Interruptions: A Criterion in the Design and Evaluation of Human-Computer Interfaces

by

Gillian Ginsburg

A thesis submitted in conformity with the requirements

for the degree of Master of Applied Science

Department of Mechanical and Industrial Engineering

Institute of Biomaterials and Biomedical Engineering

University of Toronto

Interruptions: A Criterion in the Design and Evaluation of Human-Computer Interfaces

Master of Applied Science

Gillian Ginsburg

Department of Mechanical and Industrial Engineering

Institute of Biomaterials and Biomedical Engineering

University of Toronto

2004

ABSTRACT

The main objectives of this study were to assess how interruptions affect human-computer interaction and to use interruptions to conduct representative evaluations of interfaces, the results of which can be more generalizable to the actual environments in which the interfaces are used. In order to achieve these objectives, an extensive literature review was first performed. Following this, a field study was conducted in a hospital recovery room to determine what types of interruptions nurses encounter as they care for patients. Simulated interruptions were then designed to match the type of those observed, and incorporated into representative experiments to evaluate interfaces for two infusion devices. Interruptions disrupted participants while they programmed both interfaces, although the interfaces designed by human factors principles seemed to lessen the disruptive effects of the interruptions in some cases. Limitations are discussed and recommendations are made for further improving the interfaces of both devices.

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisor, Dr. Kim Vicente, for his ongoing support, guidance, expertise, and kindness throughout this project. Many thanks are also extended to Dr. Greg Jamieson, and all members of the Cognitive Engineering Laboratory. Pavan Dadlani did a wonderful job in programming the Visual Basic simulations, and in creating the p-player for data collection.

I would also like to extend thanks and appreciation to Paula Ferguson, Patient Care Coordinator of the Post-Anesthetic Care Unit of the Toronto General Hospital, for helping me to set up the observation sessions for the field study and recruit participants for the field study and the experiments. The experienced nurses and nursing students who participated in the study are thanked for their time, effort, and helpful suggestions. I am very grateful for financial support provided by NSERC, OGS, and a grant from the University of Toronto Connaught Fund.

Above all, I would like to sincerely thank my wonderful husband, Marc Ginsburg, for his ongoing love, support, sense of humour and compassion, and his nursing expertise and insightful comments and suggestions. My incredible parents, Sarah and Haron Hillel, have always been supportive, helpful and encouraging. I am also grateful to Abby and Adam Dwosh and Jon Hillel for their love and support. Finally, thanks to my loving friends and to Smiley and Snowy for providing ongoing stress-relief.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

INTRODUCTION

When designing and evaluating human-computer interfaces, it is very important to adhere to human factors principles, to ensure that systems are safe, easy to use, and efficient (ANSI/AAMI, 2001). One essential human factors principle is to minimize the load on the user's memory (Lin, Isla, Doniz, Harkness, Vicente, and Doyle, 1998; Lin, Vicente, and Doyle, 2001; Baker, 2000). This feature of an interface is especially important in systems that are operated under complex working conditions, where a user's attention may be devoted to multiple tasks at once, and/or where a user is interrupted from performing a task and then resumes the task at a later time. An example of a complex work environment is the nursing workplace in a hospital. Interruptions occur frequently, drawing a nurse's attention away from a task, such as programming an infusion device for medication administration. If the device interface is poorly designed, putting an excessive load on a user's memory, then interruptions may result in programming errors that can severely compromise patient safety.

The purpose of this study is four-fold: (1) to explore the effects of interruptions on performance in human-computer interaction, (2) to explore the impact of interruptions on medical error and patient safety in the realm of nursing, (3) to gain insights into how to conduct interruptions experiments to evaluate interfaces representatively, and (4) to determine if medical errors can be reduced by designing human-computer interfaces for medical devices to help users deal with interruptions and reduce their disruptive effects. As a test-bed, redesigned interfaces for two patient-controlled analgesia pumps were used in the experiments.

GENERAL INTERRUPTIONS RESEARCH

The purpose of this section is to investigate general research that has been conducted in the field of interruptions and human performance, to gain insights into how to design and conduct interruptions experiments, and into why interruptions are disruptive. Much work has been done in this area. This work mainly consists of: (1) the development of a taxonomy for human interruption, (2) the conducting of experiments to determine the effects of interruptions on user performance of ongoing tasks, and (3) the conducting of experiments to determine the factors that influence the disruptive effects of interruptions.

A Taxonomy of Interruptions

A few researchers have attempted to define interruptions and establish a taxonomy that describes the different issues surrounding interruptions. This section outlines their findings. A general definition/classification of interruptions may make it easier to generalize research findings across different work domains.

McFarlane (1997, 1998, 1999) studied how and when computers doing automated tasks for people can interrupt users from other tasks when feedback is required. He defined human interruption as "the process of coordinating abrupt change in people's activities" (McFarlane 1997, p. 67). This abrupt change can involve a change in cognition, perception, or physical action. A more in-depth definition was developed by Latorella (1996): an interruption is an additional task that competes for a limited resource and redefines what is currently in active memory.

McFarlane (1997, 1998, 1999) developed a taxonomy of human interruption, as a tool for answering interruptions research questions. The taxonomy, summarized in Table 1, lists eight dimensions of human interruption. Manipulating each dimension, as discussed in the next section, can influence the disruptive effects of interruptions.

Table 1. Summary of McFarlane's (1997) taxonomy of human interruption.

Dimension of Interruption	**Example of Dimension**
Source of interruption	Self; another person; computer
Individual characteristic of person receiving interruption	Limitations of: perceptual processors, cognitive processors, motor processors, memory, focus of consciousness, processing streams; willingness and ability to be interrupted
Method of coordination	Immediate (the person must leave current task to attend to interruption); negotiated (the interruption is announced to the person, and then the person decides when to attend to it); mediated (the interruption is announced to the person's personal digital assistant (or another third party), which determines when the best time is to interrupt the person); scheduled (the person is interrupted during prearranged times only)
Meaning of Interruption	Alert; stop; divert attention (task-switching); distribute attention (task-sharing); remind; communicate information
Method of expression	Physical (i.e., verbal); type (i.e., by purpose)
Channel of conveyance	Face-to-face; mediated by a person; mediated by a machine
Human activity changed by interruption	Conscious or subconscious; individual activities; joint activities
Effect of interruption	Change in activity; change in memory; change in awareness; change in focus of attention; loss of control over activity

Figure 1 illustrates the features of a generic interruption that requires task-switching (adapted from Trafton, Altmann, Brock, and Mintz, 2003). First, a person is alerted to an interruption. There may be an interruption lag in which the person tries to come to a logical break in the primary task before switching tasks. This lag may not be substantial in safety critical environments in which immediate attention to the interruption is necessary. The person's attention then switches to the interruption task (which might be determined in experiments when the person makes the first action/decision on the interruption task). Following completion of the interruption task, there is usually a reorientation time in which the person tries to remember where he/she left off in the primary task.

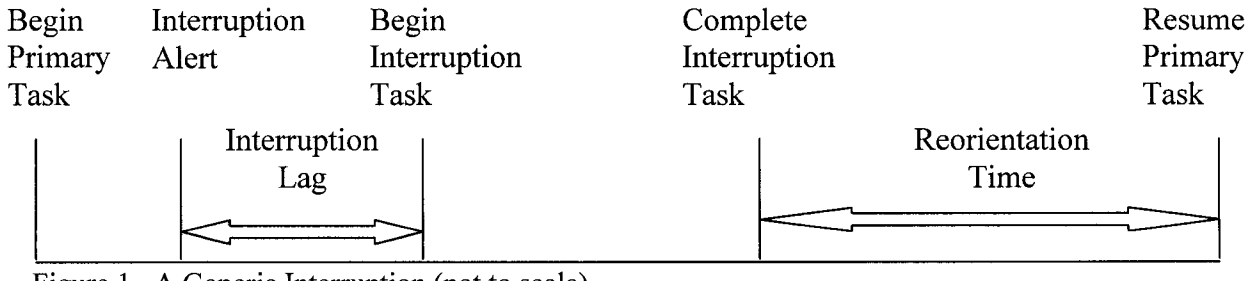


Figure 1. A Generic Interruption (not to scale).

Interruptions Experiments with Interfaces

Several researchers have performed interruptions experiments, usually to evaluate an interface and/or to determine the best time to interrupt a person. Several experiments are reviewed here. The methods used for interrupting users and the types of interruptions presented to users are focused on here, to gain insights into how to conduct interruptions experiments.

Cutrell, Czerwinski, and Horvitz (2001) conducted an interruptions experiment using database searches. Participants were asked to search through a list of book titles to find a particular book, after being given either the word title of the book (little cognitive demand required for this task), or the gist of the book (higher cognitive demand required for this task). Participants were allowed to ask for a reminder if the title or gist were forgotten. Two search methods were employed: one with a marker highlighting the title in the list that the user was looking at, and one with no marker. A simple math problem was presented as the interruption by an instant message notification, at a pseudo-random point during the search task, depending on the participant's place in the list. The main findings were that:

- performance was significantly slower in interrupted trials than non-interrupted trials,
- it took longer for participants to switch to the interruption task in gist trials than title trials, indicating that they were trying to create mental cues as to their place in the book list, and
- reminders were requested significantly more often in interrupted gist trials than non-interrupted gist trials, in gist trials than title trials, and in trials where the participant was interrupted earlier on in the search task.

Being interrupted early on in a task was even more detrimental to performance, as the task had been in the participants' short-term memory for only a minimal amount of time. Interestingly, the presence of a marker did not benefit participants in interrupted trials. This is

likely due to the fact that it provided no memory assistance in terms of the task that the participant was performing (i.e., which book to search for).

In an earlier study, Field (1987) investigated the effects of two types of interruptions on a database navigation task, where participants had to use the database to find answers to a set of questions. One interruption required participants to complete a numeric sequence, and the other was to search through a group of texts to find the title of a book. Both interruption tasks were presented after a predetermined sequence of screens in the primary task had been completed (pseudo-random timing). In the continuous task, participants were allowed to either return to any previously selected screen or return only to the last screen viewed. Field found that the interruptions had an effect on task completion. When participants were allowed to return to any previous screen, they performed better in terms of active search time, and the disruptive effects of interruptions were lessened as they were more certain of their place within the database.

A unique realistic interruptions study investigated the extent to which interruptions disrupt a pilot's activity on a simulated flightdeck (Latorella, 1996), a work domain in which disruptive effects of interruptions can prove fatal. Interruptions were air traffic control clearances that were systematically inserted into the various tasks that the participants had to perform. Interruptions increased post-interruption performance error rates. Interestingly, interruptions seemed to slightly speed up performance time, suggesting that participants adopted a compensatory strategy to work faster after an interruption, knowing that they had time constraints. This in turn could explain increased error rates after an interruption. It would be important to perform additional experiments that could determine whether errors resulted from compensatory strategies or from disruptions to a participant's memory, since other studies have shown, in contrast, that interruptions increase performance times of primary tasks, as well as error rates (Eyrolle and Cellier, 1992, 2000; McFarlane, 1999; Field,1987). Burmistrov and Leonova (2003) maintain that interruptions may not affect performance time on simple tasks, and

Zijlstra, Roe, Leonora, and, Krediet (1999) believe that people can actually over-compensate the potential performance decline, thereby performing faster on the primary task and sometimes even maintaining the same level of quality, although usually at the expense of greater psychological costs. Trafton et al. (2003) observed that participants who were forced to attend to an interruption immediately were able to adapt and eventually improved their ability to resume the primary task to the extent that they resumed as quickly as participants who were given an interruption lag.

One of the earliest studies on interruptions compared the effects of interruptions on two different calculator interfaces (Kreifeld and McCarthy, 1981). One-minute interruptions requiring participants to write down multiplication tables were presented during the regular calculation tasks. Interestingly, no significant differences were found between the two interfaces during uninterrupted trials, yet during interrupted trials, one interface resulted in longer task completion times than the other interface. The reason that one interface resulted in poorer performance could be due to the fact that it placed more memory stress on the user, as well as having an unconventional layout of digits. More errors were also committed in interrupted trials with this interface, although not significantly.

Several studies have investigated the influence of interruption task similarity, length, and complexity on the disruptive effects of interruptions (Eyrolle and Cellier, 1992, 2000; Gillie and Broadbent, 1989; Edwards and Gronlund, 1998; Latorella, 1996; Storch, 1992; Bailey, Konstan and Carlis, 2000). A study by Gillie and Broadbent (1989) used a game task that involved memorizing a list of items and then moving to locations in the game that would supply those items. Various interruptions were presented, to study the effects of interruption length, similarity to the main task, and complexity on performance of the main task. Accordingly, interruptions included:

- 30 seconds of simple mental arithmetic,
- 2.75 minutes of simple mental arithmetic,
- free recall, where participants repeated each word out loud as it was presented by the computer (similar to the primary task), with no delay between the start of the interruption and the first word to prevent participants from rehearsing their place in the main task, and
- decoding letters to numbers to perform a simple mental arithmetic problem, with participants being allowed to decide when to perform the interruption task.

The main findings were as follows: the length of an interruption and the point at which the interruption was performed did not significantly influence the disruptive effects of an interruption, and the similarity of the interruption to the primary task and the complexity of the interruption in terms of the cognitive demands it required did influence disruptiveness. The finding of interruption timing contradicts that of McFarlane (1999), described next.

McFarlane (1999) used a game task (catching falling cartoon characters by moving stretcher bearers) and an interruption graphical matching task (requiring a short focus of attention) to simulate each method of coordination that is described in his taxonomy (see Table 1). The graphic nature of the matching task was chosen to correspond to the graphical nature of the primary game task, for task similarity. He used all four methods of interruption coordination in this experiment:

- immediate (the matching task appeared, completely obscuring the view of the game task, and required completion before the game task reappeared),
- negotiated (the interruption was announced with a few brief flashes and then the game resumed, and participants decided when to begin the matching task),
- mediated (the participant's mental workload was dynamically calculated, and interruptions were presented when the workload was low), and

- scheduled (interruptions were presented at a prearranged schedule of once every 25 seconds).

It was found that performance was affected by the method used for coordinating interruptions, but there was no one best method for all performance measures. For example, the immediate method showed the worst performance in terms of accuracy, but the best performance in terms of completeness, on the interruption-matching task. The negotiated method showed the best performance in terms of accuracy on the continuous task, whereas the pre-scheduled method showed the worst. Mediation did not appear to significantly improve performance for any measure, although this may be indicative of the type of task performed.

Other researchers have also studied the timing of interruptions, and how a warning can allow a person to anticipate an interruption (Nagata, 2003; Hodgetts and Jones, 2003; Franke, Daniels and McFarlane, 2002; Miller, 2002; Trafton et al., 2003; Monk, Boehm-Davis and Trafton, 2002; Diez, Boehm-Davis and Holt, 2002; Horvitz, Jacobs, and Hovel, 1999). Warnings essentially create an interruption lag (see Figure 1), and results of these studies have shown that an interruption lag can reduce the disruptive effects of interruptions, primarily by reducing reorientation time to the primary task after the interruption task is completed and thereby reducing overall performance time of the primary task. Interruption lags in these studies allowed participants to either finish what they were working on before attending to the interruption, or encode retrieval cues to allow for better task resumption following the interruption. Most of these studies have focused on computerized work, where an automated computerized system must intermittently interrupt a user for input, while the user is focused on other tasks. However, it is important to note that in safety critical environments, such as a hospital, it may not be possible for health care workers to anticipate interruptions and have a substantial interruption lag.

Cognitive Effects of Interruptions

Numerous studies, including the ones outlined in the previous section, have investigated the disruptive effects of interruptions. It is well known that interruptions affect behaviour. These effects are related to limitations in a person's cognitive abilities to work during interruptions. Although people can execute several cognitive processes at once, their performance of a thought or action with complete control and consciousness is limited to only one at a time (McFarlane, 1997, 1998, 1999). People can, however, attempt to divide consciousness between multiple processing streams to perform multiple tasks (McFarlane, 1997). According to Oulasvirta and Saariluoma (2004), in the case of task-switching, information must be saved into long-term working memory before the switch occurs, if the interrupted task is to be resumed at a later time. They believe that interruptions requiring immediate attention can disrupt this transfer of information from short-term working memory to long-term working memory (also termed semantic elaboration), and hence have detrimental effects on performance of the interrupted task once it is resumed. They also believe that task similarity is more detrimental than when the interrupting and primary tasks are dissimilar because semantic similarity of the interrupting task retroactively interferes with the retrieval of information from long-term working memory when the primary task is resumed. These explanations support the idea that both interruption lag and retrieval cues should help reorientation to the primary task. In fact, researchers have shown that in the case that interruptions increase task completion time, it is the reorientation time that is responsible for this overall increase (Burmistrov and Leonova, 1997, 2003). Obviously, allowing individuals to come to a logical cognitive break in their primary task before having to attend to an interruption would decrease the disruptive effects of the interruption as well as the anxiety and annoyance they experience (Bailey et al., 2001), but in safety critical environments such as the nursing workplace, this would likely not be possible.

Even if users are allowed to rehearse their position in the main task before attending to an interruption, however, interruptions, even if very short, are still disruptive (Gillie and Broadbent, 1989). Thus, it is imperative for interfaces to reduce the negative effect of interruptions, and to decrease the load on the user's memory. Also, interruptions that occur early in a task (such as searching through a list) have a larger effect on a user forgetting a primary task goal than interruptions that occur later on (Cutrell et al., 2001). This may be because the primary task goal will have been in the user's memory for a shorter amount of time. In addition, the effects of interruptions are influenced by training and expertise, but training in the primary task without interruptions does not reduce disruptive effects when interruptions are actually experienced (Cutrell et al., 2001). Thus, users who are familiar with interruptive workplaces will develop coping strategies.

In developing a theory of how people remember their goals or the states of the world they want to achieve, Altmann and Trafton (2002) theorized that events during the "interruption lag", defined as the time between the onset of the alert to the interruption (i.e., the phone ringing) and the onset of the interruption itself (i.e., the conversation), are critical to the ability of a person to resume a goal after an interruption. This ability to resume a goal depends on mental or environmental cues to the goal that are developed during the interruption lag and are present at the resumption of the task. Furthermore, these cues must be of a "means-ends" nature, that is, be obvious and prime the memory of the person for the goal and not for other interfering tasks or distractions. The better a person can remember how far he/she has progressed toward achieving a goal, the more likely he/she is to accurately and efficiently resume the task, without committing potentially harmful errors. Edwards and Gronlund (1998) also found that people need associative connections between task components that can result in a mental representation of the task, to facilitate memory recall of the position in the task after an interruption occurs. Altmann and Trafton (2002) recommend that operators be taught how to search for appropriate cues and

associate them with the goal that is being interrupted. However, this active searching for cues before attending to an interruption would likely place further cognitive demands on the operator. Thus, it would be more beneficial to design user interfaces that readily provide these cues.

