
Sellen, A., Kurtenbach, G. & Buxton, W. (1992). The prevention of mode errors through sensory feedback. *Human Computer Interaction*, 7(2), 141-164.

THE PREVENTION OF MODE ERRORS THROUGH SENSORY FEEDBACK

Abigail J. Sellen, Gordon P. Kurtenbach, and [William A. S. Buxton](#)

Computer Systems Research Institute, University of Toronto

ABSTRACT

The use of different kinds of feedback in preventing *mode errors* was investigated. Two experiments examined the frequency of mode errors in a text editing task where a mode error was defined as an attempt to issue navigation commands while in insert mode, or an attempt to insert text while in command mode. In Experiment 1 the effectiveness of kinesthetic versus visual feedback was compared in four different conditions: the use of keyboard versus foot pedal for changing mode (kinesthetic feedback), crossed with the presence or absence of visual feedback to indicate mode. The results showed both kinesthetic and visual feedback to be effective in reducing mode errors. However, kinesthetic was more effective than visual feedback both in terms of reducing errors and in terms of reducing the cognitive load associated with mode changes. Experiment 2 tested the hypothesis that the superiority of this kinesthetic feedback was due to the fact that the foot pedal required subjects actively to maintain insert mode. The results confirmed that the use of a non-latching foot pedal for switching modes provided a more salient source of information on mode state than the use of a latching pedal. On the basis of these results we argue that user-maintained mode states prevent mode errors more effectively than system-maintained mode states.

1. INTRODUCTION

Mode errors as originally defined by Norman (1981) occur when a user misclassifies a situation resulting in actions which are appropriate for the analysis of the situation but inappropriate for the true situation. Mode errors in text editing are very common. Users may attempt to issue commands when the system is actually in "text insert mode" or attempt to enter text while actually in "command mode". While mode errors frequently

occur with computers, examples from diary studies of action slips (Norman, 1981; Reason and Mycielska, 1982; and Sellen, 1990) reveal mode errors occur in many other aspects of everyday experience. Examples such as trying to fast forward a videotape in the VCR when in "record mode", or turning the key in the ignition when the car engine is already running are also mode errors.

In the context of computers, the potential for mode errors exists when any given user's action can have very different effects depending on the state of the system. Fortunately the consequences of most mode errors are only minor inconveniences, and in well designed systems, are usually reversible. However, such errors in poorly designed interfaces or in highly complex systems such as aircraft and nuclear power plants can result in far more serious outcomes. In such situations, the importance of preventing such errors, or at least absorbing their effects, is critical.

Errors are not the only metric with which to measure users' problems with mode identification, however. In some cases, the user may diagnose the correct mode, but only after experiencing confusion or uncertainty. In such cases, the appropriate measure is one which reflects the cognitive effort or decision time required to deduce the system state. The amount of cognitive effort required to interact with a particular system may in turn be reflected in users' opinions of the usability of that system.

Why not just do away with modes? This was the opinion voiced strongly by Tesler (1981). But almost everything we do involves modes in one way or another, including working with so-called "modeless" computer systems such as the Apple Macintosh. Whenever dialog boxes appear, or whenever the cursor changes from an arrow to an "I-beam" depending on its location on the screen, one is in a mode. Similarly, selecting an object in the Macintosh Finder can be viewed as changing modes. When no objects are selected, typing will usually have little effect. When an object is selected, however, typing may result in renaming the object -- a common mode error in this interface. These examples serve to illustrate that what is actually meant by a "modeless" interface often refers to design in which contextual information is provided to minimize mode errors, and where modes can be easily entered and exited. There are other reasons to be concerned about the problem of managing modes. While the number of elemental actions available to interact with systems remains relatively constant, the number of functions within software applications is growing. One has only to look at applications such as Hypercard to see how more explicit modes are used to support rich functionality.

It is not clear that we can ever hope to completely eliminate the problems associated with modes, but it certainly seems possible to reduce them. One obvious solution seems to be to give users salient feedback¹ on system state. Apart from the practical importance for system designers, this raises some interesting theoretical questions: What kind of feedback is most salient to the user? Through which perceptual channel is the feedback best delivered? What are the design considerations that need to be taken into account in delivering mode information? One objective of our research is to shed some light on these issues.

There is little directly relevant literature. One exception is a study by Monk (1986) who investigated the use of auditory feedback in preventing mode errors. In this study, Monk demonstrated that mode errors could be reduced by a third by using a mode-contingent sound with each keystroke. Monk argued that sound is a good choice for system feedback in that users do not constantly look at the display while working.

There are many different design alternatives for providing mode feedback, however. For example, an alternative to Monk's method of presenting auditory mode information might be to use a sustained tone whose timbre (sound quality) depends on the current mode. In this case, we might predict that in contrast to the action-contingent sound used in the Monk experiment, subjects would be able to determine the current mode *before* initiating a possibly erroneous action. This kind of feedback might be called "proactive feedback" as opposed to the "reactive feedback" used by Monk and is an example of one dimension along which feedback may vary.

1.1. Characterizing Sensory Feedback

We can summarize some of the dimensions along which feedback can be characterized as follows:

* *Sensory modality of delivery (visual, auditory, kinesthetic)*

Through what sensory channel is the information delivered?

* *Reactive versus proactive feedback*

Does feedback occur only when an action is executed? Can one use the feedback to determine the mode before taking action?

* *Transient versus sustained delivery*

Is the feedback sustained throughout a particular mode?

* *Demanding versus avoidable feedback*

Can the user choose not to monitor the feedback?

* *User-maintained versus system-maintained feedback*

Does the user actively maintain the mode?

These dimensions are not all necessarily orthogonal. For example, feedback using the visual channel is generally avoidable: one can easily choose not to monitor visual information. Kinesthetic and audio feedback, however, are more demanding and inescapable by their very nature. What we hope to illustrate is that system designers face a variety of choices in providing mode information. In part this will be dependent on the task. All else being equal though, it seems reasonable that the more salient the feedback, the more effective it will be in preventing mode errors. Presumably feedback which is sustained, demanding, and actively-maintained is more salient than transient, avoidable, and passively-generated feedback.

2. EXPERIMENT 1: KINESTHETIC VERSUS VISUAL FEEDBACK

The first experiment was designed to test our beliefs that (1) sensory feedback to indicate mode can reduce both mode errors and the cognitive load imposed by confusion about

modes; and (2) that demanding, user-maintained feedback is a more effective way of preventing mode errors than avoidable, system-maintained feedback.

