy rates shows that the average MSD error rate across all participants and across all techniques was .02, meaning that allowing for multiple spaces between words only introduced an error of 2% into the data. Regarding the Error Rate (i.e., corrected errors excluding destructive deletions), we found that there was a significant effect of *Technique* ($F_{4,784.2}$=4.91, p<.0007) and *Bin* ($F_{2,820.7}$=12.19, p<.0001). There was no significant effect of *Technique*Bin* ($F_{8,820.7}$=1.08, p=.37) indicating that the Error Rate improvement over time was similar for all techniques. Figure 7 shows the least squares means Error Rate of each of the techniques within each bin indicating that both the Alphabet and Triplet Keyboard incurred a generally lower Error Rate than the Scroll, Triplet Scroll, and Reuse techniques (although the only significant difference in overall error rate was between Alphabet, Triplet Scroll, and Scroll, see Table 8).

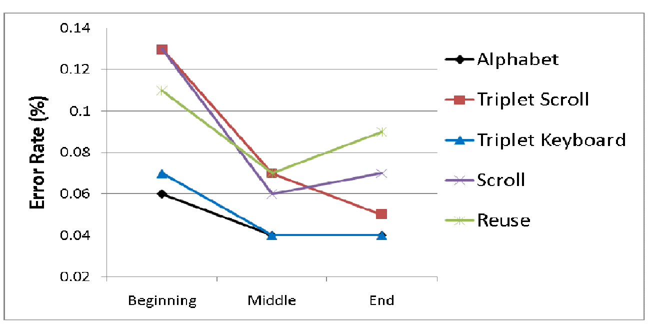

Figure 7. Error Rates over three 2-day bins.

Participants seemed to favor the scrolling techniques in both groups, choosing Scroll and Triplet Scroll as their favorite 52.8% and 47.6% of the time, respectively. In both groups, participants seemed to dislike the Alphabet Keyboard the most out of all of the techniques (choosing it as the technique they disliked most 46.2% and 63.4% of the time, respectively).

	Triplet Scroll	Triplet Keyboard	Scroll	Reuse
Alphabet	-.06*	-.003	-.07*	-.04
Triplet Scroll		.06	-.004	.02
Triplet Keyboard			-.06	-.04
Scroll				.02

Table 8. Overall least squares mean difference in Error Rate between techniques. Statistically significant differences (*p*<.05) are starred.

The final survey data showed that the majority of students in each group would choose to use Scroll and Triplet Scroll if they had to use one technique regularly in their classes (42.9% and 57.1%, respectively). Similarly, the majority of students would recommend the scrolling techniques to other students to use in their classes (57.1% for Scroll and 71.4% for Triplet Scroll). Table 9 summarizes the survey results.

	'Favorite'	'Least favorite'	'Would use'	'Would recommend'
Alphabet 1	21.4%	46.2%	28.6%	14.3%
Triplet Scroll	47.6%	33.3%	42.9%	71.4%
Triplet Keyboard	31.0%	20.5%	28.6%	14.3%
Alphabet 2	26.2%	63.4%	42.9%	28.6%
Scroll	52.8%	31.7%	57.1%	57.1%
Reuse	21.4%	4.9%	0%	14.3%

Table 9. Percentage of students choosing each technique to answer the corresponding survey question.

Discussion

Alphabet, Triplet Keyboard, and Reuse were comparably fast to learn and use after six days of practice and were all faster than the scrolling techniques. However, students demonstrated a relatively higher error rate using the Reuse technique compared to Alphabet and Triplet Keyboard, which both caused fewer errors than the scrolling techniques.

Interestingly, although the scrolling techniques (Scroll and Triplet Scroll) were significantly slower than the other techniques and produced more errors, participants seemed to favor these techniques the most. From our observations, this could be due to several factors. First, scrolling techniques present less of an occlusion problem because students can scroll and click anywhere on screen to enter letters. Second, the scrolling techniques require a focus on one part of the screen: the personal area above the blank. This made it easier to ignore other user activity. Third, all students tended to use either only their index finger or their index and middle fingers to click on buttons, while holding the mouse down with their other hand. This holding style is difficult for the non-scrolling techniques because they require targeted mouse movements.