There are two main classes of interruptions: those involving task-switching (where users are required to leave a primary task and resume it after an interruption), and task-sharing, or multitasking (where users perform multiple tasks at once) (Eyrolle and Cellier, 1992). Eyrolle and Cellier (1992) found that more errors, including intrusions, confusions, and omissions, occurred when tasks were switched. This is likely due to short term memory loss of the primary task, as opposed to task-sharing where the primary task is still being performed, yet may take longer to complete. Pawlak and Vicente (1996) used verbal and spatial secondary tasks to determine which type of cognitive processes were utilized by the primary task. For example, if the task utilized spatial resources, then performance would likely decrease with the addition of a secondary spatial task. This would seem likely, as interruption similarity to the primary task influences the cognitive disruptive effects (Eyrolle and Cellier, 1992, 2000; Edwards and Gronlund, 1998). This implies that interfaces could be designed to utilize only one of these resources, and thus free-up others to better handle interruptions. The findings from this study would also be important in designing interruptions experiments for interface evaluation, where the simulated interruptions must accurately represent those encountered in the actual workplace for the results to be generalizable. If it is known what cognitive resources are utilized by the interruptions encountered by actual interface users, it may be possible to design simulated interruptions that utilize the same resources, and thus have the same cognitive effects. Results from a laboratory setting can be better generalized to the actual work environment in which devices are used if test conditions represent the realistic environment (Kaye and Crowley, 2002).

One such work environment in which these results may be applicable and useful is the nursing workplace, a safety critical environment in which interruptions are a major concern. To our knowledge, no one has previously looked at whether or not human-computer interface design can mitigate the disruptive effects of interruptions in the nursing workplace. The next section presents a literature review that was conducted to determine what is known about the interruptions that nurses face while they care for patients.

INTERRUPTIONS IN THE NURSING WORKPLACE

This section presents a literature review of the research that has been conducted about nurses' working conditions. Studies and their findings into the effects of interruptions on nurses are first described, followed by studies and their findings of the types and frequency of interruptions encountered by nurses.

Nurses' Perceptions of Interruptions

According to the National Academy of Sciences, human medical error accounts for 44,000-98,000 preventable deaths per year in hospitals in the United States (Kohn, Corrigan, and Donaldson, 2000). Medication administration is a primary role of nurses, and can occupy nearly one-third of their time (Wakefield B.J., Wakefield D.S., Uden-Holman, and Blegen, 1998; Gladstone, 1995; Segatore, Miller, and Webber, 1994). However, medication errors, defined as events that could have led to, or did lead to, an undesirable outcome, such as increased hospital stay, permanent disability, or death (Cooper, Newbower, Long, and McPeek, 1978), are prevalent and a leading threat to patient safety (Lin et al., 1998, 2001; Gladstone, 1995). Segatore et al. (1994) cited a study that estimated that the medication error rate in hospitals is one error per patient per day. One type of error is an error in dosage, where a patient receives a dose greater or less than a predetermined amount, such as that ordered by a physician. It has been estimated that 13% to 18% of all medications administered are of the wrong dosage, and that medication errors are vastly underreported (Walters, 1992). In Gladstone's (1995) survey of incident reports, over 50% of errors were dose-related, and most commonly of an incorrect infusion rate (17.7%). The intravenous route was involved in 32.9% of errors, and infusion devices were involved in 50% of these incidents. One study (Vicente, Kada-Bekhaled, Hillel, Cassano, and Orser, 2003) found that 65-667 deaths may have occurred in the United States from 1988-2000 due to programming errors associated with a single type of intravenous patient-

controlled analgesia pump. Other researchers have also looked at programming errors associated with infusion pumps (Flynn, Mohr, and Lawlor-Klean, 2003; ECRI, 2002). Medication errors can also include: wrong route, wrong rate, omission, incorrect time, mistaken patient, and incorrect drug (O'Shea, 1999; Gladstone, 1995).

Numerous studies have been conducted to determine the causes of medication errors (Segatore et al., 1994; Davis, 1990, 1994; Gladstone, 1995; Walters, 1992; Cooper et al., 1978; Williams, 1996; Wakefield et al., 1998; O'Shea, 1999; Lin et al., 1998, 2001; McConnell, 1998; Biordi, 1993; O'Brien-Pallas, Thomson, Alksnis, and Bruce, 2001; McGillis Hall and Doran, 2001; Blegen, Goode, and Reed, 1998; Nicklin and McVeety, 2002; Blendon DesRoches, Brodie, Benson, Rosen, Schneider, Altman, Zapert, Herrmann, and Steffenson, 2002; Levy, Gopher and Donchin, 2002; Flynn, Dorris, Holman, Carnahan, and Barker, 2002). A number of these researchers interviewed and surveyed nurses themselves to find out their perceptions of why medication errors occur. Results showed interruptions and distractions as a main cause of medication errors. Walters (1992) reported that 41.6% of the nurses surveyed cited frequent interruptions as one of the three most likely causes of error, along with delay in receiving medication from the pharmacy, and RN busyness. Gladstone (1995) also reported that nurses perceived distractions by other patients/events on the ward as one of the three most likely reasons for drug error, next to incorrect patient, and poor handwriting. In Wakefield et al.'s (1998) survey, being interrupted from administering medication to perform other duties was ranked as the highest cause of medication error.

These results are not surprising, as nurses work in a chaotic environment where their attention is often divided between many tasks, and can be vied for by other nurses or patients at any time (Davis, 1994; Wakefield et al., 1998). Other perceived causes of medication error included inadequate mathematical skills (O'Shea, 1999), lack of knowledge (McConnell, 1998), poor handwriting (Williams, 1996; Gladstone, 1995), ineffective drug labels (Davis, 1990), and

equipment design and non-adherence to human factors design principles (Cooper et al., 1978; Lin et al., 1998, 2001).

It is obvious that hospital nurses in many departments view interruptions as a significant problem and one of the main contributing factors to medication errors, which can seriously jeopardize patient safety. Similar findings of interruptions causing errors are also found in other complex work domains (see previous sections). Interestingly, the studies described in this section did not attempt to classify the way in which interruptions occur or the types of interruptions, or to quantify how often these interruptions take place in the nursing workplace. It is important to know this information to design and test interfaces for devices that are used in the interruptive nursing workplace, as interfaces can potentially be designed to minimize the negative effects of interruptions and reduce errors (Kreifeld and McCarthy, 1981; McFarlane, 1997).

Types and Frequency of Interruptions

Literature on the types and frequency of interruptions encountered in the medical domain is minimal. The main studies that could be found focused on interruptions experienced by: physicians and nurses in a hospital general medical ward (Coiera and Tombs, 1998), physicians in a hospital emergency department (Chisholm, Collison, Nelson, and Cordell, 2000), and nurses in a general medical office (Paxton, Heaney, Howie, and Porter, 1996). This section briefly describes the methods used and results obtained in these three studies.

By performing observations, keeping logs and descriptions of interrupting events, and recording participants' speech, Coiera and Tombs (1998) studied how eight physicians and two nurses in a British hospital were interrupted as they performed their daily duties over approximately four months. Three types of interruptions were identified: interruptions involving calls over the telephone, interruptions involving calls over the hospital paging system, and face-

to-face conversations. An average of one phone call or page every 18.5 minutes was observed. A participant was successfully contacted in 74% of the call events that were observed. The majority of participants generated and received many interruptions of all three types.

At the conclusion of their study, Coiera and Tombs (1998) suggested several strategies for improving communication methods in hospitals to reduce interruptions. Such strategies include voicemail and email with acknowledgement, message screening, and mobile communication. A common characteristic between these methods is that they are asynchronous forms of communication, enabling the receiver to reply to the sender when it is convenient. Interestingly, however, they found that medical staff, including nurses, generated twice as many interruptions via the telephone and paging systems as they received, and often favoured interruptive methods of communication over less interruptive methods. This shows that interruptions in a hospital setting are an effective and necessary means of communication, and are inevitable when patients are being treated and information, such as patient details and answers to questions about diagnoses and treatments, is required promptly. Thus, developing strategies to minimize interruptions themselves will likely not be sufficient enough to improve patient safety on the whole. Rather, it may also be necessary to minimize the disruptive effects of interruptions, such as memory loss and attention diversion, which lead to errors.

A similar method was employed by Chisholm et al. (2000) to study the types and frequency of interruptions of thirty physicians in an emergency department, where it is estimated that 93% of medical errors may be preventable. Participants were observed and tasks, interruptions, and breaks-in-tasks were recorded. Interruptions were defined as events that briefly occupied the attention of the participant, but did not require the participant to switch to new tasks. Breaks-in-tasks were defined as events that required the participant's attention for greater than ten minutes, and hence resulted in changing tasks. It follows from these definitions that an interruption may or may not result in a break-in-task. It is important to classify

interruptions, as different types will affect people differently, and be disruptive to different degrees (Horvitz et al., 1999). As was the case with the general medical staff in the study performed by Coiera and Tombs (1998), it was observed that interruptions are an inevitable working condition in emergency departments, as patient visits are unscheduled, certain medical conditions demand the immediate attention of certain emergency physicians, drawing their attention away from other tasks including attending to other patients, and physicians often need to answer questions regarding other patients while they are performing tasks.

In this study, eight main tasks of emergency physicians were observed: patient care, viewing test results, charting, teaching, listening to reports about patients, talking with other physicians about patients, giving orders, and personal breaks. It was observed that an average of 30.9±9.7 interruptions and 20.7±6.3 breaks-in-task occurred per 180-minute observational period. A rough calculation shows that this is approximately three times the number of interruptions that occurred in the general medical ward in Coiera and Tombs' (1998) study, indicating that emergency rooms are even more interruptive workplaces than general wards. Participants performed an average of 67.6±15.7 tasks per period.

The final major study that could be found in the literature, related to the types and frequency of interruptions, involved nurses recording information about interruptions that they encountered during patient consultations in a medical office (Paxton et al., 1996). In this case, an interruption was defined as any event that disturbed the nurse's work or caused a distraction. The study was performed in two stages, the first with 34 nurses, who reported 48.5 interruptions occurring per 100 consultations, and the second, one year later, with 33 nurses reporting that 30.2 interruptions occurred per 100 consultations. A new GP contract was instated during that year, which could have led to a change in workload for the nurses, decreasing the number of interruptions experienced. The nurses felt that GPs caused most of the interruptions, and that

most interruptions involved some form of listening, which was distracting. The researchers also surveyed the patients themselves, who found the interruptions intrusive as well.

It has been demonstrated, through a review of the literature, that interruptions are a significant problem and are a cause of medical error, especially in medication administration via infusion pumps, such as patient-controlled analgesia devices. Such devices are programmed everyday by nurses in hospital recovery rooms. Lin et al. (1998, 2001) and Ford and Rollinson (2001) have demonstrated that programming errors can be significantly reduced when interfaces for patient-controlled analgesia devices are designed using human factors principles, but to our knowledge no one has investigated the ability of such an interface to reduce the disruptive effects of interruptions. Therefore, similar methods to the ones described in this section (observations and interviews) can be used to study the types and frequency of interruptions that occur in the hospital recovery room, when nurses are programming patient-controlled analgesia devices. Then the work of Lin et al. (1998, 2001) and Ford and Rollinson (2001) can be extended to test their redesigned interfaces under more representative conditions.

DESIGNING INTERFACES TO DEAL WITH INTERRUPTIONS & CONDUCTING REPRESENTATIVE EVALUATIONS

McFarlane (1997) said it best: "The effects of user-interruption in HCI are directly related to the particular design chosen for the user interface of the system. The design of the user interface directly affects the states of dimensions of the interruption process and, therefore, causally affects the results of interrupting the user" (p. 67).

Segatore et al. (1994) suggest that nurses should learn how to effectively manage distractions, starting in school. In addition to this, the devices nurses use should assist them in managing interruptions and distractions, and in doing so reduce the medication errors made. McConnell (1998) found that less than half of the articles published in the nursing literature pertaining to common medical devices that nurses use discuss how to respond to alarms, device hazards, common user errors, and malfunctions. This problem is amplified by poor device design to begin with. Lin et al. (1998, 2001) redesigned a specific user interface for a patient-controlled analgesia (PCA) device according to human factors principles. They tested the new and old interfaces with both nursing students and experienced nurses, and found that for both groups, fewer errors were made and performance was faster with the new interface than with the old. Similarly, Ford and Rollinson (2001) redesigned the user interface for an epidural patient-controlled analgesia device (EPCA) and tested the new interface with nursing students. They also found that errors and programming time were reduced for the new interface versus the old, commercially available one. However, these experiments were not conducted under completely realistic conditions, such as with interruptions. Design flaws or inefficiencies may be further expounded when interruptions occur. That is, if an interface is complicated, confusing, and non-transparent to begin with, even more errors may be made when interruptions occur. Thus, it is important to assess the effectiveness of the redesigned interfaces for the PCA and EPCA devices

under interruptive conditions, to see if features of these new interfaces decrease the load on the nurse's memory to remember where he/she was in the programming task before the interruption occurred.

To prevent user error, human-computer interfaces should be designed such that they minimize the negative effects of interruptions (McFarlane and Latorella, 2002). That is, they must be iteratively designed and evaluated with the user and his/her workplace conditions in mind, and tested under representative conditions, simulating the types and frequency of interruptions that actually occur. By doing this, as well as by adhering to other human factors principles, the disruptive effects of interruptions may be reduced. To our knowledge, there is only one other study in the literature that has evaluated the usability of intravenous infusion (IV) pumps under distractive conditions. Wiklund, Smith and Baker (2002) conducted a usability comparison of three IV pumps in an intensive care unit simulator with seventeen experienced nurses. Ambient distractions included conversation, ringing telephone, intercom announcements, staff entering and leaving, and noise generated by medical devices. Task-specific distractions included telephone calls from physicians to the participants, alarms emitted by the simulated patient's monitor and a power failure. These distractions increased the nurses' sensory and cognitive demand. The authors also performed the same study in a usability laboratory under static conditions and compared the results. They concluded that a more representative test environment enabled better identification of user errors due to a more realistic and higher cognitive demand, whereas a more static environment allowed participants to focus more on the pumps and hence directly reveal usability issues. Therefore, they stress the need to do both types of testing. The design philosophy of the pumps in question was not described. Therefore, it is not known whether human factors principles were employed when those devices were designed.

The study by Kreifeld and McCarthy (1981) showed that interruption resistance should be a criterion for the design and evaluation of interfaces. A robust interface should ideally show

insignificant differences in performance between interrupted and uninterrupted tasks. They suggest that for calculators, the last data entry and operation, as well as the current resultant, be displayed at all times, and that there should be a review key to view previous data entries and operations. This recommendation can easily be extended to other types of interfaces, to show users where they are in a task, and to allow them to view previous actions. This would provide the cues that Altmann and Trafton (2002) deem essential, to prompt their short-term memory and help to build a cognitive map, or mental representation, which would reduce the cognitive effects of interruptions. McFarlane and Latorella (2002) also purport that in the case of interruptions that are immediate in nature, with no lag, interfaces should remind users of objectives and previous activities, and have replay capability if possible. Storch (1992) found that performance after an interruption on a character-based interface tended to be superior to that on a mouse-based graphical user interface because the graphical user interface required mouse positioning, which is a more complex interaction and hence is more susceptible to disruption by interruptions. Participants spent more time looking at the graphical user interface than the character-based interface, and committed more errors. Thus, simplicity in an interface is essential if it is to be robust in the face of interruptions.

It is possible that the new interfaces developed by Lin et al. (1998, 2001) and Ford and Rollinson (2001) are already robust enough to withstand the effects of interruptions. They were designed using well-accepted human factors principles. Both of the new interfaces for the PCA and EPCA devices feature a dialogue overview that show the user's location in the programming sequence, which should provide cues to prime a user's memory. Furthermore, both of the new interfaces also have a Previous Screen (EPCA) or Review (PCA) button that allows the user to go back to the previous step to see what they had programmed last. Also, the general reduction in complexity that these interfaces exhibit versus the old interfaces should also prove beneficial when interruptions increase the cognitive demand and workload placed on users. However, the

potential benefits of these features in reducing the disruptive effects of interruptions must be verified by tests conducted under representative interruptive conditions. The goal of these experiments would be to determine whether the relative disruptive effects of interruptions are lessened with an interface that is designed based on human factors principles. That is, will interruptions be less disruptive to the new interfaces than to the old ones? As the Food and Drug Administration (FDA) acknowledges, "devices that can be used safely under conditions of low stress (i.e., low workload) could be difficult or dangerous to use under conditions of high stress" (Kaye and Crowley, 2002, p. 10). If the new interfaces do not show a reduction in the disruptive effects of interruptions, it may be necessary to further modify the designs, by providing more feedback to the user and minimizing the load on the user's memory even more, to further reduce the disruptive effects of interruptions.

In order to design such representative experiments, it was necessary to conduct a field study to answer some questions not addressed in the current literature, such as:

- what is the nature of the interruptive environment that recovery nurses work in when they program the commercially available PCA and EPCA devices and care for patients,
- what cognitive resources do these interruptions utilize,
- do recovery room nurses have the option of attending to interruptions at a later time, and
- are the interruptions of the task-sharing or task-switching type, or both?

FIELD STUDY

Since nurses frequently encounter interruptions and distractions, it is important that these interruptions be observed, quantified, and classified so that experiments in which interfaces are tested can be conducted under more representative conditions, with simulated interruptions. Thus, a field study was conducted to examine exactly what types and frequencies of interruptions occur in the nursing workplace.

Location

The field study was conducted in the Post-Anesthetic Care Unit (PACU) at the Toronto General Hospital, a teaching hospital and a member of the University Health Network, during which nurses were observed as they cared for their patients. The PACU is essentially a critical care recovery room where patients are transported after undergoing surgery in the operating room. Most patients remain in the PACU for about half an hour while their anesthetic wears off, and are then transported to the floor or sent home. The most critical cases can remain in the PACU for hours, or even overnight. The PACU differs from most other hospital units in three main ways: (a) family members are not permitted to visit, unless the patient has a prolonged stay in the PACU, (b) approximately 90% of the hospital's PCA and EPCA device usage takes place here, and (c) the PACU has an open concept layout, with a large room divided into patient bays, but nurses can see and hear each other at all times and there are no curtains surrounding the patients. The PCA and EPCA devices are frequently programmed by nurses in the PACU.