Buxton (1986) has previously argued that in articulating a command, continuity of physical motion and muscular tension can be used to provide effective feedback about the structure of the dialogue. An example of this is making a selection using a pop-up menu. Due to the continuity of the gesture used, the three steps of the transaction (Button Down: invoking the menu; Move mouse: make selection; Release Button: confirm selection) are perceived by the user as a single task. The sub-actions are "chunked" together by the continuity of motion and tension, and the overall transaction is bracketed on either side by a clear, neutral state of closure.

Our expectation, which we wanted to test in Experiment 1, was that binding through muscular tension would be effective in reinforcing knowledge about the current state over larger dialogue elements, such as modes. We chose to test this hypothesis in the context of a keyboard-based screen editor with a reputation for the number of mode errors it engenders --*vi*² (see Poller and Garter, 1984). Entering text in the modified editing environment we created requires the user to hold down a sustain pedal (from an electric piano). In order to navigate, the user releases the pedal and issues navigation commands. The expectation was that holding down the sustain pedal would bracket (or chunk) the text entry transaction much in the same way that holding down the mouse button provided the continuity of action in the pop-up menu example. Thus the active generation and maintenance of the mode state serves to bind the transaction and also to provide effective feedback on system state.

It was of interest to compare the effectiveness of kinesthetic feedback delivered via the foot pedal to visual feedback for indicating modes. In most interfaces, when mode indicators are provided at all, it is usually accomplished by changing some visual aspect of the interface (such as the cursor, for example). In this case, we maximized the saliency of the visual feedback by changing the entire screen color to a dark pink color while in edit mode. Even so, we did not expect this visual feedback to be as effective in reducing mode errors as the kinesthetic feedback because visual feedback is not only avoidable, but is passively generated by the system rather than generated by the user.

2.1. Method

Subjects. Twelve expert and twelve novice subjects were recruited from the University of Toronto and paid for their participation. An *expert* subject had extensive experience in using *vi*, a Unix-based text editing system. A *novice* subject was one who had never used *vi*, but had experience in using a computer mouse. Eleven of the experts and seven of the novices were touch typists.

Tasks. The primary task consisted of navigating through and inserting text into a pre-existing document on a Sun workstation. Subjects were instructed to insert the string "errorerror"³ following any word in the document that was printed all in capital letters. They were instructed to complete this task as quickly as possible, only correcting typing errors if they detected them within a word before leaving insert mode. Each block of text contained approximately 190 words and a total of 75 capitalized words.

A simulated *vi* text editor was created in which only a small subset of the commands were available. In order to navigate, the keys h, j, k, and l moved the cursor left, up, down, and right, respectively. In addition, the space bar was available to move the cursor right. For keyboard conditions, in order to insert text, subjects were instructed to position the cursor over the point at which the word was to be inserted, and to press the 'i' key. Once in "insert mode", the text could then be entered. After typing the text to be inserted, the escape key returned the user to "navigation mode". For foot pedal conditions, inserting text was accomplished by positioning the cursor over the insertion point, depressing the foot pedal, and keeping the pedal depressed while typing the text. Releasing the foot pedal returned the subject to navigation mode.

In addition to the primary task, subjects were also required to perform a concurrent distractor task on a Macintosh computer positioned adjacent to the Sun workstation. Thirty seconds after the editing task was begun, and after some random interval of time, beeps from the Macintosh signalled the presentation of a digit between 1 and 6 on its screen. Below the digit, six buttons numbered 1 to 6 appeared in a random order. The subjects' task was to use the Macintosh mouse to click on the button corresponding to the presented digit. Subjects were instructed to service this distractor task as quickly as possible. In order to encourage them to do so, the beeping would increase in frequency as time passed. The intervals between digit presentation were distributed according to a uniform distribution with an average interval between digits of 4.5 seconds and a range of 3 to 6 seconds.

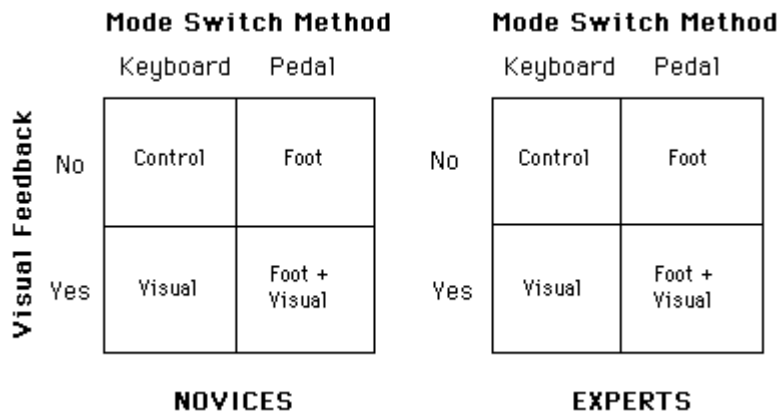


Figure 1: Schematic diagram of the experimental design. "Mode switch method" refers to the method of switching to insert mode, while "visual feedback" refers to the presence or absence of a dark pink screen color while in insert mode.

Design and Procedure. Each subject performed in each of the four conditions depicted in Figure 1. *Mode switch method* refers to the method by which insert mode was entered and exited. Keyboard mode switching means using the 'i' and 'escape' keys, while foot pedal mode switching means holding down the foot pedal to insert text. In the *visual feedback* conditions, while in insert mode, the screen changed from white to dark pink⁴. The order of the conditions for each subject was counterbalanced according to a digram-balanced Latin square.

All subjects were given a practice run on the editing task using the keyboard mode switch method immediately prior to performing the first keyboard condition, as well as a different practice run using the foot pedal mode switch method immediately preceding the first foot pedal condition. Each practice run consisted of 28 insertions into a pre-existing block of text.

At the end of the experiment, subjects were asked to rank order the conditions in terms of preference and to provide comments on the comparative usability of each "system" for text editing. The entire experiment lasted approximately an hour for expert subjects and an hour and a half for novices including a five to ten minute break halfway through.

2.2. Results

Mode Error Classification

Mode errors were operationally defined in the context of the task as follows, where <NAV> indicates switching to navigation mode and <INS> indicates switching to insert mode, by whichever method, foot pedal or keyboard:

A *navigation mode error* was defined as trying to navigate while in insert mode. Operationally, this meant the appearance of h, j, k, l, or spacebar characters while in insert mode and included any unexpected characters which could be construed as aiming errors around those keys, depending on the context. The presence of the "i" command when already in insert mode was also counted as a navigation mode error.

e.g. <INS>errorerrorllk<NAV>...