Table 10 summarizes the tradeoffs between each design factor (excluding cost, as each of our techniques was based around a single mouse per student). Figure 8 visualizes these scores after normalizing them to make direct comparisons feasible and scaling them to show tradeoffs per design factor. To compute *Screen Space Remaining* we subtract the screen footprint of each of the techniques from 1 and use an *N* value of 15 for the number of students using Scroll and Triplet Scroll. *Scalability* indicates the rate of decrease of screen space remaining as the number of students increase (i.e., the space used by Alphabet, Reuse, and Triplet Keyboard do not vary with the number of students, but the space remaining with Scroll and Triplet Scroll decreases as the number of students increase). However, as the number of students increase, the amount of occlusion on the shared Alphabet, Reuse, and Triplet Keyboards also increases. *Multiple Users* is true for the Reuse Keyboard which allows students to copy letters entered by others and false for the other techniques. *Learning Rate* is the exponential value in the fitted learning curve for each technique. *Speed* and *Accuracy* are the overall least square means values for each technique for KSPS and Error Rate, respectively. *Preference* is the percentage of students indicating a technique as their favorite (averaging across Groups 1 and 2 for Alphabet).

	Scroll	Triplet Scroll	Triplet Keyboard	Reuse	Alphabet
Space Remaining	0.95	0.88	0.87	0.84	0.84
Scalability	-0.30	-0.80	0.00	0.00	0.00
Multiple Users	0.00	0.00	0.00	1.00	0.00
Learning Rate	0.10	0.16	0.22	0.17	0.15
Speed	0.28	0.27	0.39	0.47	0.44
Accuracy	0.87	0.87	0.93	0.89	0.94
Preference	0.53	0.48	0.31	0.21	0.24

Table 10. Summary of scores for each text-entry technique according to seven relevant design factors for multiple mouse text-entry in single-display groupware.

From Figure 8, for example, we can see that Scroll leaves the most screen space available for displaying actual content out of the five techniques. It slowly varies with the number of students making it relatively scalable. It has one of the slowest learning rates, has one of the slowest text entry speeds, is one of the least accurate text entry techniques, but was most preferred.

From these results, Triplet Keyboard appropriately balances most of our initial design criteria to be an ideal candidate for multiple-mouse text entry in single display groupware. It leaves more screen space available for content than Alphabet, is scalable because it does not vary with the number of students, has the fastest learning rate, has a relatively fast text entry speed, and is very accurate. However, the Triplet Keyboard was favored less than the scrolling techniques, but more than Alphabet and Reuse.

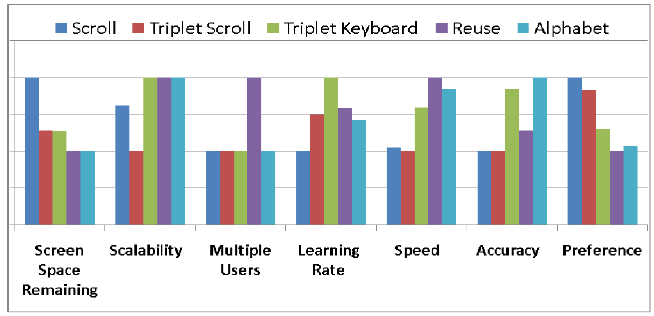

Figure 8. Normalized and scaled scores. Higher bars indicate larger values, measured as stated in the text for each factor.