Participants

Ten nurses in total were observed as they cared for their patients. These nurses have worked in the PACU for an average of 6.4 years (ranging from 7 months to 22 years). All are Registered Nurses with completion of a critical care course and/or critical care experience.

Informed consent was obtained from each nurse who was observed (see Appendix 1). The nurses were not specifically told that the observer was looking at interruptions, but rather that the overall work environment was being assessed. It was emphasized that their performance was not being evaluated so as not to alter the behaviour of the nurses during the observations.

Method

Ethics approval to perform the field study in the PACU was first obtained from the University of Toronto and then from the University Health Network. In total, 25 hours of observations were conducted over several days. These observations were performed during the busiest times in the PACU, which are typically Tuesday-Thursday from 2-9pm. During these times nurses can care for up to three patients at once. The following information was recorded for each interruption that each nurse encountered:

- The time at which the interruption occurred,
- The time at which the interruption was attended to by the nurse,
- The time at which the nurse finished attending to the interruption,
- The time at which the nurse returned his/her full attention to the primary task,
- Whether the interruption resulted in task-switching (requiring the nurse to leave the primary task to attend to the interruption), or task-sharing (requiring the nurse to attend to both tasks at the same time),
- A description of the interruption task,
- A description of the primary task,
- The source of the interruption (i.e., another nurse, etc.),
- How the interruption was announced to the nurse (i.e., face-to-face, via a pager, etc.), and
- Observed detrimental effects of the interruption on the nurse's performance of the primary task.

Results

Table 2 summarizes the type and length of the interruptions that were observed. All observed interruptions were attended to immediately by the nurses, and for all of the interruptions that resulted in task-switching, the nurse resumed his/her primary task immediately following completion of the interruption task. Primary tasks included charting, hooking up intravenous (IV) lines, administering medications, and drawing blood. In addition, all of the verbal interruption tasks, such as answering a question or listening to a report, with the exception of the phone call, arose from "face-to-face" interactions, meaning that the source of the interruption was close to the nurse, who was made aware of the interruption task verbally.

Table 2. Type and length of observed interruptions.

Interruption Task	Source	Length	Type
Answer a question	Another PACU nurse	Very brief (i.e., under 1 minute)	Task-sharing or task switching
Move 10 feet away from X-ray being taken	Nearby patient	Very brief	Task-switching
Keep an eye on the patient's monitor, while doing other tasks (i.e., injecting medications, drawing blood, charting, etc.)	Patient's monitor	Continuous	Task-sharing
Listen to verbal report while hooking patient up to PACU monitors, connecting IVs, etc (when patient first arrives)	Another PACU nurse, operating room nurse, or anesthesiologist	Brief (i.e., under 3 minutes)	Task-sharing
Talk with patient while performing various tasks	Patient	Brief or continuous	Task-sharing
Help another nurse, (i.e., check another patient's ventilator, draw blood, etc.)	Another PACU nurse	Brief	Task-switching
The nurse has two or three patients (usually stable ones), and therefore has to monitor all of them and perform tasks on each	Patients	Continuous	Task-sharing or task switching
Give verbal report on the patient's status	Another PACU nurse or a physician on rounds	Brief	Task-sharing or task switching
Phone call	Anyone	Brief	Task-switching

Figure 2 illustrates the number of times each of the interruptions in Table 2 were observed in total, for those interruptions that were not of a continuous nature.

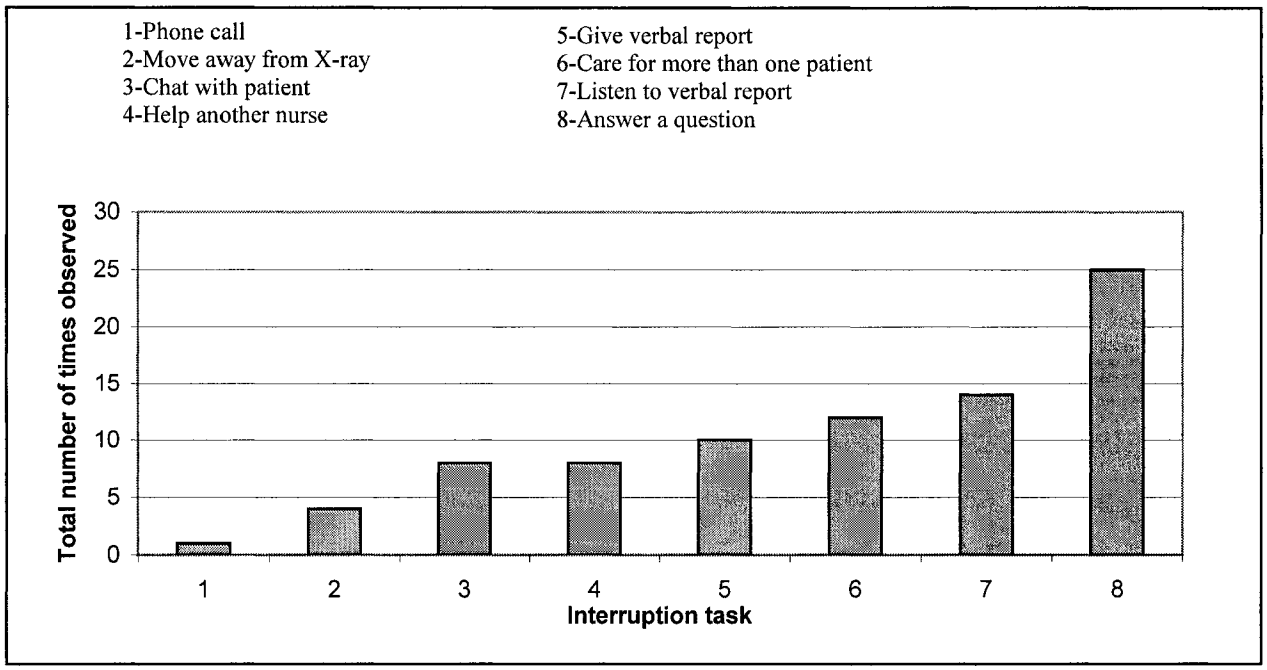


Figure 2. Number of times interruption task observed.

Discussion

The observation that all interruptions were attended to immediately is a characteristic consistent with the nursing workplace, where issues concerning patient care must be dealt with as they arise. Furthermore, as Figure 2 illustrates, the most frequent interruption tasks required verbal cognitive resources, resulted from face-to-face interactions, and either required the nurses to recall information from memory or commit information to memory. Thus, simulated interruptions used in realistic experiments that mimic the PACU should be designed to use this same cognitive resource and place the same cognitive demand on participants.

No obvious detrimental effects of the interruptions on performance, in terms of errors committed, were observed during the field study. However, the nurses who were interrupted occasionally exhibited frustration from the increased workload and mental demand imposed by

the interruption. In addition, it is important to note that throughout this field study the PACU was not operating at maximum capacity. There are 21 operating rooms serviced by the PACU, seven of which were closed due to a shortage of anesthesiologists. Therefore, it would be useful to conduct a follow-up study when all 21 operating rooms are open and functioning. It is likely that the same interruption tasks that were observed in this field study would occur even more frequently, and thus PACU nurses could potentially experience an even greater increase in workload and cognitive demand, and performance may be affected to a greater degree.

Finally, individual differences were a limitation of this study. Personality traits and individual processing and focusing abilities of each nurse may influence the interruptions that they experience and their consequences.

Conclusions

Despite the limitations involved with this field study, its purpose, to examine the interruptions experienced by nurses in the actual workplace in order to design representative experiments so that findings can be best generalized outside of the laboratory, was achieved. The experiments that were designed with the simulated interruptions are described next.

INTERFACE EXPERIMENTS

Using the results from the field study described in the previous section, laboratory experiments were conducted to explore the effects of representative interruptions on performance by nursing students and experienced nurses programming the new and old interfaces of both devices (PCA and EPCA). Ethics approval for these experiments was obtained from the Ethics Review Board at the University of Toronto, prior to their commencement. Pilot studies were conducted to determine how to make the experiments as representative as possible (for maximum generalizability of findings to the actual nursing workplace) while maintaining as much experimental control as possible. Following the pilot studies, three experiments were conducted:

- Experiment 1 involved nursing students performing programming tasks with the PCA interfaces,
- Experiment 2 involved nursing students performing programming tasks with the EPCA interfaces, and
- Experiment 3 involved experienced nurses performing programming tasks with both the PCA and EPCA interfaces.

Hypothesis

It was hypothesized that the new interfaces of both devices would reduce the disruptive effects of interruptions over the old interfaces, primarily owing to their less complex programming sequences, more intuitive user-friendly layouts, button groupings and labels, and feedback features. It was also hypothesized that both the nursing students and the experienced nurses would take longer to complete programming tasks, commit more programming errors, and experience a higher workload when interruptions occur.

Participants

Two groups of participants were selected: nursing students from the University of Toronto who already have had some nursing education but have never programmed a PCA or EPCA device before (Experiments 1 and 2), and experienced nurses who work in the PACU and program these devices daily (Experiment 3). All participants signed informed consent forms prior to participating in the experiments. Twenty-two nursing students (twenty females, two males) between the end of their first and final years of their nursing degree agreed to participate in the study (eleven for Experiment 1 and eleven for Experiment 2). Seven experienced nurses (all female) agreed to participate in the study for Experiment 3. These nurses have worked in the PACU for an average of 8.2 years (ranging from 1.5 to 20 years) and program the PCA device used in this experiment an average of approximately five times per day and the EPCA device used in this experiment an average of approximately three times per day. All participants were compensated $20/hr.

Materials

All three experiments were conducted in a quiet room, so that no external distractions or interruptions would be present. Computer simulations of the new and old interfaces of both devices were used. Computer simulations of the new and old interfaces of the PCA device existed previously in Toolbook (version 1.5) (Lin et al., 1998, 2001) but were re-coded into Visual Basic for the purposes of these experiments (see Figures 3a and 3b for examples of the computer simulations for the new and old PCA interfaces), whereas those of the EPCA device already existed in Visual Basic (Ford & Rollinson, 2001) (see Figures 3c and 3d for examples of the computer simulations for the new and old EPCA interfaces). A tool to assist in the collection of data was developed in Visual Basic for use in these experiments. This tool, called the p-player, records the time taken to complete each programming task, as well as each button pressed

by the participant in sequence, so that the experimenter can accurately pinpoint the programming errors that were committed (see Figure 3e for an example).

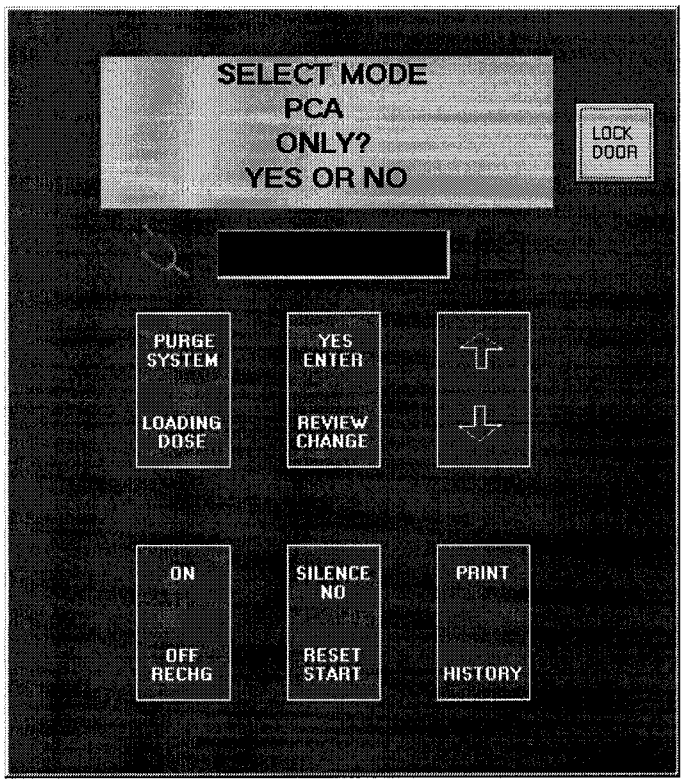


Figure 3a. Screen shot of the old PCA interface (Lin et al., 1998).

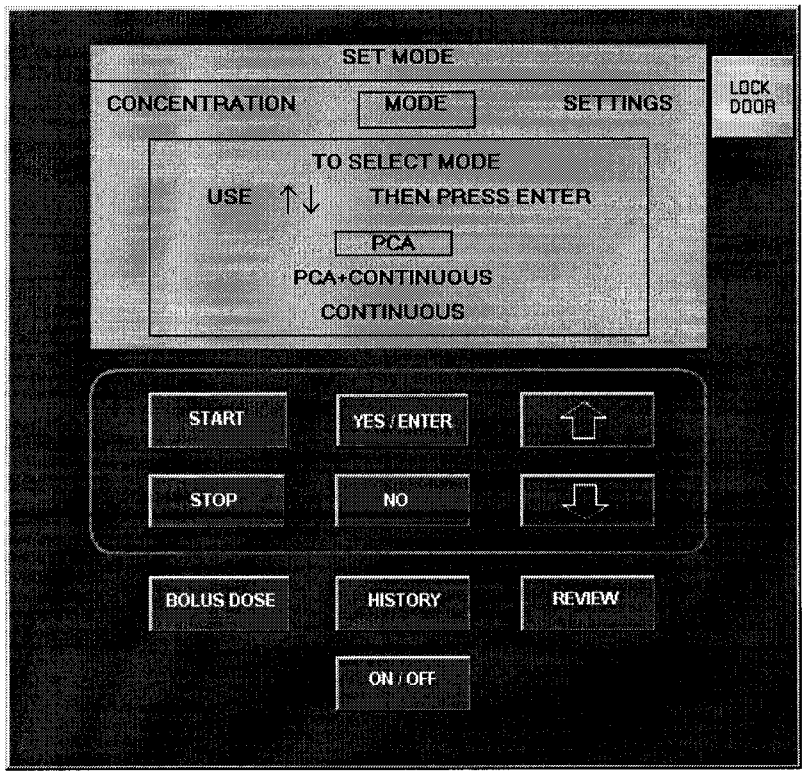


Figure 3b. Screen shot of the new PCA interface (Lin et al., 1998).

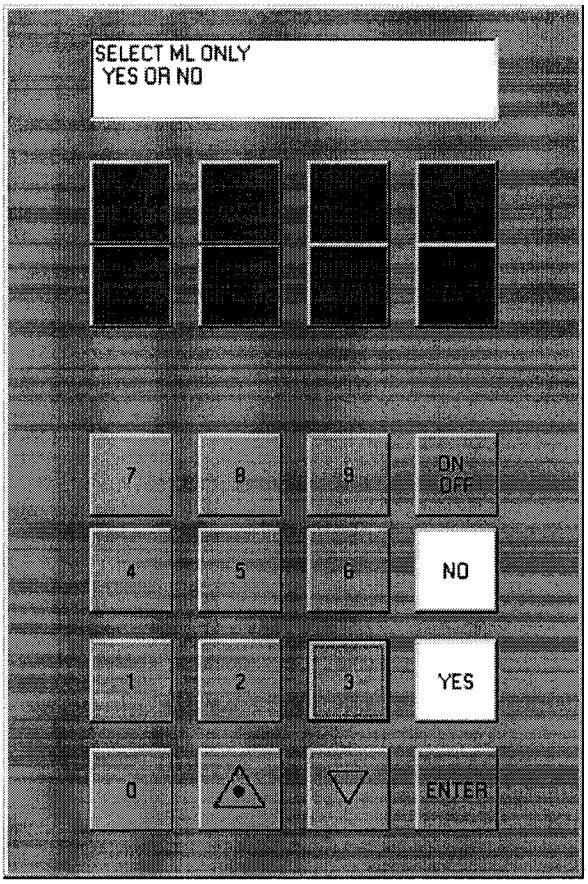


Figure 3c. Screen shot of the old EPCA interface (Ford & Rollinson, 2001).

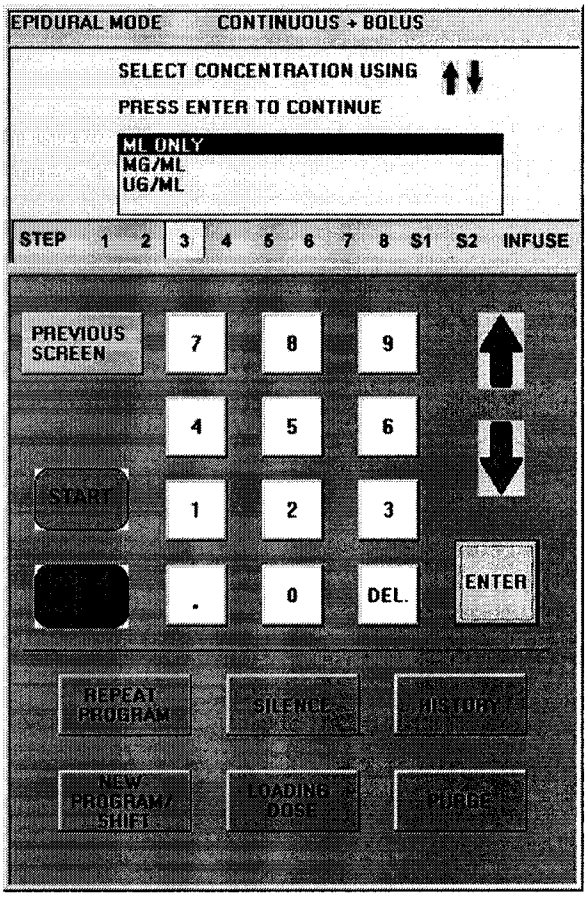


Figure 3d. Screen shot of the new EPCA interface (Ford & Rollinson, 2001).