An *insertion mode error* was defined as trying to insert text while in navigation mode. This meant the appearance of any portion of the string "errorerror" while in navigation mode and also included anything which might be an aiming error around those keys.

e.g. lljjjjlerr<INS>errorerror...

There was some question as to whether the appearance of an additional "escape" character when in navigation mode in the keyboard mode switch conditions constituted a mode error. One could argue that such a response is due to uncertainty about the mode. However, normally in *vi* there is no cost (except a system beep and an extra keystroke) to making this "error". Many experts thus adopt the strategy of hitting escape to "make sure" they are in navigate mode. Indeed, in many of the examples where this was found to happen, experts hit "escape" after servicing the distractor. However novices also made similar errors from time to time, albeit infrequently. Since novices would not be as likely to have adopted this strategy, this suggests that some errors of this sort may, in fact, have been mode errors. Because of this questionable classification, the statistical analyses were run on the data both with and without the errors involving extra escape characters. The classification method which include these errors is henceforth referred to as using the "liberal" classification criterion, and the method which does not include them is referred to as using the "conservative" classification criterion.

In addition to mode errors, a class of errors we call *synchronization errors* occurred in the foot pedal conditions. A synchronization error looked very similar to a mode error in that a navigation command would sometimes precede the release of the foot pedal, or the letter "e" would sometimes precede depression of the pedal. It was clear, though, that these errors were different from mode errors in that the time between the erroneous

keystroke and the response of the pedal was very short (less than 200 msec). Thus these errors arose because of problems in synchronizing the action of the pedal with the keystrokes. Errors with times less than 200 msec. were therefore classified as synchronization errors.

Mode Errors

The mean number of mode errors using the liberal criterion is shown in Figure 2. Figure 3 shows the number of mode errors using the conservative criterion. While experts made more errors on average than did novices, this difference did not quite reach significance at the $\alpha = .05$ level using either criterion. Note that the choice of a liberal versus conservative classification only affected the means in the keyboard conditions, and that these effects were most pronounced in the case of experts.

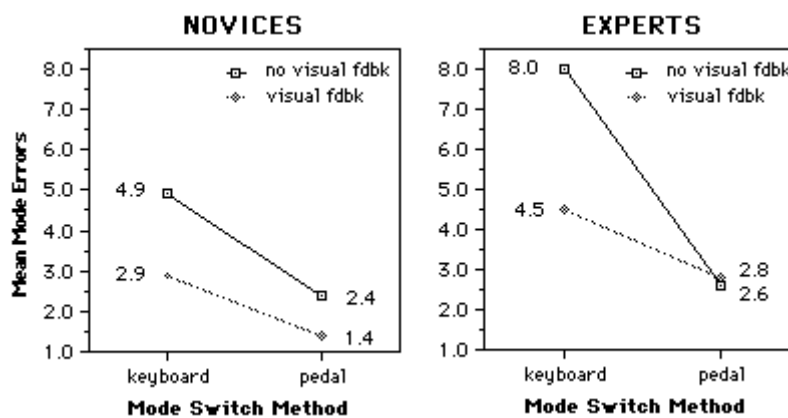


Figure 2: Liberal criterion. Mean number of mode errors for novices and experts plotted as a function of method of mode switch method (keyboard versus foot pedal) and visual feedback (present versus absent).

Both mode switch method and visual feedback affected the number of mode errors committed. For both the novices and the experts, the pedal method of mode switching resulted in significantly fewer mode errors than the keyboard using both criteria (liberal $F(1, 11) = 20.74, p < .001$; conservative $F(1,11) = 13.51, p < .001$). In addition, there were significantly fewer mode errors in conditions with visual feedback than those without for both novices and experts using both criteria (liberal $F(1, 11) = 11.40, p < .003$; conservative $F(1,11) = 8.45, p < .008$). Omega-squared tests were performed to assess the relative magnitude of the main effects. Mode switch method accounted for 15.6% of the variance using the liberal criterion, and 11.0% using the conservative criterion. Visual feedback, however, accounted for only 4.8% of the variance using the liberal criterion and 4.1% using the conservative criterion.

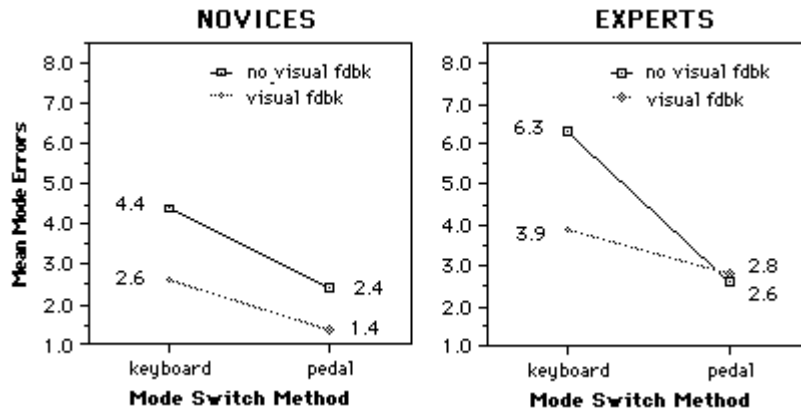


Figure 3: *Conservative criterion. Mean number of mode errors for novices and experts plotted as a function of method of mode switch method (keyboard versus foot pedal) and visual feedback (present versus absent).*

Finally, the effect of visual feedback depended on the mode switch method. Using both criteria, there was a significant interaction present between mode switch method and visual feedback (liberal $F(1, 11) = 9.56, p < .005$; conservative $F(1, 11) = 5.13, p < .03$). In order to understand the source of these interactions better, separate analyses were run on the expert and novice groups. The result was a significant mode switch method by visual feedback interaction for experts (liberal $F(1, 11) = 10.16, p < .009$; conservative, $F(1,11) = 4.67, p < .05$) but not for novices. This indicates that for experts, while visual feedback was effective in reducing mode errors when the method of mode switching was the keyboard, it was redundant in the case of the foot pedal.