Triplet Scroll could also be a potential candidate because it leaves almost the same screen real-estate available as the Triplet Keyboard (when used by 15 students), has a higher learning rate than Scroll, and is preferred by students despite being slow and quite inaccurate. The tradeoff between Triplet Keyboard and Triplet Scroll is that the former requires students to share the text entry area (leading to higher occlusion by others' cursors) whereas in the latter, students have control over their personal space for entering text. Furthermore, Triplet Scroll requires less gaze shifting than Triplet Keyboard because students only have to look above their Blank for letters to enter, making it ideal for in-context text entry activities such as diagram labeling and filling in the blanks. In contrast, Triplet Keyboard may be better suited for activities such as spelling out words spoken by a teacher, identifying images, or free-form question answering. However, since these techniques are similar in that they both require students to use three mouse buttons mapped to three character sequences, the skills learned for either one should transfer easily to the other. Therefore, it might be desirable to make both techniques available to teachers who could then choose which one to use based on the number of students participating or the activities being performed. Both Triplet Scroll and Triplet Keyboard could be also enabled concurrently, allowing students to enter text via the Triplet Keyboard or by scrolling through the Triplet Scroll alphabet. Note that Reuse, which can leverage input from other users, could be used in combination with either of these. Reuse might prove valuable by encouraging more engagement with other students.

CONCLUSION

The goal of this research was to identify and evaluate text-entry techniques for students using mice in single display groupware scenarios. After examining related work, we presented several design dimensions important to multiple mouse text entry in a classroom environment. We presented 13 multiple mouse text entry techniques in five categories: on-screen keyboards, multi-letter keyboards, scrolling, collaborative, and advanced techniques. We evaluated these in a three-phase study that showed the Triplet Keyboard balanced most of our design dimensions well, while the Triplet Scroll was most preferred by students. Therefore, we suggest that a combination of these two (optionally augmented with Reuse) can effectively enable mouse based text entry in single display groupware systems for developing region classrooms.

We have deployed a version of Mischief [11] with the Triplet Keyboard in combination with the Reuse technique in one rural Indian school. Early feedback indicates that students have been able to effectively learn and use this combination for actual text entry activities created by the teacher. Our future goals include examining the role of collaboration in text-entry and examining the pedagogical effectiveness of enabling text-entry in multiple mouse single-display groupware settings.

ACKNOWLEDGMENTS

We thank Jacob O. Wobbrock, James Fogarty, and Daniel Avrahami for their insights in many aspects of this work and their guidance on the analyses. We also thank the MSRI interns for their assistance during the formal study. Finally, we thank Bhagya Rangachar, Jaya George and the teachers and students at the Children's Lovecastles Trust and the Christel House in Karnataka for their support and participation in this research.