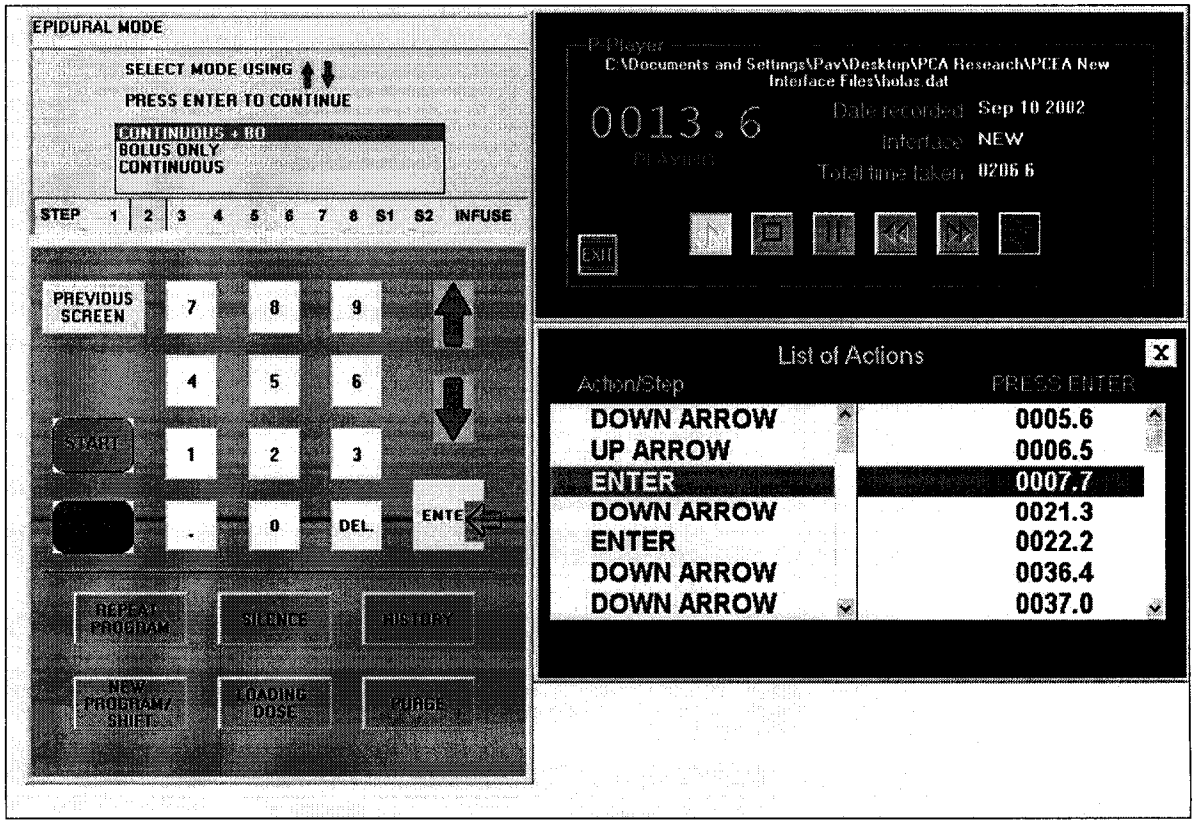


Figure 3e. Screen shot of the p-player tool with the new EPCA interface.

The computer simulations of the interfaces for both devices were run on an Intel Pentium II processor (448MHz, 128MB of RAM) PC with Windows XP as the operating system. A standard Logitech mouse was used by the participants to operate the simulations that were displayed on a 12-inch colour monitor.

PCA and EPCA order forms that are currently used by nurses in the PACU were used in these experiments. One order form was used for each mode that was tested (see Appendix 2). One order form was used for each mode for practice purposes as well, with different parameter values than the experimental order forms. The PCA order forms were modified slightly to include a continuous dose parameter, so that the pca+continuous and continuous only modes could be tested. Currently nurses in the PACU only operate the PCA device in pca only mode (the PCA order form used in the PACU was revised since Lin et al.'s (2001) study to reflect this). However, to make the results of these experiments comparable to those of Lin et al. (2001), all three modes were tested. Similarly, nurses in the PACU always program the EPCA device in the continuous+bolus mode, but all three modes were tested in these experiments.

Two patient charts were obtained from the PACU and photocopied (without any identifying information) to be used for the interruption tasks in these experiments. An audiotape was used to record the interrupted trials as well as the interviews at the end of each experimental session. A stopwatch was used to time the interruptions.

Experimental Designs

For the nursing students, a 2x3x2 within-participant design was used for each device (11 nursing students programmed the PCA device in Experiment 1 and 11 programmed the EPCA device in Experiment 2), with interface (new and old), mode (pca only, continuous only, and pca+continuous) and interruption (yes and no) as the within-participant factors. For the experienced nurses, a 2x2x3x2 within-participant design was used, with the same factors and

levels as the novice participants, but with the addition of device as another within-participant factor because it was more difficult to recruit experienced nurses to participate in the study. Counterbalancing was used to eliminate order effects of device (for experienced nurses), interface, mode, chart used, and interruption questions asked.

Procedure

Each participant in each experiment participated in two experimental sessions, scheduled on two different days to avoid effects from fatigue. At the beginning of her/his first session, each participant read an information sheet about the study (see Appendix 3), read and signed an informed consent form (see Appendix 4), filled in a demographics questionnaire (see Appendix 5), and read written instructions and received supplementary verbal instructions on how to fill out a mental workload questionnaire (see Appendix 6). In Experiments 1 and 2, the nursing students programmed one interface in their first session and the alternate interface in their second session. One chart was used in the first session and the alternate chart in the second session. In Experiment 3, the experienced nurses programmed both interfaces of one device in their first session and both interfaces of the alternate device in their second session. This was done so that if an experienced nurse ended up only attending one session, she would have at least programmed both interfaces of one device and her data could be used. This in fact was the case for one of the experienced nurses who only programmed the interfaces of the PCA device. Both charts were used for the experienced nurses in their first session (one for the first interface and the alternate one for the alternate interface). Although the same charts were used in the second session, different questions were asked. All participants performed six programming trials with each interface (each of the three modes interrupted and not interrupted). Interrupted and uninterrupted programming trials were alternated, but all participants always started with a non-interrupted trial. For each mode, the same order form (with the same parameters) was used for

both the interrupted and uninterrupted trials for experimental control purposes, but the six trials were ordered such that participants would perform two trials on two different modes between programming the same mode again. Also, the same sequence of programming trials was used for both interfaces of each device, again for control purposes.

Prior to the first programming trial being performed for each interface, each participant was given basic training on the interface by the experimenter and unsupervised practice on the interface (with the practice order forms), with the experimenter present to answer questions. For the PCA device interfaces, the participants were also told during the training session to answer "no" on the purging step and to not administer a loading dose, as this is the PACU policy and adhering to the PACU policy as much as possible would make the experiments as realistic as possible. For the EPCA interfaces, participants were also told to choose epidural mode for all trials and not to set or deliver a loading dose, for the same reasons. Participants were also given two minutes to review the patient's chart that was to be used for the interrupted trials for that interface. This was to simulate the participant having an actual patient and knowing something about that patient.

The pilot studies showed that some participants take longer to program the same mode than others, or to answer the same question than others. Originally it was hoped that interruptions could be timed to occur during certain pre-determined subtasks. However, due to these timing issues, and the fact that participants used different strategies to cope with each interruption they encountered, it was impossible to always squeeze an interruption in at the same time for each participant, and consequently each participant would experience a different number of interruptions for the equivalent trial, thereby confounding the results. To remedy this situation, frequency was used as a control, rather than timing the interruptions to coincide with specific subtasks. When this frequency method was tested in the pilot studies, six seconds was found to be a worst-case scenario that would place enough load on the participants during the

shortest trials. This is more frequent than what we observed in the field study, however since no one has performed such experiments before, an interruption frequency of six seconds was found to be a good starting point to investigate the effects of interruptions on performance. If we were to observe no impact under these high frequency conditions, then it would not be worth investigating less frequent interruptions. In interrupted trials on the PCA device, the first interruption was given immediately following the participant answering yes/no to the purge prompt. In interrupted trials on the EPCA device, the first interruption was given immediately following the participant selecting either epidural or pca mode. Subsequent interruptions for both devices were given six seconds after the participant resumed the primary programming task (i.e., six seconds were counted starting when the experimenter observed the participant clicking a button on the interface after completing the previous interruption task). Thus, participants were allowed to reorient themselves to the interfaces and continue programming for six seconds before the next interruption occurred.

The interrupted trials consisted of the experimenter verbally interrupting the participant with a question related to information in the chart every six seconds of programming time (see Appendix 7 for the list of questions that was used for each chart). The experimenter pretended to be either another nurse or a physician on rounds requesting information about the participant's "patient". It was emphasized at the beginning of each session that the participants should work as quickly and accurately as possible when programming the interfaces and when answering the questions (to mimic the actual work environment which is very fast paced while requiring a high degree of accuracy as lives are at stake if mistakes are made). Participants were also instructed to respond to the questions immediately (to mimic the immediate nature of the interruptions observed in the field study). Participants were allowed to deal with each interruption using whatever strategy they preferred to accomplish the goals of the experiments (to program as accurately and quickly as possible while answering the questions immediately and as accurately

and quickly as possible). Participants' answers to the interruption questions were recorded onto audiotape as well as written by the experimenter.

At the end of each programming trial, participants filled out a NASA-TLX questionnaire (adapted from Hart and Staveland, 1988) (see Appendix 8). This is a widely-used and extensively investigated subjective assessment of workload, in which participants rate how much each of the six workload factors contributed to their workload for that trial, and these ratings are then weighted by the number of times each factor is circled in the pairwise comparisons. At the end of each session, participants were interviewed by the experimenter to obtain their subjective comments and preferences, which were recorded on audiotape. The following questions were asked after the six programming trials were completed with the first interface:

- Do you have any comments in general about the interface?
- How did you feel during the trials when you were not being asked questions?
- How did you feel during the trials when you were being asked questions?
- Did answering the questions affect your performance on the primary programming task? If so, how?
- What strategies did you use to cope with the interruptions?
- Do you have any other comments?

The same questions were asked after the six programming trials were completed with the alternate interface, with the addition of the following questions:

- How does this interface compare to the first one you programmed?
- Which interface do you prefer?
- Which interface is easier to program when interruptions occur?

Finally, if the participant had just programmed the new interface of either device, she/he was also asked about whether or not the feedback features were used and under what circumstances. Participants were debriefed at the conclusion of their second session.

Performance Measures

Three measures were used to assess performance on the primary programming task: programming time, subjective mental workload, and the number of programming errors committed.

Programming time was measured by the p-player as the total time it took the participant to complete the primary programming task. For the PCA device, this was the total time from when the participant clicked the ON button (old interface) or ON/OFF button (new interface) until the participant clicked the LOCK DOOR button (new and old interface). For the EPCA device, this was the total time from when the participant chose the mode (epidural or pca for both interfaces) until when the participant clicked START (new interface) or RUN STOP (old interface). Hence, programming time for these experiments included the time taken to attend to the interruptions in interrupted trials.

Any action deviating from the correct, and hence most efficient, set of programming actions within any given trial was considered to be an error. Further to this are two classifications of errors: corrected and uncorrected. Uncorrected errors are defined as being uncorrected before the end of the trial. These can include an uncorrected error in the units chosen, concentration, mode, dose, rate, lockout interval, four hour limit, or total amount. Corrected errors can include any of the above errors that were corrected before the end of the trial, as well as setting but not delivering a loading dose, and actions resulting in a decrease in programming speed/efficiency, such as pressing an incorrect key when trying to get to the next screen or previous screen. Cascading errors are errors that result directly from a previous error (such as neglecting to set the continuous dose due to choosing pca only mode instead of pca+continuous mode). In the case of cascading errors, only the first error was counted. The total number of errors is the number of uncorrected plus corrected errors.

Statistical Analysis

SAS (version 8.2) for Windows 2000 server was used to perform statistical procedures on the data. An alpha level of 0.05 was used for all statistical tests. 95% confidence limits were computed where possible, to display on graphs and to explore significant interactions (Loftus and Masson, 1994). Nonparametric one-tailed sign tests were performed to complement ANOVA results (Siegel, 1956). Data was averaged over all three programming modes for the sign tests.

Results and Discussion (Experiment 1)

Programming time. For novice participants programming the PCA device, significant main effects were found for Interface ($F(1,10)=23.36$, $p=0.0007$) and Interruption ($F(1,10)=129.46$, $p<0.0001$), and significant 2-way interaction effects were found for Interface*Mode ($F(2,20)= 6.15$, $p=0.0083$) and Interface*Interruption ($F(1, 10)=12.51$, $p=0.0054$). Thus, as shown in Figure 4, the new PCA interface significantly reduced programming time overall compared to the old PCA interface, which was over 1.5 times slower. Figure 5 illustrates that interruptions increased programming time overall when compared to no interruptions for the PCA device, although it should be noted that programming time here includes the time to attend to the interruption task as well as the actual interface programming time, as it was not possible to separate the two because participants used both task-switching and task-sharing during the interrupted trials. A closer look at the Interface*Interruption interaction in Figure 6a reveals that when no interruptions were present, programming time for both interfaces was not significantly different (although 10 out of 11 participants programmed the new interface faster than the old ($p=0.006$, one-tailed sign test)), whereas the new PCA interface was significantly faster than the old PCA interface when interruptions were present (all 11 participants programmed the new interface faster than the old ($p<0.001$, one-tailed sign test)). It also shows that interruptions significantly increased programming time for both the new and old

interfaces (when programming the same interface, all 11 participants had faster programming times during uninterrupted trials ($p<0.001$, one-tailed sign test)). These results support our hypothesis that the new interface reduces the disruptive effects of interruptions on programming time relative to the old interface, which is clearly shown in Figure 6b.

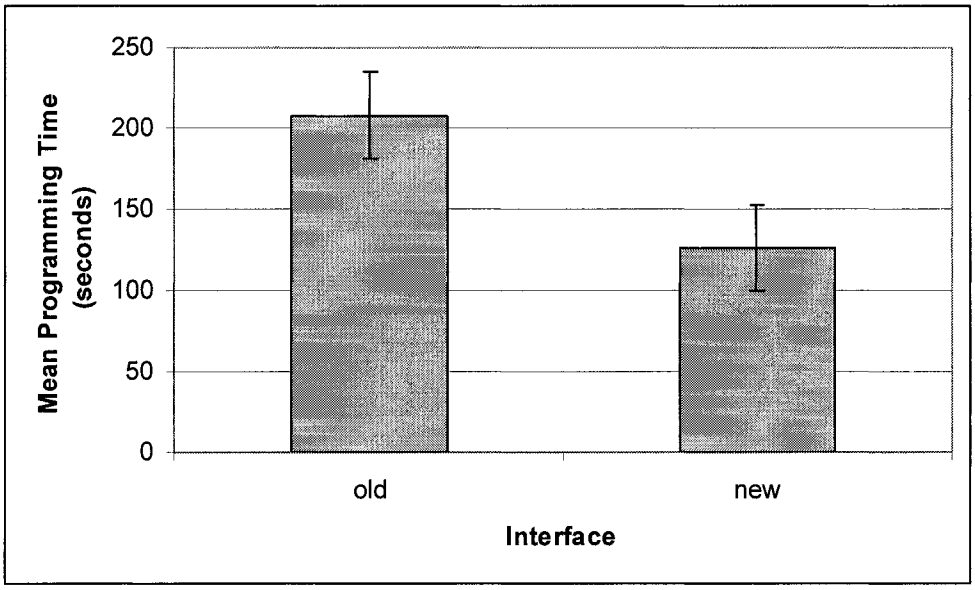


Figure 4. The effect of interface on programming time for novice PCA users, plotted with 95% confidence limits.

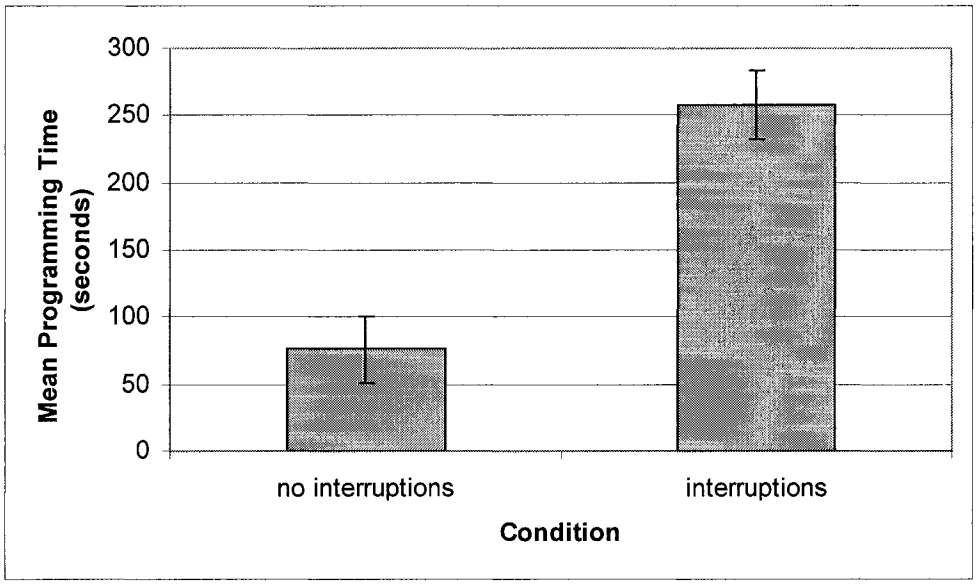


Figure 5. The effect of interruptions on programming time for novice PCA users, plotted with 95% confidence limits.