Task Completion Time

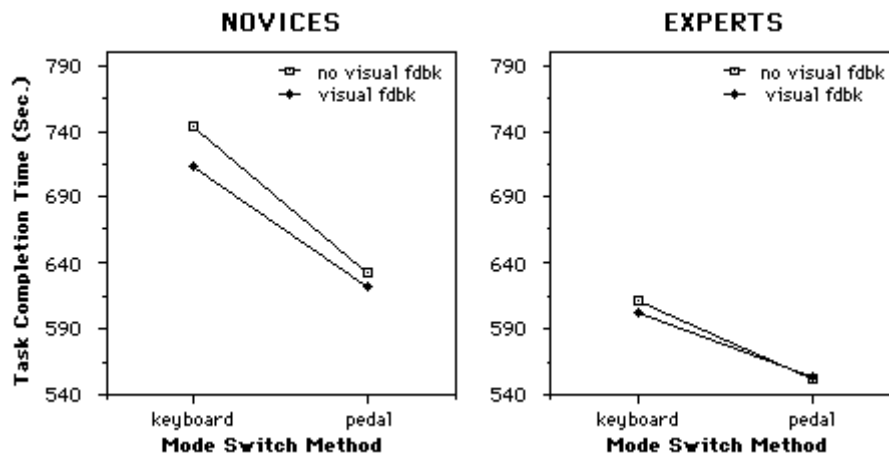


Figure 4: *Mean task completion times for novices and experts plotted as a function of method of mode switch (keyboard versus foot pedal) and visual feedback (present versus absent).*

The total time to complete the task in each condition is shown in Figure 4. Experts were significantly faster than novices ($F(1, 11) = 4.88, p < .038$). The only other significant result was a main effect of mode switch method, with the foot pedal being faster than the keyboard ($F(1, 11) = 25.85, p < .001$).

Effects of Switching Between Tasks

Resume time was defined as the total length of time between servicing the distractor task (clicking on a number on the Macintosh) and taking the first overt action in the editing task. In the keyboard conditions, this first action always consisted of a keystroke of some sort. In the foot pedal conditions, resuming the editing task sometimes began with a keystroke, and sometimes began by depressing or releasing the foot pedal. Resume time was taken to be a measure of confusion about the mode in the editing task, since choosing which action to take on resuming the main task of editing would tend to depend on assessing the mode first. The means are shown in Figure 5.

The pedal resulted in a significantly faster mean resume time than the keyboard ($F(1, 11) = 13.98, p < .001$). There were no significant effects of visual feedback, no differences between novices and experts, and no interactions found.

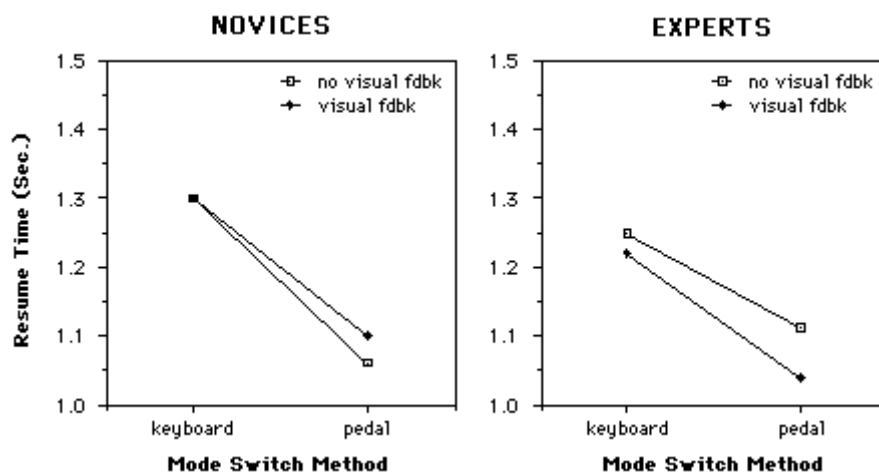


Figure 5: Mean "resume time" for novices and experts plotted as a function of method of mode switching (keyboard versus foot pedal) and visual feedback (present versus absent).

Service time for the distractor task was also examined. This was defined as the time between the occurrence of an audio interruption by the distractor task and the mouse click cancelling the number on the Macintosh screen. There were no significant differences found among conditions or between novices and experts.

Questionnaire Ranking Data

At the end of the experiment subjects were asked to imagine that each condition represented a system that they might use to do text editing on a daily basis and to rank order each of the four "systems" according to their preference. It was clear that subjects fell into three main categories: those who preferred the foot pedal, those who preferred

the keyboard, and those who preferred the "systems" with visual feedback, regardless of mode switch method. Operationally, they were classified as pedal-oriented, keyboard-oriented, or visual-oriented according to which conditions they chose as their first and second preferences versus their third and fourth choices.

One expert and one novice failed to complete the ranking task properly and could not be classified. Of the remaining eleven experts, five were keyboard-oriented, five were pedal-oriented, and one was visual-oriented. Of the eleven novices, eight preferred the foot pedal systems, two preferred the visual feedback systems, and one was unclassifiable. This last novice subject preferred either visual feedback, or the foot pedal, but not both.

2.3. Discussion

The results show the effectiveness of both visual and kinesthetic feedback in preventing the occurrence of mode errors. The benefits of visual and kinesthetic feedback were found regardless of whether or not subjects were experienced users of standard *vi*, a system with no explicit mode indicator. Thus, even though many of the expert subjects commented that they were used to keeping track of the mode "in their head", feedback of both kinds significantly reduced their mode errors nonetheless.

Kinesthetic versus Visual Feedback

While both kinds of feedback were beneficial, the results make a stronger case for kinesthetic over visual feedback for the prevention of mode errors in this particular task. The omega-squared tests showed that the mode switch factor accounted for approximately three times more variance than the visual feedback factor (11.0% to 15.6% versus 4.1% to 4.8%). In this experimental context, therefore, the magnitude of the foot pedal effect was greater than the magnitude of the visual feedback effect.⁵

The analysis of resume time also supports the conclusion that pedal feedback was more beneficial than visual feedback. It seems reasonable to assume that the amount of time required to resume the editing task in part reflected decision time during which subjects were attempting to diagnose the state of the system. Such cognitive processes are effortful and increase the mental workload of the task. Any differences in resume time among conditions must reflect a difference in cognitive operations since there are no differences in the physical actions required to return to the editing task after servicing the distractor.

Use of a foot pedal led to a significantly faster resume time than the keyboard while the presence of visual feedback made no difference. Further, these results are independent of level of skill, since novices and experts both benefitted from the foot pedal and not from visual feedback. The results therefore indicate that pedal feedback effectively reduced the cognitive load imposed by the system, at least with respect to confusion about system mode, while visual feedback did not.