REFERENCES

1. Banks,D.A. (2006). Audience Response Systems in Higher Education: Applications and Cases. Information Science Publishing.
2. Beatty, I. (2004). Transforming Student Learning with Classroom Communication Systems. *EducCause Center for Applied Research (ECAR), Research Bulletin*, 3.
3. Card, S.K., Moran, T.P. and Newell, A.S. (1983) The power law of practice. In The Psychology of Human-Computer Interaction. Hillsdale, NJ: Lawrence Erlbaum, 57-65.
4. Fleetwood, M.D., and Fick, C.S. (2004) Input rates for a one-handed input device (OHAI) for Chinese text entry. Human Factors and Ergonomics Society Annual Meeting.
5. Gay, L. R. (1980). The Comparative Effects of Multiple-Choice versus Short-Answer Tests on Retention. *Journal of Educational Measurement 17* (1): 45-50
6. Inkpen, K., Ho-Ching, W., Keuderle, O., Scott, S.D., and Shoemaker, G.. (1999). This is fun! We're all best friends and we're all playing: supporting children's synchronous collaboration. *CSCL 1999*.
7. Isokoski, P., and MacKenzie, I.S. (2003). Metrics for Text Entry Research: Combined Model for Text Entry Rate Development. *CHI 2003,* 752-753.
8. MacKenzie, I.S. (2002). Mobile Text Entry Using Three Keys. *NordiCHI 2002,* 27-34.
9. MacKenzie, I.S. and Soukoreff, R.W. (2003). Phrase Sets for Evaluating Text Entry Techniques. *CHI 2003,* 754-755.
10. MacKenzie, I.S. Tanaka-Ishii, K. (2007). *Text Entry Systems: Mobility, Accessibility, Universality*. Morgan Kauffman.
11. Moraveji, N., Kim, T., Ge, J., Pawar, U.S., Mulcahy, K., and Inkpen, K. (2008). Mischief: Supporting Remote Teaching in Developing Regions. *CHI 2008*, 353-362.
12. Moraveji, N., Inkpen, K., Cutrell, E., Balakrishnan, R. (2009) A Mischief of Mice: Examining Children's Performance in Single Display Groupware Systems with 1 to 32 Mice. *CHI 2009,* 2157-2166.
13. Moraveji, N., Lindgren, R., Pea, R. (2009). Organized Mischief: Comparing Shared and Private Displays on a Collaborative Learning Task. *Extended abstracts of CSCL 2009*. Rhodes, Greece.
14. Negroponte, N. One Laptop Per Child. http://laptop.org.
15. Newell, A. and Rosenbloom, P.S. (1981). *Mechanisms of Skill Acquisition and the Law of Practice*. In Anderson J.R. (ed.) Cognitive Skills and their Acquisition. Erlbaum, Hillsdale, NJ.
16. Pal, J., Pawar, U.S., Brewer, E., and Toyama, K. (2006) The case for multi-user design for computer aided learning in developing regions. *WWW 2006,* 781-789
17. Pawar, U.S., Pal, J., Gupta. R., Toyama, K. (2007) Multiple Mice for Retention Tasks in Disadvantaged Schools. *CHI 2007,* 1581-1590.
18. Pawar, U.S., Pal, J., Uppala, S., and Toyama, K. (2006) Effective Educational Delivery in Rural Computer Aided Education: Multimouse. *DL 2006*.
19. Patra, R., Pal, J., Nedevschi, S., Plauche, M., and Pawar, U. (2007). Usage Models of Classroom Computing in Developing Regions. *ICTD 2007,* 158-167.
20. Pavlovych, A., Stuerzlinger, W. (2002) Model for Non-Expert Text Entry Speed on 12-Button Phone Keypads. *CHI 2002,*.
21. Pawar, U.S., Pal, J., Toyama, K. (2006). Multiple Mice for Computers in Education in Developing Countries. *ICTD*,.
22. Pierpoint, W. G. (2002). The Art and Skill of Radio Telegraphy: A Manual for Learning, Using, Mastering and Enjoying the International Morse Code as a Means of Communication. http://www.qsl.net/n9bor/n0hff.htm.
23. Raynal, M., and Vigouroux, N. (2005). Genetic Algorithm to Generate Optimized Soft Keyboard. *CHI* 2005, 1729-1732.
24. Singer, A. J., Murphy, M., S. Hines, M. (2003). Teaching to Learn, Learning to Teach: A Handbook for Secondary School Teachers. *Lawrence Erlbaum.*
25. Soukoreff, R.W., and MacKenzie, I.S. (2003). Metrics for Text Entry Research: An Evaluation of MSD and KSPC, and a New Unified Error Metric. *CHI 2003.* p. 113-120.
26. Stewart, J., Bederson, B. B., & Druin, (1999) A. Single display groupware: a model for co-present collaboration. *CHI 1999*,.
27. Swan, K., Kratcosi, A., Hooft, M.V., and Campbell, D. (2007). Technology Support for Whole Class Engagement. *Journal of Research Center for Educational Technology, 3*(1).
28. Wobbrock, J.O. and Myers, B.A. (2006). Trackball text entry for people with motor impairments. *CHI 2006*, 479-488.
29. Wilensky, U. and Stroup, W. (1999). Learning through Participatory Simulations: Network-Based Design for Systems Learning in Classrooms. *CSCL 1999.*
30. Wobbrock, J.O. (2007) Measures of text entry performance. Chapter 3 in Text Entry Systems: Mobility, Accessibility, Universality, I. S. MacKenzie and K. Tanaka-Ishii (eds). San Francisco: Morgan Kaufmann, pp. 47-74.
31. Wobbrock, J.O. and Myers, B.A. (2006) Analyzing the input stream for character-level errors in unconstrained text entry evaluations. *ACM Transactions on Computer-Human Interaction (TOCHI) 13* (4): 458-489.
32. Zhai, S., Hunter, M., Smith, B.A. (2002). Performance Optimization of Virtual Keyboards*. Human Computer Interaction 17*(3): 229-269.