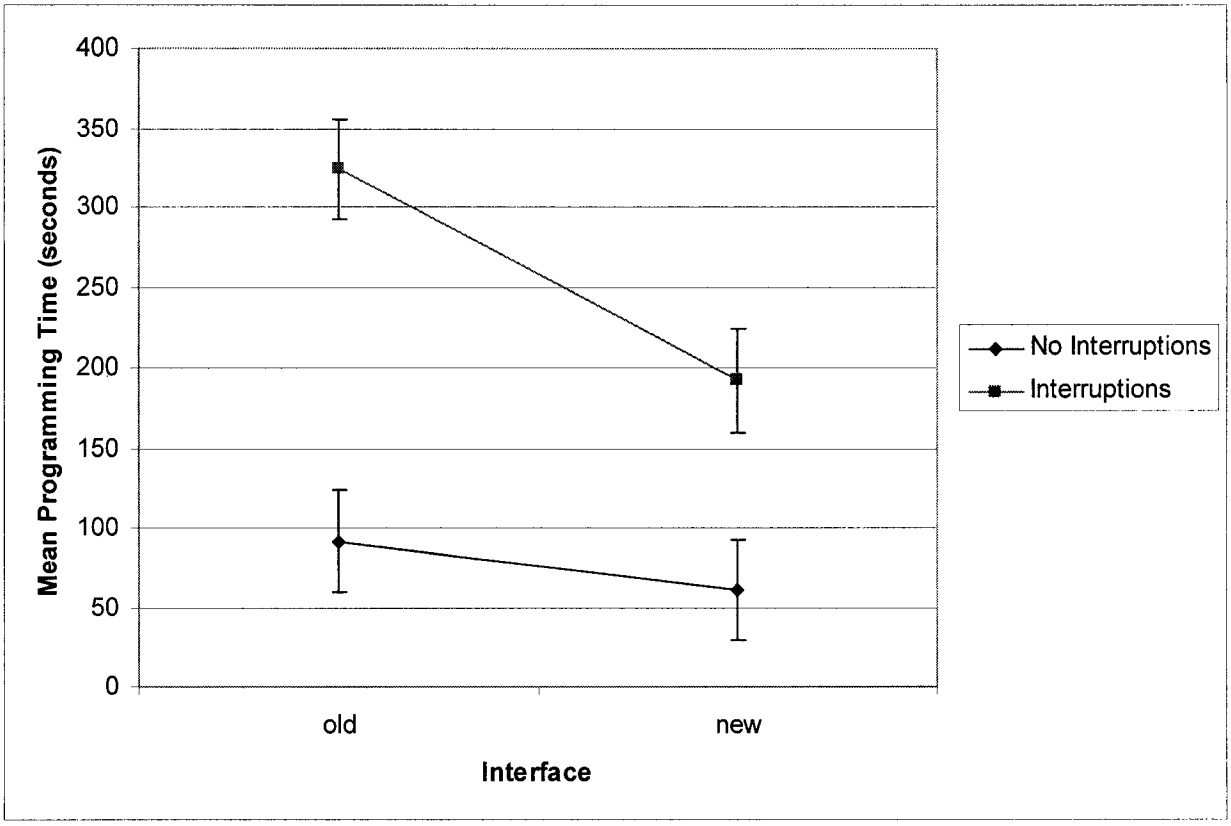


Figure 6a. The Interface*Interruption interaction for novice PCA users, plotted with 95% confidence limits.

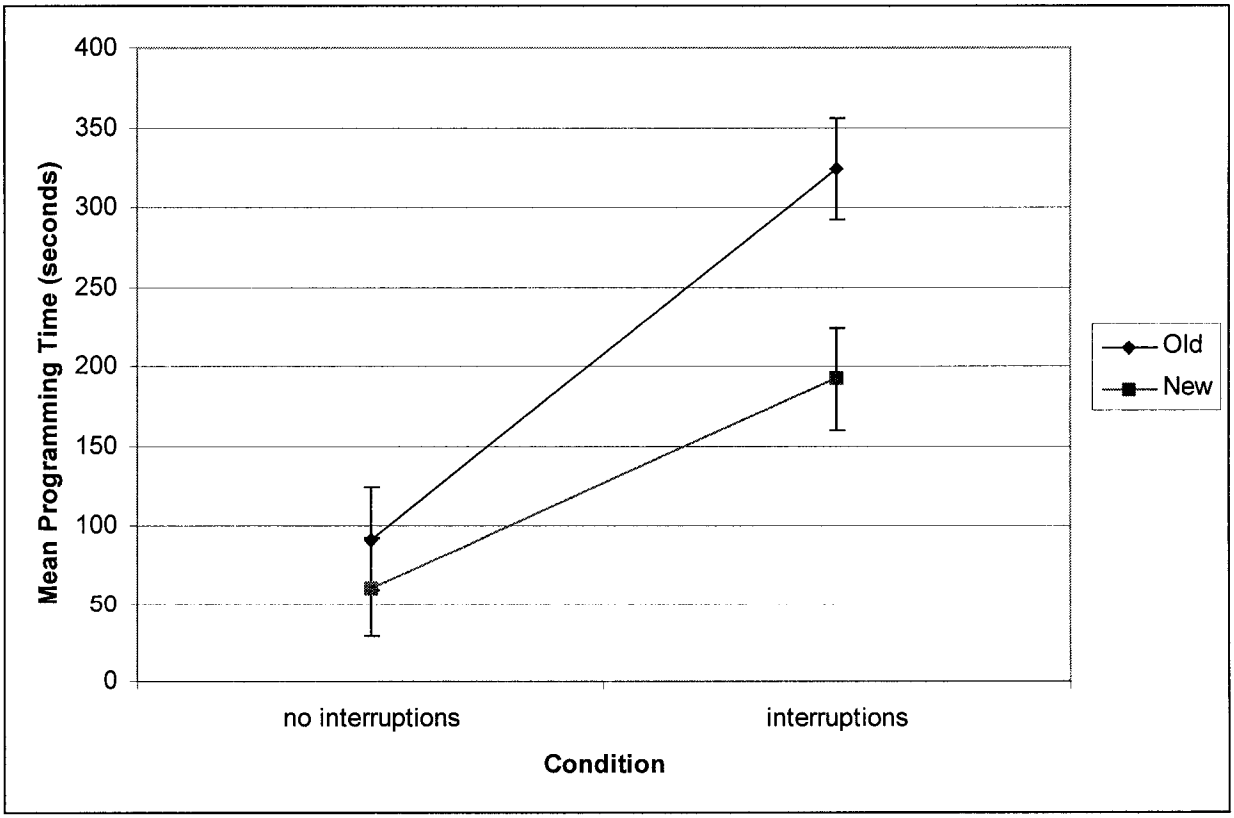


Figure 6b. Another representation of the Interface*Interruption interaction for novice PCA users, plotted with 95% confidence limits.

Workload. For novice participants who programmed the PCA device, a significant main effect was found for Interruption ($F(1,10)=93.30$, $p<0.0001$), and a marginally significant main effect was found for Interface ($F(1,10)=4.75$, $p=0.0543$). Thus, as shown in Figure 7, the new PCA interface reduced workload overall when compared to the old interface. Figure 8 reveals that workload was significantly higher in interrupted trials than in uninterrupted trials for the PCA device, with interruptions more than doubling the workload experienced by novice participants when programming the PCA interfaces (all 11 participants experienced higher workload when programming the new interface with interruptions than without ($p<0.001$, one-tailed sign test), while 10 out of 11 participants experienced higher workload when programming the old interface with interruptions than without, and one participant experienced comparable levels of workload ($p=0.001$, one-tailed sign test)). When no interruptions were present, 8 out of 11 participants experienced lower workload when programming the new interface than the old, and one participant experienced comparable levels of workload (not significant with a sign test). When interruptions were present, 8 out of 11 participants experienced lower workload when programming the new interface than the old (not significant with a sign test). These results do not support our hypothesis of the new interface reducing the disruptive effects of the interruptions in terms of workload because the Interface*Interruption interaction was not significant. However, this could be attributed to small sample size. It could also be that participants were rating the workload associated with the interruption tasks as well as that associated with the programming tasks when they were filling out the NASA-TLX questionnaires, since they filled them at the end of each trial, which could have confounded the workload results.

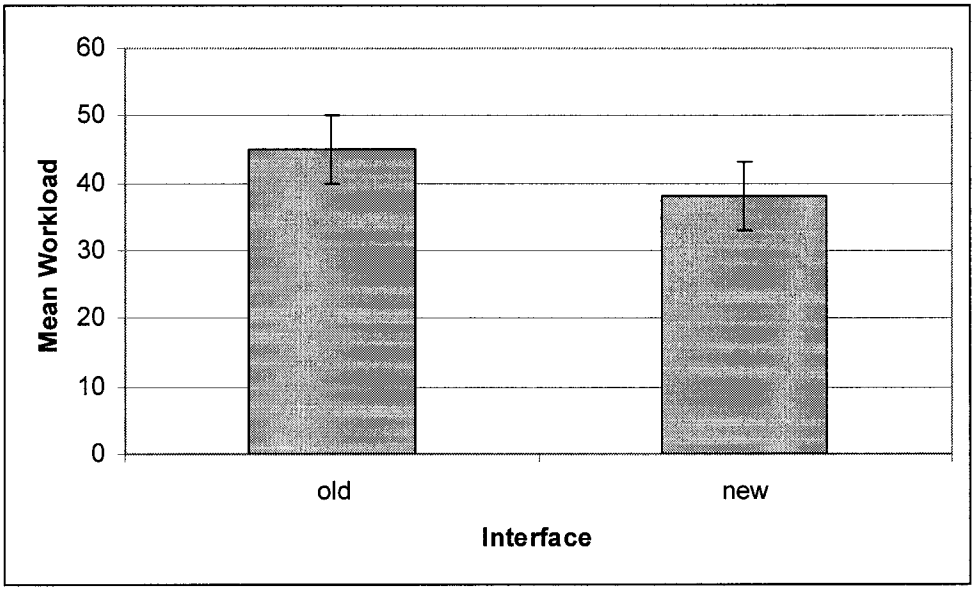


Figure 7. The main effect of interface on workload for novice users of the PCA device, plotted with 95% confidence limits (p=0.0543).

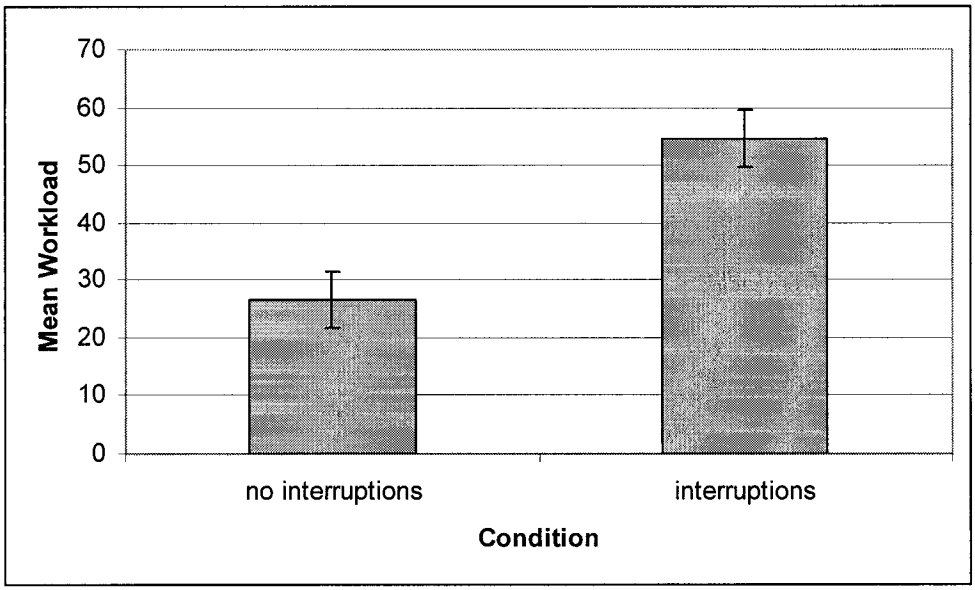


Figure 8. The main effect of interruptions on workload for novice users of the PCA device, plotted with 95% confidence limits.

Errors. Tables 3 and 4 describe the errors that were committed by the novice participants while programming the old and new PCA interfaces, respectively. The errors that are highlighted in grey represent uncorrected errors. To summarize these tables, novice participants committed seven errors on the old PCA interface (all were corrected), six of which occurred during interrupted trials, and nine errors on the new PCA interface (three were uncorrected), five of which occurred during interrupted trials (all uncorrected errors occurred

during interrupted trials). One concentration programming error was made with the old interface during an interrupted trial. The error data do not support our hypothesis of the new interface reducing the disruptive effects of interruptions. This will be discussed in detail in the General Discussion section.

Table 3. Errors committed by novice participants on the old PCA interface.

Participant	Mode	Interrupted?	Description
1	pca	yes	pca dose error (entered 2mg/mL instead of 1mg/mL), detected in history and corrected through review/change
1	pca+continuous	yes	after entering all the settings, redid program through review/change even though no mistake was made
6	pca	yes	pca dose error (entered 2mg/mL instead of 1mg/mL), detected in history and corrected through review/change
6	pca+continuous	no	after entering all the settings, went through review/change even though no mistake was made
8	pca+continuous	yes	concentration error (entered 1.5mg/mL instead of 2.5mg/mL), realized right away and corrected right away through review/change
10	pca	yes	lockout interval error (entered 5min instead of 6min), detected in history and eventually corrected in review/change but made another mistake in the process (see next)
10	pca	yes	entered 'yes' to administer loading dose prompt, but instead of setting it pressed review/change and fixed previous mistake (entered 'no' to administer loading dose prompt this time)

Table 4. Errors committed by novice participants on the new PCA interface.

Participant	Mode	Interrupted?	Description
4	pca+continuous	yes	continuous dose error (entered 1.5mg/h instead of 1mg/hr), not detected and not corrected
6	continuous	yes	mode error (chose pca+continuous instead of continuous only), realized on next screen and corrected right away using review button
7	pca+continuous	no	mode error (chose continuous instead of pca+continuous), detected on summary screen and corrected using review button

8	pca+continuous	yes	mode error (chose pca instead of pca+continuous), not detected and not corrected (therefore continuous rate was never set)
9	pca+continuous	yes	mode error (chose continuous instead of pca+continuous), not detected and not corrected (therefore pca dose and lockout interval were never set)
10	continuous	no	after entering all the settings, pressed review button to get back to beginning and reprogrammed interface, even though no mistake was made initially; made a mistake during reprogramming (see next)
10	continuous	no	mode error (chose pca instead of continuous), detected and eventually corrected after making another mistake (see next)
10	continuous	no	attempted to correct previous mode error by entering 0mg for pca dose, detected and corrected
11	pca+continuous	yes	after entering all settings, pressed review button to check them instead of history button

Comments and preferences. For the novice participants who programmed the PCA interfaces, 6 out of 11 participants preferred the new interface over the old, when programming with and without interruptions, while the other five participants' interface preference could not be retrieved from the audiotapes as they were inaudible ($p=0.016$, one-tailed sign test, for participants whose responses were audible). All nursing students reported that they were disrupted by the interruptions. The most common ways in which they reported being negatively affected were in terms of being slowed down, having to concentrate more when programming, making mistakes, being frustrated, flustered and confused, forgetting where they were in the programming task before the interruption occurred, being anxious, having to recheck what was previously programmed, having difficulty finding their place after the interruption, and having difficulty focusing on the programming task. Interestingly, the feedback features of the new interface were not reported to be used as often as hypothesized to minimize the disruptive effects of the interruptions and to assist the participants in reorienting to the interface. Only three out of

seven of the nursing students whose responses were retrievable from the audiotapes reported ever looking at the feedback features.

Experiment 2 was conducted to see whether or not results found for the type of infusion pump used in Experiment 1 (PCA device) would be similar to those obtained for a different type of infusion pump (EPCA device). The results of Experiment 2, in which novice participants programmed the EPCA device, are presented next.

Results and Discussion (Experiment 2)

Programming time. For novice participants programming the EPCA device, significant main effects were found for Mode ($F(2, 20)=12.75$, $p=0.0003$) and Interruption ($F(1, 10)=459.72$, $p<0.0001$), a significant 2-way interaction was found for Mode*Interruption ($F(2,20)=8.74$, $p=0.0019$), and a significant 3-way interaction was found for Interface*Mode*Interruption ($F(2,20)=4.82$, $p=0.0196$). Thus, as shown in Figure 9, interruptions significantly increased programming time overall versus when there were no interruptions for the EPCA device (again, the programming time in interrupted trials includes the time taken to complete the interruption tasks). Although programming time for the old interface (mean=123.18 seconds) was higher than that for the new interface (mean=93.94 seconds), this difference was not significant. A closer look at the significant 3-way Interface*Mode*Interruption interaction in Figures 10, 11 and 12 reveals that for all three modes, interruptions significantly increased programming time for the new EPCA interface compared to when no interruptions were present. However, interruptions significantly increased programming time for the old interface in the continuous+bolus mode only, although the same trend was observed in the other two modes (all 11 participants programmed the new interface faster in uninterrupted trials than interrupted trials ($p<0.001$, one-tailed sign test), and all 11 participants programmed the old interface faster in uninterrupted trials than interrupted trials

($p<0.001$, one-tailed sign test)). Also, when no interruptions were present, the new interface tended to decrease programming time over the old interface for all three modes, although this difference was not significant (9 out of 11 participants programmed the new interface faster than the old ($p=0.033$, one-tailed sign test)). Surprisingly, the new interface appeared to increase programming time over the old interface when interruptions were present for the continuous only and bolus only modes, although these differences were also not significant (9 out of 11 participants programmed the new interface faster than the old when interruptions were present ($p=0.033$, one-tailed sign test)). These results do not support our hypothesis, both in terms of the new interface lessening the effects of interruptions (except in the continuous+bolus mode instance) as well as in terms of results being generalizable from one device to the other (i.e., PCA to EPCA). The fact that some of our findings are not significant could be because both EPCA interfaces are fairly quick to program to begin with (they usually take less than a minute to program when no interruptions are present), and hence a benefit of the new interface over the old may be difficult to detect. The continuous+bolus mode is the longest mode to program, and was the only mode to show a lessening of the disruptive effects of interruptions in the new interface. Also, because of the speed with which both interfaces can be programmed, fewer interruptions could be presented in interrupted trials. Sample size could also be a factor.

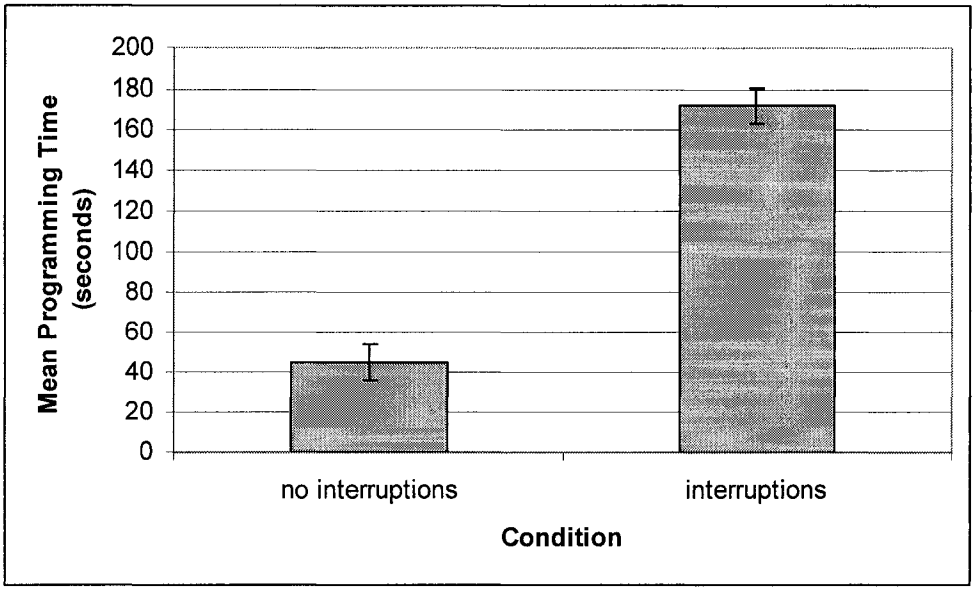


Figure 9. The main effect of interruptions on programming time for novice EPCA users, plotted with 95% confidence limits.