Why might the foot pedal be a better way of reducing the cognitive load imposed by confusion about modes than visual feedback? Both methods of delivering feedback can be described as sustained and proactive -- they are both present throughout the mode state

and do not depend on the execution of an action before providing feedback. But there are some major differences between them which can be enumerated:

- 1) *User-maintained versus system-maintained feedback.* The kinesthetic feedback in this context required that subjects actively maintained the mode by holding down the foot pedal. In the visual case, the feedback was maintained by the system. Thus in this condition subjects received the mode information passively.
- 2) *Kinesthetic versus visual sensory channels.* Visual feedback is inherently more avoidable than kinesthetic feedback. Subjects in this task could have chosen not to attend to the visual cues provided. However, kinesthetic feedback is more difficult to ignore, even though the possibility of habituation to this kind of sensory feedback exists, especially over longer periods of time.
- 3) *Relationship between mode switching and feedback.* In the case of the foot pedal, the effector for articulating mode switching was also the limb through which sensory feedback on mode status was received. This was not the case for visual feedback where the effectors for mode switching (the fingers) were distinct from the sensory channel receiving the feedback.
- 4) *Competition for visual attention.* The visual feedback may have competed with the visual nature of the editing task. It may be the case that searching the screen or monitoring the outcome of one's keystrokes during text-editing means that fewer attentional resources are available to allocate to monitoring the color of the screen. Thus it may be that using a different "channel" for receiving sensory feedback about mode is more effective since it does not compete with task-specific resources.
- 5) *Distribution of tasks over limbs.* Another possibility is that the foot pedal may have effectively prevented mode errors due to the fact that mode switching and maintenance of mode was allocated to a limb independent of the effectors used for the main task (namely, the fingers). In the visual feedback condition, the task of mode switching was accomplished via the fingers, potentially interfering with the primary typing task. It is conceivable (but we think improbable) that assigning the task of mode switching to a non-interfering limb provides the subject with a more effective memory of the mode.

The alternatives listed above point out some interesting and sometimes subtle distinctions between different kinds of feedback. They also serve to emphasize the need to consider the relationship between the type of feedback used and the nature of the cognitive and physical constraints of the primary task. However, of most interest to us is the user-maintained versus system-maintained distinction because it is closely related to the notions of chunking and phrasing discussed by Buxton (1986). The results are consistent with the theory that user-generated muscular tension not only provides continuity of physical motion to bind transactions but also provides effective feedback on mode state. Conventional visual feedback is system-maintained and thus does not confer these advantages. Experiment 2, discussed later, provides a more direct investigation of the active/passive maintenance distinction.

Other Issues of Usability

We found other interesting behavioral differences among conditions. One clear difference was the shorter time in which the task was completed for foot pedal versus keyboard conditions. Somewhat surprising was the fact that this was true not only for novices but

also for experts, most of whom had many years of experience with standard keyboard mode switching in *vi*.

The fact that keyboard mode switching caused more mode errors and therefore may have incurred more cost in terms of error recovery time probably contributed to the task time difference. Another contributing factor doubtless was the fact that assigning the task of mode switching to the foot pedal meant that this could be accomplished without physically interfering with the other tasks of navigating and typing. Both subjects who could and could not touch type commented that having to alternate between "i" and "escape" and the navigation keys meant having to constantly re-position the fingers on the keyboard. Many of them felt that this led to more errors in typing. Many of them also said that using "i" and "escape" meant they had to spend more time searching the keyboard, which made the task more effortful. In the keyboard case, controlling the mode was accomplished through two different physical controls ("i" and "escape") whereas mode switching with the foot pedal involved the same control to enter and exit edit mode. Whatever the various contributing factors might have been, subjects (both novices and experts) commented that they liked the fact that editing with the foot pedal seemed much faster.

There were different problems associated with the foot pedal. Most notably, from time to time there were synchronization errors where subjects would either depress or release the pedal a fraction of a second too early or too late. These were fairly infrequent though, averaging less than one error per subject during the entire experiment. In addition, some subjects commented that they thought that eventually their foot would become tired (although none stated that their foot actually *was* tired). One wonders how much of a problem this might be with good ergonomic design, however. By analogy, holding down the accelerator pedal while driving tends not to be tiring except for extremely long distances.

Finally, we had expected that there would be some differences in service time and in "chunking" behavior across conditions. Chunking behavior refers to the tendency to finish one sub-task before attending to another (Buxton, 1986). In this case, chunking was defined as the tendency to delay servicing the distractor task until completion of a sub-task within the editing task. For example, subjects had a strong tendency to complete navigation to the next word, or to complete typing of the inserted word before attending to the distractor. They did this even though they were instructed to service the distractor task as quickly as possible. We predicted that with improved feedback subjects might feel secure enough to interrupt their primary task (text editing) mid-stream, in order to service the distractor. This was not the case, however, and perhaps speaks to the strength of the tendency to chunk in all conditions.

Experts versus Novices

There was some question as to whether the experts in this study were truly "experts" since subjects could use only a restricted set of commands in the editing task. However differences between experts and novices suggest that they were in fact drawn from distinct populations.

First, experts completed the task much faster than novices in *all* conditions. Note that this was the case even though they were as naive as novices with regard to the foot pedal.

This suggests that experts had no trouble integrating the new device with their previously established skills in *vi*.

Second, experts exhibited a different pattern of behavior with regard to mode errors. It is interesting to note that experts made more mode errors on average than novices, contrary to what might be predicted. Also, unlike the novices, experts did not benefit from visual feedback in combination with the foot pedal. This is somewhat surprising given that all but one of the experts were touch typists and frequently monitored the screen. Conversely the beneficial effect of visual feedback for the novices independent of mode switch method was also surprising given that five of the twelve were not touch typists and constantly monitored the keyboard. One might therefore expect that visual cues would be less effective for this group. This may be explained by the fact that we frequently observed novices making deliberate visual checks to ascertain the mode when returning from the distractor task. It could be that experts were more likely to be looking at the screen but not necessarily for the purpose of making a visual check on the mode. Looking at the screen does not necessarily indicate that subjects were attending to the visual mode indicator.

One explanation for the differences in error behavior across groups is that for experts there is less overhead involved in correcting errors. Expert users of *vi* make errors all the time, and are highly skilled at recovering from them. Because the cost of an error for an expert is considerably lower than the cost of an error for a novice, experts can afford to commit them more often. Novices, on the other hand, must be considerably more cautious and therefore tend to make fewer errors. Increased cautiousness on the part of novices may also account for why they benefitted from visual feedback given that they were already receiving feedback from the foot pedal. Novices might be motivated to exploit every available cue in an effort to avoid errors. The error results taken together suggest that while avoiding mode errors does not necessarily become easier with experience, recovering from them does.

Finally, there was a preponderance of mode errors consisting of an additional "escape" character when in navigation mode for experts in the keyboard conditions. This was not the case for novices, nor for the experts in the foot pedal conditions. Hitting an extra escape character is a way to make sure one is in navigate mode which normally incurs no cost to the user. It emphasizes the fact that that experts clearly had built up this error avoidance strategy specific to the keyboard interface in order to cope with the tendency to make mode errors when no mode feedback is available.

3. EXPERIMENT TWO: SUSTAINED VERSUS LATCHING FOOT PEDALS

A second experiment was designed to provide a clearer answer to the question: Why was the foot pedal in the first experiment such an effective means of delivering mode information? We wished to argue that it was not the fact that the foot was used *per se*, but that it was the active generation and maintenance of the mode state through the foot pedal which reduced the cognitive load imposed by mode switching. Thus, we predicted that if

the foot was used to accomplish mode switching, but was not used to actively sustain the mode state (as in a latching foot pedal), then errors and the cognitive load due to mode switching would be equivalent to the keyboard condition. Furthermore, we predicted that in comparison to the use of a sustained foot pedal, both a latching foot pedal and keyboard would be inferior.

3.1. Method

Subjects. Fifteen expert users of *vi* were recruited from the University of Toronto and paid for their participation. Eight of the experts were touch typists.

Tasks. Both the primary, editing task and the distractor task were identical to the tasks used in Experiment 1. In addition, the keyboard condition and foot pedal condition (henceforth referred to as the "sustained pedal") were identical to the conditions in Experiment 1 without visual feedback. The difference was that a latching foot pedal was also introduced as a method of changing modes. In this condition, inserting text was accomplished by positioning the cursor over the insertion point, and depressing and releasing the foot pedal once. Depressing and releasing the foot pedal again returned the subject to navigation mode.

Design and Procedure. Each subject performed in each of the three conditions outlined above: keyboard, sustained pedal, and latching pedal. The order of conditions for each subject was counterbalanced according to a Latin square. Note that in this experiment all subjects were experts, and no visual feedback was used.

All other aspects of the procedure were identical to the procedure used in Experiment 1: Each subject received a training period on the device about to be used prior to performing each condition. At the end of the experiment, subjects filled out a questionnaire ranking the three conditions and providing comments on their usability. This experiment was somewhat shorter than the previous one, lasting less than an hour.

3.2. Results

Mode Errors

Mode errors were operationally defined using the same criteria described in Experiment 1. Again two different sets of criteria were used: in the conservative classification, errors involving extra "escape" characters were not included. Similarly, synchronization errors were defined as before (as errors in which the time between erroneous keystroke and mode switching was less than 200 msec.). Synchronization errors occurred both in the sustained and latching pedal conditions. The total frequency of synchronization errors in the sustained pedal condition was 18 (an average of just over 1 per person) and in the latching case the total frequency was 2.

The mean number of mode errors in the three conditions is shown in Figure 6. The results from the experts in Experiment 1 (no visual feedback) are also shown for comparison purposes. The difference between conditions is highly significant (liberal $F(2,28) = 25.8$, $p < .0001$; conservative, $F(2, 28) = 17.50$, $p < .0001$). Using the *liberal* criterion, planned comparisons of the difference between means showed the keyboard condition to yield

significantly more mode errors than the latching pedal ($p < .007$), and the latching pedal to produce significantly more errors than the sustained pedal ($p < .0002$). Using the *conservative* criterion, however, planned comparisons revealed that while the sustained pedal produced significantly fewer errors than both the keyboard and latching pedal conditions ($p < .0001$), there was no difference between the keyboard and the latching conditions ($p < .33$).

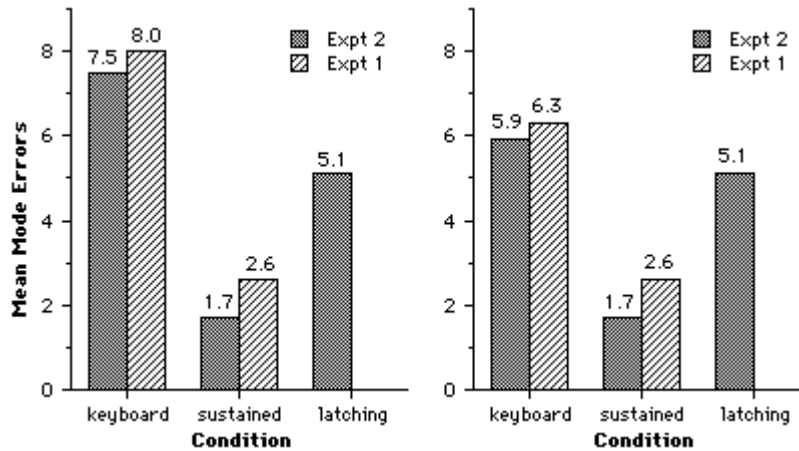


Figure 6: Mean frequency of mode errors in the three conditions showing the results using the liberal criterion on the left and conservative classification on the right. Results from experts in Experiment 1 are also shown.

Task Completion Time

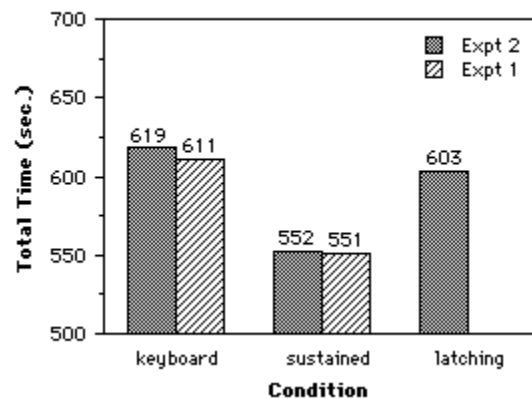


Figure 7: Mean task completion time for the three conditions in Experiment 2. Task completion times for the identical conditions for the experts in Experiment 1 are also shown.

Total time to complete the task in each condition is shown in Figure 7, again showing the times of the comparable conditions to those of the experts in Experiment 1 (no visual feedback). The difference among conditions is significant ($F(2,28) = 6.6, p < .004$).

Planned comparisons of the difference between means reveal that the sustained pedal yielded faster tasks completion times than both the latching pedal ($p < .013$) and keyboard conditions ($p < .002$). However, the times for the keyboard versus latching pedal were not significantly different from each other ($p < .42$).

Effects of Switching Between Tasks

Mean *resume times* are shown in Figure 8. Unlike Experiment 1, differences among conditions do not reach significance in this experiment ($F(2,28) = 1.89, p < .17$) even though the absolute magnitude of the difference between keyboard and sustained pedal conditions in Experiment 1 is similar to that found in Experiment 2. This result is due to larger variability in Experiment 2. In addition, resume times in general are shorter than those produced by the subjects in Experiment 1.