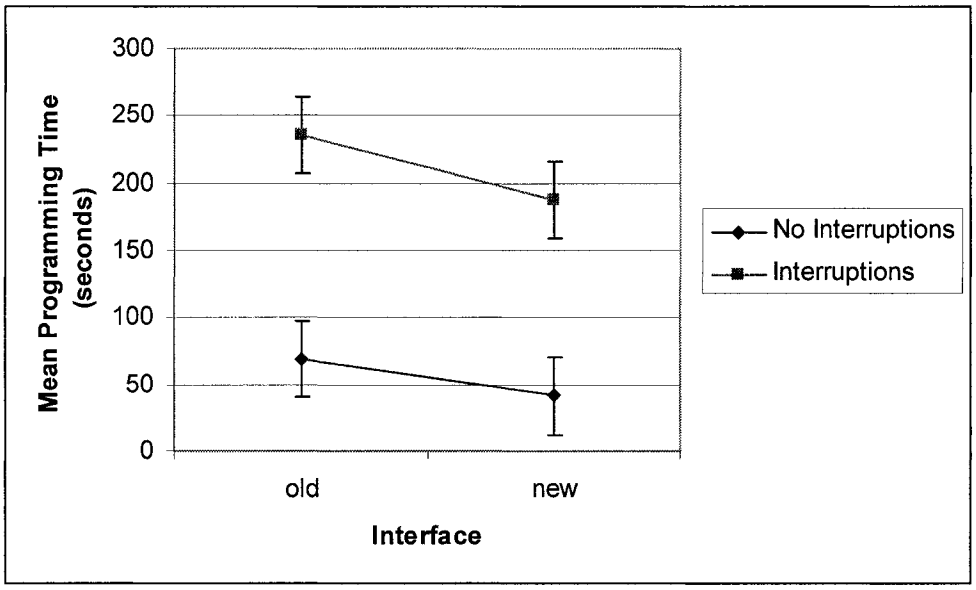


Figure 10. The Interface*Mode*Interruption interaction for novice EPCA users (continuous+bolus mode), plotted with 95% confidence limits.

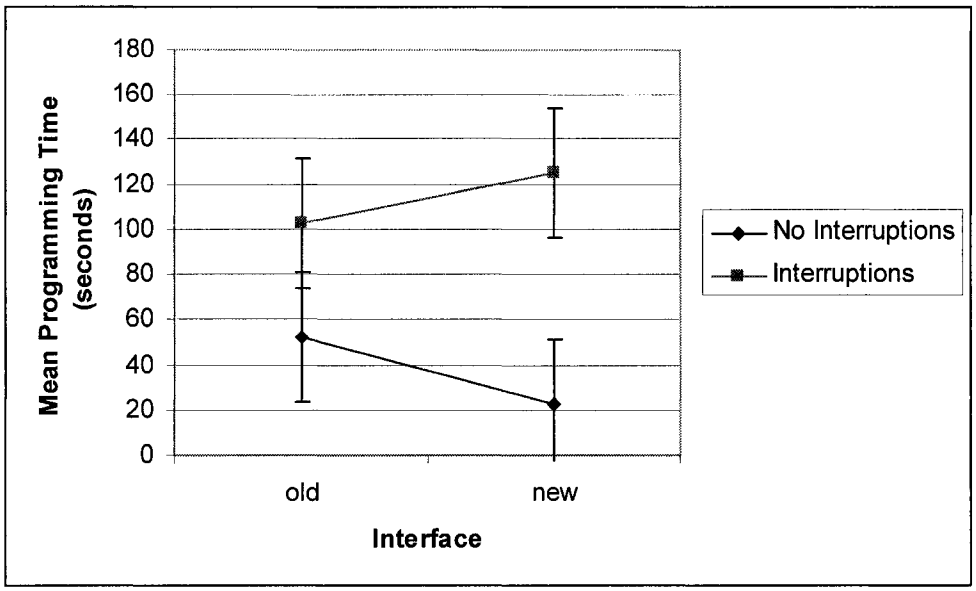


Figure 11. The Interface*Mode*Interruption interaction for novice EPCA users (continuous only mode), plotted with 95% confidence limits.

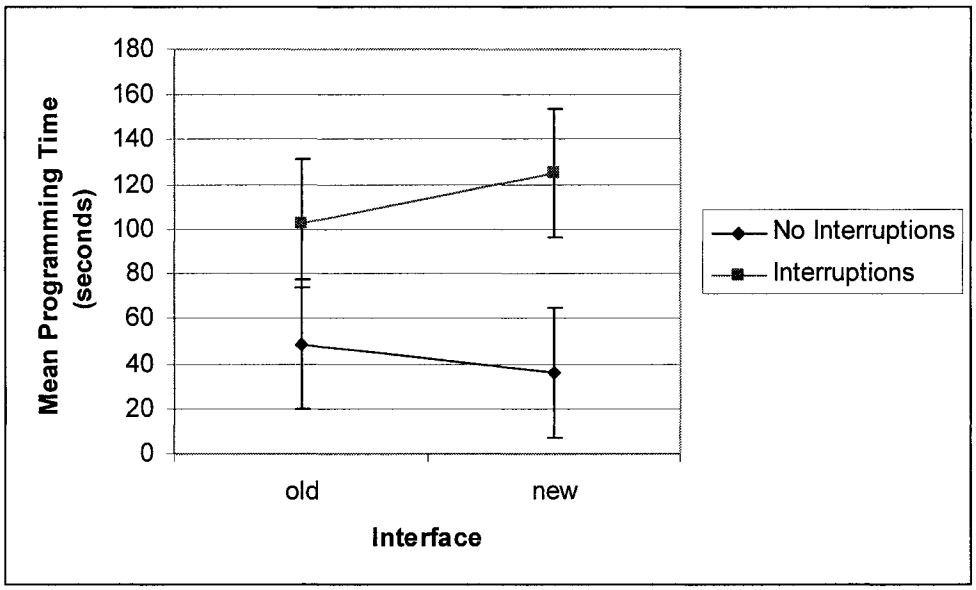


Figure 12. The Interface*Mode*Interruption interaction for novice EPCA users (bolus only mode), plotted with 95% confidence limits.

Workload. For novice participants who programmed the EPCA interfaces, significant main effects for Mode ($F(2,20)=15.27$, $p<0.0001$) and Interruption ($F(1, 10)=36.13$, $p<0.0001$) were found. As shown in Figure 13, interruptions were found to increase workload overall; all 11 participants experienced higher workload when programming the old interface when interruptions were present compared to no interruptions ($p<0.001$, one-tailed sign test), and 10

out of 11 participants experienced the same for the new interface, while one experienced comparable levels of workload with and without interruptions (p=0.001, one-tailed sign test). The new interface was observed to decrease workload over the old interface with and without interruptions, although these differences were not significant. When no interruptions were present, 9 out of 11 participants experienced lower workload when programming the new interface than the old, and one participant experienced comparable levels for both interfaces (p=0.011, one-tailed sign test). When interruptions were present, 7 out of 11 participants experienced lower workload when programming the new interface than the old, and one participant experienced comparable levels of workload when programming both interfaces (not significant with a one-tailed sign test). These findings are similar to those of workload for the PCA interfaces found in Experiment 1.

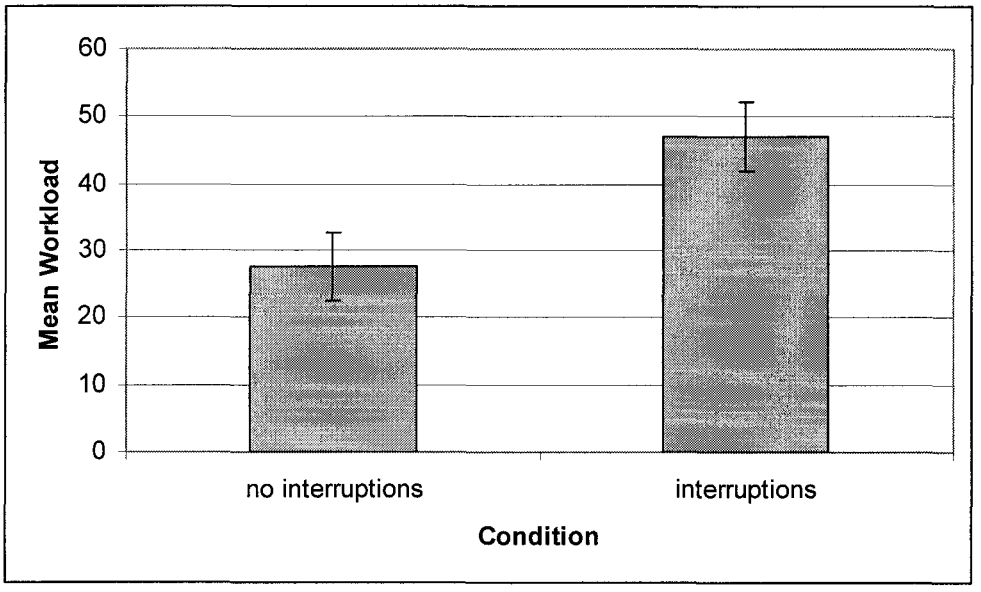


Figure 13. The main effect of interruptions on workload for novice users of the EPCA device, plotted with 95% confidence limits.

Errors. Tables 5 and 6 describe the errors that were committed by the novice participants while programming the old and new EPCA interfaces, respectively. The errors that are highlighted in grey represent uncorrected errors. To summarize these tables, novice participants committed 37 errors on the old EPCA interface (all were corrected), 22 of which occurred during interrupted

trials, and 14 errors on the new EPCA interface (7 were uncorrected), 7 of which occurred during interrupted trials (3 uncorrected errors occurred during interrupted trials). These data seem to support our hypothesis, since interruptions had less of an effect on performance, in terms of errors, for the new interface than the old. However, all errors made with the old interface were corrected, whereas 7 uncorrected errors were committed with the new interface. The majority of these were either errors in choosing the wrong mode for the device to operate in, or entering the wrong value for the total amount parameter. It is unclear why similar errors were not committed with the old interface.

Table 5. Errors committed by novice participants on the old EPCA interface.

Participant	Mode	Interrupted?	Description
16	pca+continuous	yes	entered 'yes' to loading dose prompt
16	pca+continuous	yes	entered a loading dose of 2mL, but entered 'no' to deliver loading dose prompt
16	pca	no	entered 'yes' to loading dose prompt
16	pca	no	entered a loading dose of 4.5mL, but entered 'no' to deliver loading dose prompt
16	pca+continuous	no	entered 'yes' to loading dose prompt
16	pca+continuous	no	entered a loading dose of 2mL, but entered 'no' to deliver loading dose prompt
16	pca	yes	entered 'yes' to loading dose prompt
16	pca	yes	entered a loading dose of 4.5mL, but entered 'no' to deliver loading dose prompt
18	pca+continuous	no	four hour limit error (entered 0mL instead of 50mL), detected in history and eventually corrected through review/change, but made many errors in the process (see next)
18	pca+continuous	no	attempted to correct previous error by typing 50 while in history screen, which had no effect on interface
18	pca+continuous	no	entered 'yes' to loading dose prompt but instead of setting it pressed review/change which started a new program
18	pca+continuous	no	bolus dose error (entered 5mL instead of 2mL), detected in history and corrected through review/change
18	pca+continuous	no	four hour limit error (entered 0mL instead of 50mL), detected in history and corrected through review/change

18	pca	yes	bolus dose error (entered 5mL instead of 4.5mL), realized on next screen and corrected right away using down arrow button
18	pca+continuous	yes	bolus dose error (entered 0mL instead of 2mL), detected in history and eventually corrected using review/change, but made another mistake in the process (see next)
18	pca+continuous	yes	entered 'yes' to loading dose prompt
18	pca+continuous	yes	entered a loading dose of 0mL but entered 'no' to deliver loading dose prompt
20	pca	no	bolus dose error (entered 0.5mL instead of 4.5mL), realized on next screen and corrected right away using down arrow button
20	pca+continuous	yes	bolus lockout error (entered 0min instead of 20min),detected in history and corrected through review/change
20	pca+continuous	no	entered 'yes' to loading dose prompt, but instead of setting it pressed review/change which restarted program
21	continuous	no	mode error (chose bolus only instead of continuous), realized on next screen and corrected right away using down arrow button
21	continuous	no	total amount error (entered 0mL instead of 200mL), detected in history and eventually corrected through review/change but made another mistake in the process (see next)
21	continuous	no	unit error (chose mg/mL instead of mL only), detected and corrected through review/change
23	pca	no	bolus lockout error (entered 5min instead of 25min) realized on next screen and corrected right away using down arrow button
23	pca	yes	bolus dose error (entered 45.45mL instead of 4.5mL), realized on next screen and corrected using down arrow button but made another mistake in the process (see next)
23	pca	yes	bolus dose error (entered 0.5mL instead of 4.5mL), realized on next screen and corrected right away using down arrow button
23	pca+continuous	yes	four hour limit error (entered 5mL instead of 50mL), detected in history and eventually corrected using review/change but made many other errors in the process (see next)
23	pca+continuous	yes	attempted to correct previous error by pressing down arrow button while in history
23	pca+continuous	yes	attempted to correct first error by pressing reset button while in history
23	pca+continuous	yes	attempted to correct first error by pressing down arrow button again while in history

23	pca+continuous	yes	scrolled all the way to the end of the history screens
23	pca+continuous	yes	attempted to correct first error by pressing down arrow button again while in history
23	pca+continuous	yes	scrolled all the way to the end of the history screens
23	pca+continuous	yes	attempted to correct first error by pressing down arrow button again while in history
23	pca+continuous	yes	entered 'yes' to loading dose prompt
23	pca+continuous	yes	entered a loading dose of 2mL but entered 'no' to deliver loading dose prompt
23	pca+continuous	yes	four hour limit error (entered 55mL instead of 50mL), realized on next screen and detected right away using down arrow button

Table 6. Errors committed by novice participants on the new EPCA interface.

Participant	Mode	Interrupted?	Description
13	pca	yes	bolus dose error (entered 45mL instead of 4.5mL), not detected and not corrected
13	pca	no	bolus dose error (entered 45mL instead of 4.5mL), detected on summary screen and corrected using previous screen button
15	pca+continuous	yes	rate error (entered 2mL/hr instead of 5mL/hr), realized on next screen and corrected right away using previous screen button
18	pca+continuous	yes	mode error (chose bolus only instead of both), realized on next screen and corrected right away using previous screen button
20	continuous	yes	total amount error (entered 20mL instead of 200mL), detected in summary screen and corrected using previous screen button
21	continuous	no	total amount error (entered 250mL instead of 200mL), not detected and not corrected
21	pca+continuous	yes	total amount error (entered 0mL instead of 200mL), not detected and not corrected
21	pca	no	device mode error (chose pca instead of epidural), not detected and not corrected
21	pca	no	total amount error (entered 250mL instead of 200mL), not detected and not corrected
21	pca+continuous	no	device mode error (chose pca instead of epidural), not detected and not corrected
21	pca	yes	device mode error (chose pca instead of epidural), not detected and not corrected
22	pca+continuous	no	mode error (chose bolus only instead of both), detected in summary screen and corrected using previous screen button

22	pca+continuous	no	bolus dose error (entered 5mL instead of 2mL), detected in summary screen and corrected using previous screen button
23	pca	yes	total amount error (entered 40mL instead of 200mL), detected in summary screen and corrected using previous screen button

Comments and preferences. For the novice participants who programmed the EPCA interfaces, 9 out of 11 participants preferred the new interface over the old, one participant did not indicate a preference, and one participant felt both interfaces were equally easy-to-use, when programming with and without interruptions (p=0.002, one-tailed sign test). All nursing students reported that they were disrupted by the interruptions, for the same reasons as in Experiment 1. The feedback features of the new EPCA interface were also reportedly not used as often as hypothesized to minimize the disruptive effects of the interruptions and to assist the participants in reorienting to the interface. Only 1 out of the 11 participants reported ever looking at the feedback features.

Experiment 3 was conducted to see whether or not results found for the novice participants in Experiments 1 and 2 would be similar to those obtained for an experienced participant group. The results of Experiment 3, in which experienced nurses programmed the PCA and EPCA device interfaces, are presented next.

Results and Discussion (Experiment 3)

Programming time. For the experienced nurses, significant main effects were found for Device ($F=(1,5)=33.26$, p=0.0022), Interface ($F=(1,6)=10.77$, p=0.0168), and Interruption ($F=(1,6)=55.31$, p=0.0003), significant 2-way interactions were found for Device*Interface ($F=(1,5)=9.92$, p=0.0254) and Device*Mode ($F=(2,10)=4.84$, p=0.0339), and a significant 3-way interaction was found for Device*Interface*Mode ($F(2,10)=7.83$, p=0.009). As shown in Figure 14, the new interface significantly decreased programming time overall compared to the old

interface for the PCA device only, although a similar trend was observed for the EPCA device. Breaking up the significant 3-way interaction by mode revealed that the new EPCA interface significantly reduced programming time over the old EPCA interface only for the continuous mode, although similar trends were observed for the other two modes. As shown in Figure 15, interruptions significantly increased programming time overall compared to when no interruptions were present (for both interfaces of both devices, all participants completed the trials faster when no interruptions occurred (p=0.008, one-tailed sign test (PCA), p=0.016, one-tailed sign test (EPCA))). The new interfaces appeared to decrease programming time over the old interfaces with and without interruptions, although these differences were not significant as the Device*Interface*Interruption interaction was not significant. When no interruptions occurred, all seven participants programming the PCA device were faster with the new interface than the old (p=0.008, one-tailed sign test), and five out of six participants programming the EPCA device were faster with the new interface than with the old, while one participant programmed both at comparable speeds (p=0.031, one-tailed sign test). When interruptions were present, five out of seven participants programming the PCA device were faster with the new interface than the old (not significant with a sign test), and four out of six participants programming the EPCA device were faster with the new interface than with the old (not significant with a sign test). These results do not support the hypothesis that the new interfaces reduced the disruptive effects of interruptions in terms of programming time, but this may be due to a small sample size. However, trends found in this Experiment for the PCA device are comparable to those found in Experiment 1, and these results also seem to favour the EPCA device for the experienced nurses more so than for the nursing students in Experiment 2. The experienced nurses tended to benefit from the new interface of the PCA device versus the old overall. This is important, as they already have much experience in programming the old interface, and they are also used to dealing with interruptions.