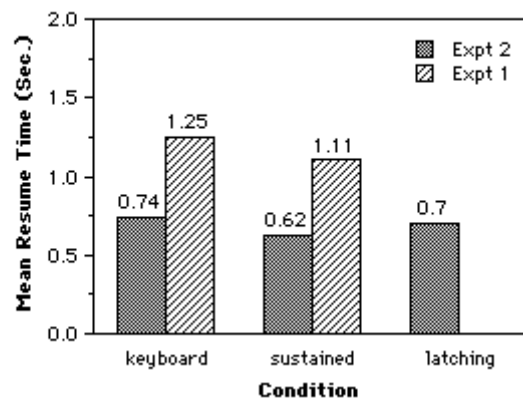


Figure 8: Mean time taken to return to the editing task (resume time) after servicing the distractor task. Results from experts in Experiment 1 shown for purposes of comparison.

As in the first experiment, there were no significant differences in *service time* among conditions ($p < .715$). Contrary to the findings for resume time, service times in general were *longer* than those of the experts in Experiment 1.

Questionnaire Ranking Data

As in Experiment 1, subjects were asked to rank each configuration (keyboard, sustained and latching foot pedal), according to their preference as a feature within a text editing system they might use on a daily basis. The rankings indicated a strong preference for the sustained foot pedal and consistent dislike of the latching foot pedal. Eleven of the fifteen of the subjects preferred the sustained pedal over either the latching pedal or keyboard. Of those eleven, eight preferred the keyboard and three preferred the latching pedal as second choice. Three subjects ranked the keyboard highest and sustained pedal second. Finally one subject preferred either foot pedal giving precedence to the latching.

3.3. Discussion

Our hypothesis that the keyboard and the latching pedal would produce equally error-prone behavior was borne out, using the conservative criterion. The latching pedal was less error-prone than the keyboard using the liberal criterion, but the fact that the liberal criterion includes errors due to extra escape characters calls into question the validity of this classification. Extra escape "errors" in both experiments were committed almost exclusively by experts, almost always in the keyboard conditions. Thus, they are most probably the result of error avoidance habits rather than true mode errors.

Furthermore, as we predicted, neither the keyboard nor the latching pedal matched the performance of the sustained pedal. This finding, in combination with the finding of equivalence between keyboard and latching pedal, argues for the effectiveness of user-maintained feedback in preventing mode errors. Experiment 2 makes clear the important point that it was not the use of a foot pedal *per se* that resulted in the performance improvements with respect to mode errors, but the fact that actively-sustained feedback was available to the subjects.

Using the foot, however, has the added benefit that mode switching can be performed and sustained feedback can be maintained without interfering with the primary hand-dominated keyboard task. This probably contributed to the short task completion time for the sustained foot pedal. Other contributing factors may include the fact that the sustained foot pedal caused fewer errors and hence less time was needed for error recovery. In addition, only one overt action was involved in switching modes (depressing *or* releasing the pedal) whereas the latching pedal involved two (depressing *and* releasing the pedal). The keyboard method involved only one action (hitting "i" or "escape") but these two controls were distinct. As in Experiment 1, subjects mentioned that searching the keyboard for "i" and "escape" and having to reposition their hands on the home row in order to resume typing probably slowed them down. The fact that both the latching pedal and the keyboard conditions yielded longer task completion times was likely due to some combination of these factors (errors, number of mode switch controls, and physical interference with the main task).

A further aspect of usability was revealed by subjects comments and rankings of the three devices. Most of the users liked the sustained pedal best, and all but one ranked it first or second in order of preference. The most common complaint was that the latching foot pedal was very tiring. Subjects often kept their foot hovering above the pedal in anticipation of switching modes.

Finally, in this experiment the kind of criterion used to classify the mode errors significantly affected the results. The fact that experts in the keyboard conditions in both experiments often hit an extra escape character to "make sure" of the mode they were in shows that they develop strategies to attempt to compensate for their uncertainty. Note that no such comparable strategy was possible in the latching foot pedal case because making an extra foot pedal response would result in the undesired effect of switching modes. On the other hand, given the properties of the sustained foot pedal, the need for developing such strategies is greatly reduced, if not eliminated.

4. DESIGN IMPLICATIONS AND CONCLUSIONS

These experiments not only make the case for providing feedback on mode state to the user, but also point out two important design considerations. One consideration has to do with the nature of the feedback provided. We have attempted to show that actively-maintained feedback is more effective than passively-maintained feedback in preventing mode errors. This kind of feedback is most naturally provided in the kinesthetic domain. It is conceivable that actively-maintained mode states could be provided through other channels such as visual or auditory channels although the scenarios one can imagine usually have a kinesthetic or at least proprioceptive component. For example, humming to maintain a mode state or looking in a particular place to maintain a mode have proprioceptive components to the feedback. The other side of the coin is to attempt to provide kinesthetic feedback passively (for example, by applying pressure to the arm in a particular state) and to see whether this is also an effective form of feedback. This potential experiment might help distinguish between the active-passive dimension and the dimension of sensory channel.

The other design consideration concerns minimizing the interference between mode switching and feedback, and task performance. With word processing tasks such as the one used in these experiments, where users are seated at a desk and fairly static, the sustained foot pedal is a viable design solution to this problem. However, a similar solution would not necessarily work when word processing using a portable computer in an airplane, for example. Furthermore, it is conceivable that there may be cases where one may want to distinguish among more than two modes. A design challenge motivated by the above experiments, therefore, is to explore other methods of providing the feedback to similar effect. First, other approaches to creating actively-sustained feedback need to be explored. Second, other forms of feedback, such as the use of sound, deserve more attention than they have received up until now.

While the use of visual feedback is obviously important, the channel is over-used, in our opinion, when compared to other modalities. Besides the issue of channel overload, our experiments suggest that information delivered through the visual channel is simply not as *salient* as information delivered kinesthetically. This may be the case even though the visual cues in the first experiment involved changing the entire screen area pink. This has important implications for systems which rely on more subtle visual cues such as changing the shape of the cursor or the color of the menu bar.

An underlying motivation of these experiments was to test some of the concepts introduced by Buxton's (1986) paper on chunking and phrasing. One of the principle points of that paper is that the nature of the initiation or termination of a transaction is not as important in structuring a task as the nature of the feedback maintained during a transaction. In the experiments reported here, the purpose of this feedback or muscular tension was to maintain a sense of task mode or system state awareness. The data from these experiments support Buxton's suggestion that user-maintained feedback through pressure is an effective mechanism for binding transactions. Pressure is one of two chunking mechanisms that Buxton suggested. The second is that of continuity of motion, what he calls "kinesthetic continuity" and which should more properly be called "kinetic continuity". Both kinetic continuity and pressure share the property of user-maintained active feedback. Consequently the results of the experiment suggest that this is worth a more detailed exploration, having implications for stylus-driven, line-drive interfaces and gesture-based systems.

How might we go about incorporating kinesthetic, user-maintained feedback into human-computer interfaces? To some extent, kinesthetic feedback is already being used successfully to convey mode information in existing interfaces. Examples of this include both pull-down and pop-up menus. Holding down the shift key when typing characters, and holding down the mouse button while dragging icons or sweeping out areas on the screen such as window boundaries are other examples where it is used. In retrospect, the findings of these experiments suggest that carrying out these kinds of operations using latching mechanisms, (such as a Caps lock key), will tend to produce more mode errors. Our own everyday experience suggests that this is indeed the case.

In prospect, the results can be applied by first performing a task analysis. If we look at the above examples, they can all be characterized by an A-B-A structure where A is the primary task, and B is a temporary or interim task. One alternative is to assign the B task to a different effector such as the other hand (see Buxton and Myers, 1986; Buxton, 1990). Another is to use a non-interfering effector to maintain a temporary mode state for task B. These interim tasks are prime candidates for situations where user-generated, kinesthetic feedback can be applied since they tend to be frequent but temporary. Because they are temporary, the fatigue factor does not become an issue. Because they are frequent, it becomes all the more important to prevent mode errors from occurring.

4.1. Conclusions

To conclude, the research reported makes three main points:

- That by providing sensory feedback, a common class of error (mode errors) can be significantly reduced for both novices and experts. Both visual and kinesthetic feedback can significantly improve performance.
- That the modality (sensory channel) used for feedback, and the nature of that feedback are important design considerations. While we do not wish to discourage the use of visual feedback where appropriate, these experiments have pointed out the benefits of other kinds of feedback. Kinesthetic feedback has the important property that it tends to be actively-maintained and thus is especially effective in preventing mode errors.
- That expanding the range of input devices and effectors used for interacting with computers can lead to many improvements in system usability. In particular, input methods which do not interfere with the primary task and which can be carried on concurrently with the primary task can lead to substantial improvements in terms of task completion time and reducing the cognitive load imposed on the user.

As the complexity and functionality of systems grow, we must learn to anticipate the errors users will make and to design interfaces to minimize their occurrence. In order to cope with this growing responsibility, we feel strongly that interface design will be served well by looking beyond the traditional "mouse-keyboard-display" configuration and investigating other channels and modalities of interaction. We believe that the work of Monk (1986) and the research reported in this paper support this view, and hope that it will stimulate additional research and activity in this direction.

Acknowledgements. The members of the Input Research Group at the University of Toronto provided the forum for the design and execution of this project. We especially thank Gary Hardock and Scott MacKenzie for comments and assistance. We also thank Steve Draper and Don Norman for comments and advice on the first experiment and an earlier version of this paper. In addition, we are grateful to our colleagues at the University of Toronto: Daniel Read and Ian Spence of the Department of Psychology for their advice in statistical matters, and Alison Lee and George Drettakis of the Dynamic Graphics Project for implementation advice.

Support. We gratefully acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada, Digital Equipment Corporation, Apple Computer and Xerox PARC.

REFERENCES

- Buxton, W. (1986). Chunking and phrasing and the design of human-computer dialogues. In H. J. Kugler (Ed.) Information Processing '86, Proceedings of the IFIP 10th World Computer Congress (pp. 475-480). Amsterdam: North Holland Publishers.
- Buxton, W. (1990). The "natural" language of interaction: A perspective on non-verbal dialogues. In B. Laurel (Ed.), The Art of Human-Computer Interface Design (pp. 405-416). Reading, MA: Addison Wesley.
- Buxton W. and Myers B. (1986). A study in two-handed input. Proceedings of the CHI '86 Conference on Human Factors in Computing Systems, 321-326. New York: ACM.
- Keppel, G. (1982). Design and analysis: A researcher's handbook. (Second edition.) Englewood Cliffs, N. J.: Prentice-Hall Inc.
- Monk, A. (1986). Mode errors: A user-centred analysis and some preventative measures using keying-contingent sound. International Journal of Man-Machine Studies, 24, 313-327.
- Norman, D. A. (1981). Categorization of action slips. Psychology Review, 88 (1), 1-15.
- Poller, M.F., & Garter, S.K. (1984). The effects of modes on text editing by experienced editor users. Human Factors, 26(4), 449-462.
- Reason, J., & Mycielska, K. (1982). Absent-minded? The psychology of mental lapses and everyday errors. Englewood Cliffs, N. J.: Prentice-Hall Inc.
- Sellen, A.J., Kurtenbach, G.P., and Buxton, W.A.S. (1990). The role of visual and kinesthetic feedback in the prevention of mode errors. Proceedings of the Third IFIP Conference on Human-Computer Interaction, (INTERACT), 667-673. Cambridge, England.
- Sellen, A. J. (1990). Mechanisms of human error and human error detection. Unpublished doctoral dissertation. University of California, San Diego, La Jolla, CA.
- Tesler, L. (1981). The Smalltalk environment. Byte, August, pp. 90-147.

FOOTNOTES

- ¹We define "feedback" in the general sense as information about system state received through any of the human sensory channels.
- ²Joy, W. and Horton, M. (1986). An introduction to editing with Vi. In "UNIX User's Supplementary Documents (USD)." Berkeley, CA: Computer Systems Research Group, University of California, USD.
- ³The string 'errorerror' was chosen so that mode errors consisting of attempts to insert these characters in navigation mode could be clearly distinguished from navigation commands or aiming errors around the navigation keys, and similarly for errors in attempting to navigate in insertion mode.
- ⁴The use of the pink color was intended to be consistent with the convention of "red for record mode" which is prevalent in many other kinds of artifacts such as VCR's. In addition, note that the technique of holding down a foot pedal to record is used in other kinds of systems such as dictaphones.
- ⁵As statisticians (e.g., Keppel, 1982) have pointed out, comparing the relative strength of two factors should be treated with caution since the measures in part depend on the choice of levels of each factor, and thus are partly under the experimenter's control. In this case, we feel we suitably maximized the difference between the condition with no visual feedback and the condition with visual feedback in order to permit some degree of generalization. However, it could still be argued that the visual feedback would have been more salient had it been flashing, for example. We agree that these kinds of design considerations may improve the quality of the feedback but that they may also run the risk of becoming too distracting.