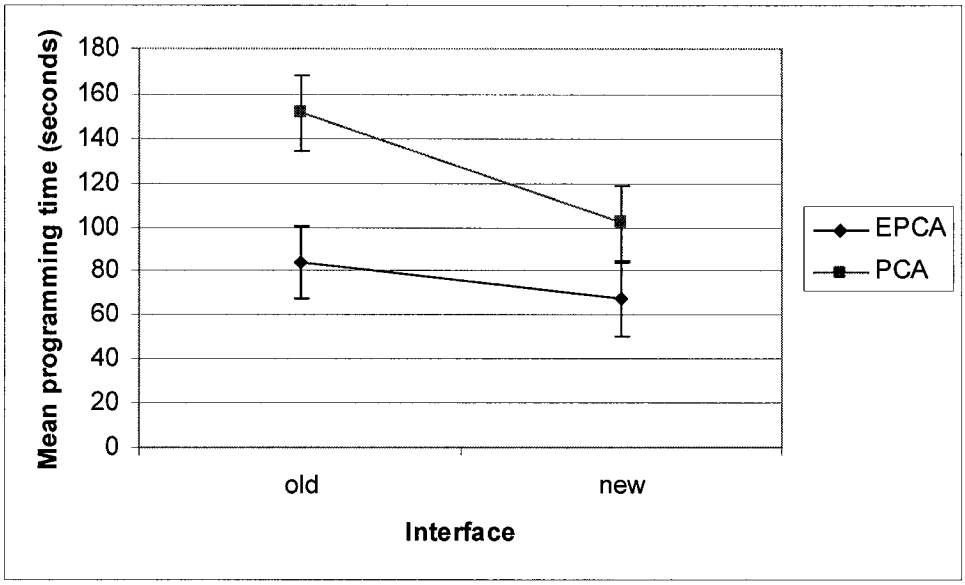


Figure 14. The Device*Interface interaction for experienced nurses, plotted with 95% confidence limits.

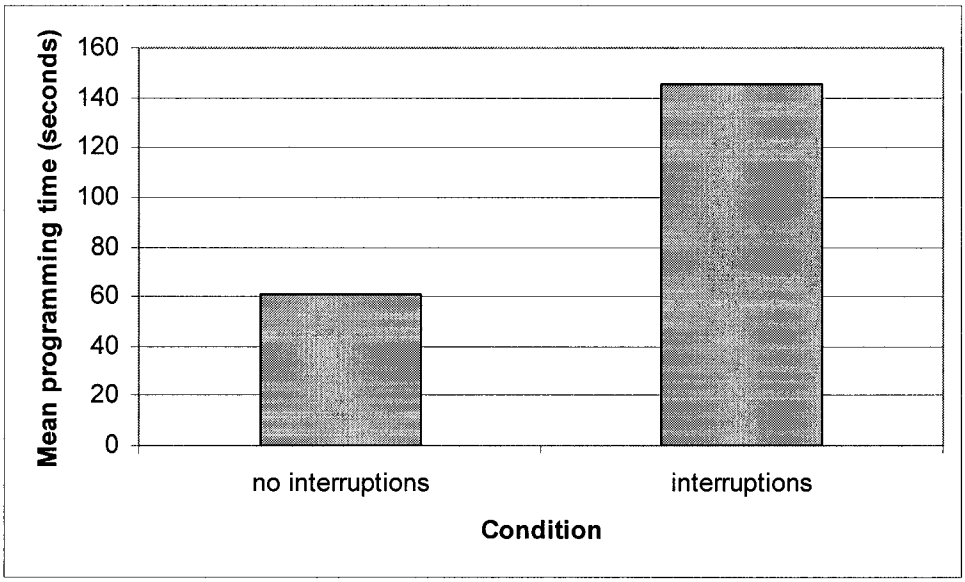


Figure 15. The main effect of interruptions on programming time for experienced nurses (it was not possible to obtain confidence intervals due to missing data (one participant only programmed the PCA device), but results are significant).

Workload. For the experienced nurses, significant main effects were found for Mode ($F(2,12)=5.25$, $p=0.0230$) and Interruption ($F(1,6)=13.68$, $p=0.0101$), and a significant 2-way interaction was found for Interface*Mode ($F(2,12)=5.36$, $p=0.0217$). As shown in Figure 16, interruptions significantly increased workload overall. For the EPCA device, all six participants experienced a higher level of workload when interruptions occurred while programming the old interface versus without interruptions ($p=0.016$, one-tailed sign test). This was also the case for

five out of six participants while programming the new EPCA interface (not significant with a one-tailed sign test). All participants experienced higher levels of workload when programming both interfaces of the PCA device when interruptions occurred, except for one who experienced comparable levels of workload when programming the old interface ($p=0.016$, one-tailed sign test). The workload experienced by participants was lower for the new interfaces of both devices than the old interfaces with and without interruptions, although these differences were not significant. When no interruptions were present, five out of seven participants programming the PCA device experienced lower workload with the new interface than the old, while one experienced comparable levels for both interfaces (not significant with a one-tailed sign test), and four out of six participants programming the EPCA device experienced lower workload with the new interface than with the old, while one experienced comparable levels for both interfaces (not significant with a one-tailed sign test). When interruptions were present, five out of seven participants programming the PCA device experienced lower workload with the new interface than the old (not significant with a one-tailed sign test), and three out of six participants programming the EPCA device experienced lower workload with the new interface than with the old, while one participant experienced comparable levels of workload when programming both interfaces (not significant with a one-tailed sign test). Again, participants rating the workload associated with the interruptions tasks as well as the programming tasks could have confounded results.

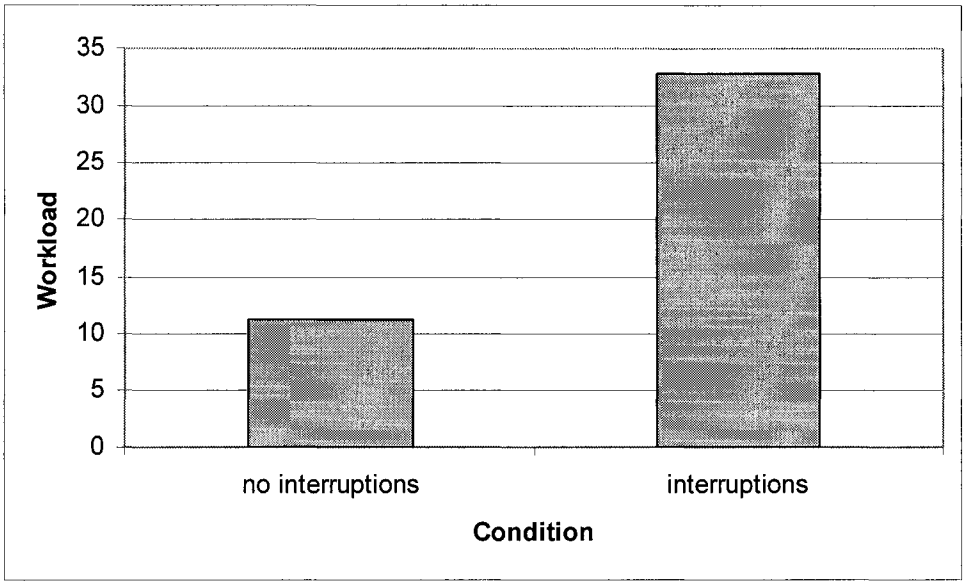


Figure 16. The main effect of interruptions on workload for experienced nurses (it was not possible to obtain confidence limits due to missing data, however results are significant).

Errors. Tables 7 to 10 describe the errors that were committed by the experienced participants while programming the old and new PCA interfaces and the old and new EPCA interfaces, respectively. The errors that are highlighted in grey represent uncorrected errors. For the experienced nurses, two errors were committed on the old PCA interface (all were corrected), both occurring during uninterrupted trials, and eight errors were committed on the new PCA interface (two were uncorrected), all occurring during interrupted trials. Two concentration programming errors were made by the same experienced nurse with the old PCA interface on an uninterrupted trial, due to choosing a default value. Experienced nurses committed 18 errors on the old EPCA interface (12 were uncorrected), 11 of which occurred during interrupted trials (6 uncorrected errors occurred during interrupted trials), and 20 errors on the new EPCA interface (13 were uncorrected), 12 of which occurred during interrupted trials (7 uncorrected errors occurred during interrupted trials). For both EPCA interfaces, a majority of the errors committed by the experienced nurses were total amount errors, and all of the uncorrected errors were total amount errors. One possible cause of this may be that those participants who committed total amount errors may not have read the order sheets properly for that parameter. It could also be

that experienced nurses program 250mL as the total amount more often than 200mL in the PACU, and hence did so in this experiment out of habit. No obvious differences seem to exist between the two EPCA interfaces in terms of errors.

Table 7. Errors committed by experienced participants on the old PCA interface.

Participant	Mode	Interrupted?	Description
31	pca	no	concentration error (chose default option of 1mg/mL morphine instead of entering 2mg/mL), detected in history and corrected through review/change
31	continuous	no	concentration error (chose default option of 1mg/mL morphine instead of entering 1.5mg/mL), detected in history and corrected through review/change

Table 8. Errors committed by experienced participants on the new PCA interface.

Participant	Mode	Interrupted?	Description
27	pca	yes	pca dose error (entered 2mg instead of 1mg), not detected and not corrected
28	pca+continuous	yes	mode error (chose continuous instead of pca+continuous), detected in history and eventually corrected but made many errors in the process (see next)
28	pca+continuous	yes	attempted to correct previous error by pressing review button, but pressed enter and got back to summary screen
28	pca+continuous	yes	mode error again (chose continuous instead of pca+continuous), detected in summary screen and eventually corrected but made more errors in the process (see next)
28	pca+continuous	yes	attempted to correct previous error by pressing review button, but pressed enter and got back to summary screen
28	pca+continuous	yes	rate error (entered 2.5mg/hr instead of 1mg/hr), detected in summary screen and corrected eventually using review button but made another error first (see next)
28	pca+continuous	yes	attempted to correct previous error by pressing review button, but pressed enter and got back to summary screen
29	pca+continuous	yes	pca dose error (entered 2.5mg instead of 1.5mg), not detected and not corrected

Table 9. Errors committed by experienced participants on the old EPCA interface.

Participant	Mode	Interrupted?	Description
26	continuous	no	total amount error (entered 250mL instead of 200mL), not detected and not corrected
26	pca	yes	total amount error (entered 250mL instead of 200mL), not detected and not corrected
26	pca+continuous	no	total amount error (entered 250mL instead of 200mL), not detected and not corrected
26	continuous	yes	total amount error (entered 250mL instead of 200mL), not detected and not corrected
26	pca	no	total amount error (entered 250mL instead of 200mL), not detected and not corrected
26	pca+continuous	yes	total amount error (entered 250mL instead of 200mL), not detected and not corrected
27	continuous	no	total amount error (entered 0mL instead of 200mL), detected in history and corrected using review/change
29	continuous	yes	rate error (entered 0mL/hr instead of 5mL/hr), detected in history and eventually corrected using review/change but made another error in the process (see next)
29	continuous	yes	started a new program in review/change instead of just changing the old one
30	pca+continuous	no	total amount error (entered 250mL instead of 200mL), not detected and not corrected
30	continuous	yes	total amount error (entered 250mL instead of 200mL), not detected and not corrected
30	pca	no	total amount error (entered 250mL instead of 200mL), not detected and not corrected
30	pca+continuous	yes	total amount error (entered 250mL instead of 200mL), not detected and not corrected
30	continuous	no	total amount error (entered 250mL instead of 200mL), not detected and not corrected
30	pca	yes	four hour limit error (entered 2mL instead of 40mL), realized on next screen and corrected right away through review/change but made more errors in the process (see next)
30	pca	yes	tried to correct previous error using down arrow button but was unsuccessful
30	pca	yes	pressed review/change to correct first error but started new program instead of changing old one
30	pca	yes	total amount error (entered 250mL instead of 200mL), not detected and not corrected

Table 10. Errors committed by experienced participants on the new EPCA interface.

Participant	Mode	Interrupted?	Description
26	continuous	no	total amount error (entered 250mL instead of 200mL), not detected and not corrected
26	pca	yes	mode error (chose both instead of bolus only), realized on next screen and detected right away using previous screen button
26	pca	yes	total amount error (entered 250mL instead of 200mL), not detected and not corrected
26	pca+continuous	no	total amount error (entered 250mL instead of 200mL), not detected and not corrected
26	continuous	yes	total amount error (entered 250mL instead of 200mL), not detected and not corrected
26	pca	no	total amount error (entered 250mL instead of 200mL), not detected and not corrected
26	pca+continuous	yes	total amount error (entered 250mL instead of 200mL), not detected and not corrected
27	pca	yes	mode error (chose both instead of bolus only), not detected and not corrected (therefore rate was never set)
27	pca+continuous	yes	bolus lockout error (entered 2min instead of 20min), realized on next screen and eventually corrected using previous screen button but made mistakes in the process (see next)
27	pca+continuous	yes	pressed history button while trying to correct previous error
27	pca+continuous	yes	pressed up arrow on bolus dose screen while trying to correct previous error
30	pca+continuous	no	total amount error (entered 250mL instead of 200mL), not detected and not corrected
30	continuous	yes	total amount error (entered 250mL instead of 200mL), not detected and not corrected
30	pca	no	total amount error (entered 250mL instead of 200mL), not detected and not corrected
30	pca+continuous	yes	total amount error (entered 250mL instead of 200mL), not detected and not corrected
30	continuous	no	total amount error (entered 250mL instead of 200mL), not detected and not corrected
30	pca	yes	total amount error (entered 250mL instead of 200mL), not detected and not corrected
31	pca	yes	bolus dose error (entered 45mL instead of 4.5mL), detected on summary screen and corrected using previous screen button
31	pca+continuous	no	rate error (entered 2mL/hr instead of 5mL/hr), realized right away and corrected right away using previous screen button

31	pca+continuous	no	bolus dose error (entered 20mL instead of 2mL), realized right away and corrected right away using previous screen button

Comments and preferences. For the experienced nurses, five out of six participants preferred the new EPCA interface over the old with and without interruptions, and one preferred the old interface (not significant with a sign test). Similarly, six out of seven experienced nurses preferred the new PCA interface over the old with and without interruptions, and one preferred the old interface (not significant with a sign test). Most nurses reported being disrupted by interruptions in terms of being slowed down and having to find their place after an interruption. For the experienced nurses, only one out of five participants reported ever looking at the feedback feature of the new EPCA interface (one participant's response was inaudible), while three out of seven participants reported ever looking at the feedback feature of the new PCA interface.

General Discussion

To reiterate, the main question of interest that drove this research was whether or not the human factors improvements in the new interfaces would significantly reduce the disruptive effects of interruptions. It was hypothesized that this would be the case in general, and specifically that the feedback features of both of the new interfaces would help users to know where they were in the programming sequence before an interruption occurred, and facilitate their reorientation to the interface and the primary programming task following an interruption.

First, it should be noted that the findings of Lin et al. (1998, 2001) and Ford and Rollinson (2001) were not entirely reproduced in this study. For example, with novice participants programming the EPCA device, Ford and Rollinson found a significant reduction in programming time for the new interface versus the old, whereas the reduction in this study was not significant (Experiment 2). The same difference occurs between nursing students and

experienced nurses programming the PCA device in this study (Experiments 1 and 3) and in Lin et al.'s study. Differences also exist in the error results. One reason for the differences in results could be a difference in the number of participants. Lin et al. had 12 novice participants and 12 experienced participants for the PCA device, and Ford and Rollinson had 12 novice participants for the EPCA device. In contrast, there were 11 participants in Experiments 1 and 2 of this study, and only 7 participants in Experiment 3. Another reason could be that the exact parameters to be programmed in each task were different between the studies. Lin et al. and Ford and Rollinson did not explicitly state which parameter values they used. Furthermore, a within-subject design was used for all factors for the experienced nurses in this study (Experiment 3), whereas the experienced nurses in Lin et al.'s study only programmed the PCA device, and hence this difference in study designs could have caused a difference in results. A difference in the level of programming experience of the experienced nurses could also have affected the results. If, for example, the average number of years of experience of the participants used in this study with the existing interfaces for the PCA and EPCA devices was higher than that of Lin et al's study, then performance with the old interfaces may have been better, making it more difficult to find a significant benefit to the new interfaces. Finally, the specific instructions given to participants could have had an impact. In this study, participants were instructed to program the interfaces as fast and accurately as possible while answering the chart questions immediately and as accurately as possible during interrupted trials. It is possible that the results could have been affected if participants were anticipating interruptions throughout their programming tasks.

The programming time results of this study generally seem to favour the new interface of the PCA device over the old interface, with and without interruptions, for nursing students and experienced nurses. All participants were disrupted by the interruptions, as expected. In the cases where a favourable trend for the new interfaces was non-significant, a small number of

participants could be the cause of the lack of significance, since the nonparametric sign tests showed that a majority of subjects performed better with the new interfaces than with the old ones. It is important to realize when interpreting the data that programming time is directly influenced by the errors that are made, as having to correct an error obviously increases programming time. The workload results tended to favour the new interfaces of both devices for both groups of participants, with and without interruptions, although with the method we used to elicit participants' subjective perceptions of workload, we were not able to distinguish between workload associated with the interfaces and workload associated with the interruption tasks.

The error results appear to be far more complex. For novice participants programming the PCA device, interruptions appeared to increase the total number of errors made. In fact, interruptions seemed to affect both interfaces to the same degree. However, more uncorrected errors were made with the new interface than the old, all occurring during interrupted trials. This may seem surprising in light of the findings of Lin et al. (1998). When put into context of the actual features of the new PCA interface, though, this result is not surprising at all. One such feature is the summary screen that is automatically displayed on the interface after the last parameter has been programmed. The fact that it automatically appears is safer and more efficient, as users do not have to press a button to bring it up, which they may forget to do. However, the summary screen disappears automatically as well, after a few seconds. It was observed during interrupted trials that sometimes a participant would be interrupted right when the summary screen appeared, and when they returned their attention to the interface after completing the interruption task (in the case of task-switching), the summary screen would have disappeared. Some participants subsequently forgot to press History to review their settings, and hence did not detect errors they had made, causing these errors to be uncorrected. This point illustrates the need to conduct realistic experiments, because under uninterrupted conditions,

having the summary screen automatically disappear is more efficient, yet this is very dangerous under interruptive conditions.

A majority of the errors made with the new interface were mode errors, whereas no mode errors were committed with the old interface. It could be that because the nursing students had no experience in programming infusion devices, they were not used to determining the correct programming mode from order sheets. Having each mode option presented on separate screens in the old interface may force the participants to be sure of the mode that was ordered, whereas all modes presented on one screen in the new interface may cause participants to make a quicker, less accurate decision. For the experienced nurses programming the PCA device, all errors with the new interface occurred during interrupted trials, and again the summary screen disappearing automatically caused some errors to be uncorrected. Several of the errors made with the new interface were committed while trying to correct previous errors, but it is important to note that all of these errors were made by the same participant, indicating that this inability to efficiently correct errors may be attributed to an individual difference. Reducing individual differences in future experiments may be achieved by giving participants a more structured and standardized training session with the interfaces, although possibly at the expense of representativeness because nurses usually receive minimal training on infusion devices in hospitals.

Although the feedback features (i.e., the displays that show the current programming step and the current position in the programming sequence) of both of the new interfaces were not specifically designed by Lin et al. (1998) and Ford and Rollinson (2001) for interruptions handling purposes, it was hypothesized at the outset of the study that these features would reduce the disruptive effects of interruptions. It was surprising, then, to find that very few participants reported using them. Perhaps more training with these features would have caused them to be used more often. Another possibility is that participants subconsciously used them and therefore did not report doing so, but only eye-tracking experiments would be able to determine if this was

the case. More research needs to be done in this area, and in determining what information is the most useful to users who are recovering from an interruption, and what is the best way to display this information in an interface.

Several other phenomena observed during the experiments are worth mentioning. First, when resuming a programming task after an interruption that resulted in task-switching, participants took several seconds to reorient themselves to the programming task, during which time they usually looked between the interface and the order sheet several times. Reorientation to a primary task after an interruption is a common finding in the literature (Burmistrov and Leonova, 1997, 2003; Trafton et al., 2003). It would be interesting in follow-up experiments to compare the reorientation time between the new and old interfaces for each device. It would also be worthwhile to break down the programming time measured into several parts, such as the actual time to program the interface, the time to attend to the interruption (which could be measured as the time between when the question is asked and when the answer is given), the reorientation time (which could be measured as the time between when the answer to the interruption question is given by the participant and when the next action is performed on the interface), and the time taken to switch to the interruption task (similar to Burmistrov and Leonova, 1997, 2003), but to do so, task-switching would have to be enforced for all interruptions.

Another interesting observation was that despite explicit instructions to attend to interruptions immediately, participants nevertheless still employed certain coping strategies when faced with an interruption. These strategies included hurrying to finish programming the current subtask, or hurrying to complete the current programming trial if they were close to the end. Participants admitted to using these strategies in the interviews following their sessions, and said they felt that if they were to leave the interface in the middle of a step, they would surely either take longer to reorient to the interface after the interruption, and/or make a mistake.

Several other researchers have also observed similar strategies in their experiments (Burmistrov and Leonova, 1997; Eyrolle and Cellier, 2000; Zijlstra et al., 1999).

CONCLUSION

The contributions of this study are far reaching and have implications in nursing, interruptions research, and interface design and evaluation. Previously, no one has looked specifically at what types of interruptions occur in a hospital recovery room, nor at whether interfaces designed specifically by human factors principles can decrease the disruptive effects of interruptions, relative to existing interfaces. This study has also shown the value of conducting representative experiments. Whenever interfaces are used in interruptive workplaces, they should be designed to help users handle interruptions and minimize their negative effects. Furthermore, interfaces should be evaluated under realistic conditions, for results to be generalizable from the laboratory setting to the actual work environment. This is paramount when designing and evaluating safety critical systems, such as medical technologies, as lives are at stake if errors are made. The results from the field study portion of this project can be used to design representative usability experiments to test many other devices used in the nursing workplace. Furthermore, the methods employed in this field study can serve as a benchmark for future studies to investigate interruptions in other health care areas or other work domains in which interruptions exist but have not yet been studied.

Results of the experiments support findings in the literature on human performance in human-computer interaction when interruptions are present, and interesting phenomena that other researchers have observed were also reproduced in this study. Specifically, we found that interruptions negatively affected our participants in terms of programming time, workload, and errors, and that participants adopted coping strategies to handle the interruptions. Our hypothesis that the new interfaces would reduce the disruptive effects of interruptions relative to the old interfaces, and that results would be consistent across devices and participant groups was only partially confirmed. Limitations that existed in these experiments are described next, followed

by recommendations as to how to make the new interfaces more robust in the face of interruptions.

Limitations

There were some limitations to these experiments. Because we were using computer simulations of the interfaces, the experiments were not completely realistic. Thus, clinical trials would need to be performed to test for generalizability. Another limitation is that small sample sizes were used in this experiment, which could have affected the results. Furthermore, interfaces for two patient-controlled analgesia devices were investigated in this study, and so results obtained for these devices may differ from results that would be obtained from experiments with other types of infusion pumps and other medical devices. This should also be investigated. In addition, other methods of measuring workload or cognitive demand that tasks place on users should be explored, as we believe that the workload caused by the interruption tasks themselves may have confounded our results. Also, the programming time measured in these experiments included the time taken to complete the interruption tasks, as participants used both task-switching and task-sharing strategies. Future work should try to enforce only task-switching, as it would be interesting to see if the new interfaces reduce the disruptive effects of interruptions when only this method is used. Finally, only one frequency of interruption was used in this study, as it was the first of its kind and so it was meaningful to employ a worst-case scenario. Further research should be conducted on how the frequency of interruption affects performance in this work domain.

Recommendations

Based on the results of this study, it is recommended that medical devices be customizable to match hospital unit protocols, so that inapplicable steps and options, and hence

avoidable errors, can be eliminated. Thus, hospital-specific templates should be available for programmable devices. For the University Health Network PACU in which both the PCA and EPCA devices are used, several things should be customized in the new interfaces:

- Only one operating mode is ever programmed for each device in the PACU, the other modes should be removed from the programming options.
- Only mL units are used for the EPCA device in the PACU, so the other unit options should be removed.
- PACU nurses do not purge the device while programming, so this step should be removed from the programming sequence.
- PACU nurses do not deliver bolus doses while programming the devices, so this step should be removed from the programming sequence.
- The EPCA device is only ever programmed in epidural mode, not pca mode, therefore this option should be removed. This is particularly confusing for users, as one of the modes within the epidural device mode is patient-controlled (called 'bolus only' on the interface), which is referred to as 'epidural pca' on the order sheet. The order sheet should also be revised to match the interface's terminology.

There are a few more potential areas of improvement that were revealed through this study. For instance, many participants confused the Yes and Enter buttons of the old EPCA interface. These should be combined into a single button for increased programming efficiency. Also, while the Review button on the new PCA interface is very useful, the label could perhaps be changed to Back, or Previous Screen as it is labelled on the new EPCA interface, to be even clearer. Furthermore, there are two summary screens for the new EPCA interface. The rate parameter is listed on the first screen, while all other value-based parameters are listed on the second screen. The rate parameter should be listed with the others as it was observed in the

experiments that some participants looked mainly at the second summary screen and were confused as to why the rate was not shown there for modes involving continuous infusion. As previously mentioned, the summary screen of the PCA device should not disappear automatically (as error data show that this can cause errors to go undetected and thus uncorrected), but should disappear only once the user presses a button to acknowledge that they have checked their settings. In addition, the error data also reveal that many participants made substitution errors in interrupted trials. That is, for a particular parameter, they programmed in a value that appeared on the order sheet for another parameter. This occurred far more frequently with the PCA device interfaces, all occurring during interrupted trials. To reduce the likelihood of this type of error occurring, it is recommended that the parameters on the PCA order sheet be spread out more, as they are on the EPCA order sheet. Finally, the open-concept layout of the PACU itself could be conducive to unnecessary interruptions, as nurses can see and hear each other at all times, and so disruptive conversation not pertinent to effective patient care can take place. Future research should look at the PACU layout to determine if the workplace can be redesigned to reduce the occurrence of unnecessary interruptions, while still allowing the unit to function efficiently and in accordance with resource constraints. Future research should also look into methods of making patient information, doctor's orders, and test results accessible to all health care professionals, so that interruptions of this nature can also be reduced. Examples of such methods may be electronic patient charts and point-of-care devices.

REFERENCES

Altmann, E., and Trafton, J. (2002). Memory for goals: an activation-based model. Cognitive Science, 26, 39-83.

ANSI/AAMI. (2001). Human factors design process for medical devices. HE74.

Bailey, B.P., Konstan, J. A., and Carlis, J. V. (2000). Measuring the effects of interruptions on task performance in the user interface. In Proceedings of the 2000 IEEE Conference on Systems, Man, and Cybernetics (pp. 757-762). Los Alamitos, CA: IEEE Press.

Bailey, B.P., Konstan, J. A., and Carlis, J. V. (2001). The effects of interruptions on task performance, annoyance, and anxiety in the user interface. In Proceedings of Human Computer Interaction INTERACT '01 (pp. 593-601). Orlando, FL: Elsevier Science Publishers.

Baker, J. A. (2000). A pathway to elimination of drug and medial errors: a study of the impact on medication error reduction related to Pyxis products. San Diego, CA: Pyxis Corporation.

Blegen, M. A., Goode, C., and Reed, L. 1998. Nurse staffing and patient outcomes. Nursing Research, 47, 43-50.

Blendon, R. J., DesRoches, C. M., Brodie, M., Benson, J. M., Rosen, A., B., Schneider, E., Altman, D. E., Zapert, K., Herrmann, M. J., and Steffenson, A. E. (2002). Views of practicing physicians and the public on medical errors. New England Journal of Medicine, 347, 1933-1940.

Biordi, D. L. (1993). Nursing error and caring in the workplace. Nursing Administration Quarterly, 17, 38-45.

Burmistrov I., and Leonova A. (1997). Interruptions in the computer aided office work: implications to user interface design. In From Experience to Innovation: Proceedings of the 13th Triennial Congress of the International Ergonomics Association, Vol. 5 (pp. 77-

79). Tampere, Finland: Finnish Institute of Occupational Health.

Burmistrov I., and Leonova A. (2003). Do interrupted users work faster or slower? The micro-analysis of computerized text editing task. In Human-Computer Interaction: Theory and Practice (Part I) - Proceedings of HCI International 2003, Vol. 1 (pp. 621-625). Mahwah, NJ: Lawrence Erlbaum Associates.

Cellier, J-M., and Eyrolle, H. (1992). Interference between switched tasks. Ergonomics, 35, 25-36.

Chisholm, C., Collison, E. K., Nelson, D. R., Cordell, W. H. (November 2000). Emergency department workplace interruptions: are emergency physicians "interrupt-driven" and "multitasking"? Academic Emergency Medicine, 7, 1239-1243.

Coiera, E. and Tombs, V. (1998). Communication behaviors in a hospital setting: An observational study. British Medical Journal, 316, 673-6.

Cooper, J.B., Newbower, R. S., Long, C. D., and McPeek, B. (1978). Preventable anesthesia mishaps: A study of human factors. Anesthesiology, 49, 399-406.

Cutrell, E. Czerwinski, M., and Horvitz, E. (2001). Notification, disruption and memory: Effects of messaging interruptions on memory and performance. In Proceedings of Human-Computer Interaction INTERACT '01 (pp. 263-269). Orlando, FL: Elsevier Science Publishers.

Davis, N. (1990). Detection and prevention of ambulatory care pharmacy dispensing errors. Hospital Pharmacy, 25, 18-28.

Davis, N. (1994). Concentrating on interruptions. American Journal of Nursing, 94(3), 14.

Diez, M., Boehm-Davis, D. A., and Holt, R. W. (2002). Model-based predictions of interrupted checklists. In Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting (pp. 250-254). Santa Monica, CA: Human Factors and Ergonomics Society.

ECRI. (2002). Medication safety: PCA pump programming errors continue to cause fatal

overinfusions. Health Devices, 31(9), 342-347.

Edwards, M., and Gronlund, S. (1998). Task interruption and its effects on memory. Memory, 6, 665-687.

Eyrolle, H., and Cellier, J-M. (2000). The effects of interruptions in work activity: Field and laboratory results. Applied Ergonomics, 31, 537-543.

Field, G. (1987). Experimentus interruptus. ACM SIGCHI Bulletin, 19(2), 42-46.

Flynn, E. A., Dorris, N. T., Holman, G. T., Carnahan, B. J., and Barker, K. N. (2002). Medication dispensing errors in community pharmacies: A nationwide study. In Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting (pp. 1448-1451). Santa Monica, CA: Human Factors and Ergonomics Society.

Flynn, F., Mohr, L., and Lawlor-Klean, P. (2003). Right programming of pumps to prevent errors in the infusion process. Joint Commission Journal on Quality and Safety, 29(1), 37-39.

Ford, L., and Rollinson, A. (2001). Interface redesign for a patient-controlled epidural analgesia device. Unpublished B.A.Sc. Thesis, University of Toronto, Toronto, Ontario.

Frank, J. L., Daniels, J. J., and McFarlane, D. C. (2002). Recovering context after interruption. In Proceedings of the 24th Annual Meeting of the Cognitive Science Society (pp. 310-315). Austin, TX: Cognitive Science Society.

Gillie, T., and Broadbent, D. (1989). What makes interruptions disruptive? A study of length, similarity and complexity. Psychological Research, 50, 243-250.

Gladstone, J. (1995). Drug administration errors: A study into the factors underlying the occurrence and reporting of drug errors in a district general hospital. Journal of Advanced Nursing, 22, 628-637.

Hart, S. G., and Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index):

Results of empirical and theoretical research. In Hancock, P. A., and Meshkati, N. (Eds.), Human mental workload. Amsterdam, North Holland: Elsevier Science Publishers.

Hodgetts, H. M., and Jones, D. M.. (2003). Interruptions in the tower of London task: can preparation minimise disruption? In Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting (pp. 1000-1004). Santa Monica, CA: Human Factors and Ergonomics Society.

Horvitz, E., Jacobs, A., and Hovel, D. (1999). Attention-sensitive alerting. In Proceedings of UAI '99 Conference on Uncertainty and Artificial Intelligence (pp. 305-313). San Francisco, CA: Morgan Kaufmann.

Kaye, R., and Crowley, J. (2000). Medical device use-safety: Incorporating human factors engineering into risk management. Rockville, MD: U.S. Department of Health and Human Services, Food and Drug Administration.

Kreifeldt, J., and McCarthy, M. (1981). Interruption as a test of the user-computer interface. In Proceedings of the 17th Annual Conference on Manual Control (pp. 655-67). JPL Publication 81-95.

Kohn, L. T., Corrigan, J. M., and Donaldson, M. S. (Eds.). (2000). To err is human: Building a safer health system. Washington, DC: National Academy Press.

Latorella, K. (1996). Investigating interruptions: An example from the flightdeck. In Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting (pp. 249-253). Santa Monica, CA: Human Factors and Ergonomics Society.

Levy, J., Gopher, D., and Yoel, D. (2002). An analysis of work activity in the operating room: Applying psychological theory to lower the likelihood of human error. In Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting (pp. 1457-1461). Santa Monica, CA: Human Factors and Ergonomics Society.

Lin, L., Isla, R., Doniz, K., Harkness, H., Vicente, K. J., and Doyle, D. J. (1998). Applying

human factors to the design of medical equipment: patient-controlled analgesia. Journal of Clinical Monitoring and Computing, 14, 253-263.

Lin, L., Vicente, K. J., and Doyle, D. J. (2001). Patient safety, potential adverse drug events, and medical device design: A human factors engineering approach. Journal of Biomedical Informatics, 34, 274-284.

Loftus, G. R., and Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. Psychonomic Bulletin & Review, 1, 476-490.

McConnell, E. A. (1998). Medical device use by nurses: A review of published nursing literature. Biomedical Instrumentation & Technology, 32, 471-478.

McFarlane, D. C. (1997). Interruption of people in human-computer interaction: A general unifying definition of human interruption and taxonomy (Tech. Report NRL/FR/5510-97-9870). Washington, DC: US Naval Research Lab.

McFarlane, D.C. (1998). Interruption of people in human-computer interaction. Unpublished doctoral dissertation, George Washington University, Washington, DC.

McFarlane, D. C. (1999). Coordinating the interruption of people in human-computer interaction. In Proceedings of Human-Computer Interaction INTERACT '99 (pp. 295-303). Orlando, FL: Elsevier Science Publishers.

McFarlane, D. C. & Latorella, K.A. (2002). The scope and importance of human interruption in HCI design. Human Computer Interaction, 17, 1-62.

McGillis-Hall, L., and Doran, D. I. (2001). A study of the impact of nursing staff mix models and organizational change strategies on patient, system and nurse outcomes. Unpublished report, University of Toronto, Toronto, Ontario.

Miller, S. L. (2002). Window of opportunity: Using the interruption lag to manage disruption

in complex tasks. In Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting (pp. 245-249). Santa Monica, CA: Human Factors and Ergonomics Society.

Monk, C. A., Boehm-Davis, D. A., and Trafton, J. G. (2002). The attentional costs of interrupting task performance at various stages. In Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting (pp. 1824-1828). Santa Monica, CA: Human Factors and Ergonomics Society.

Nagata, S. F. (2003). Multitasking and interruptions during mobile web tasks. In Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting (pp. 1341-1345). Santa Monica, CA: Human Factors and Ergonomics Society.

Nicklin, W., and Mcveety, J. E. (2002). Canadian nurses' perceptions of patient safety in hospitals. Canadian Journal of Nursing Leadership, 15(3), 11-21.

O'Brien-Pallas, L., Thomson, D., Alksnis, C., and Bruce, S. (2001). The economic impact of nurse staffing decisions: Time to turn down another road? Hospital Quarterly, Spring, 42-50.

O'Shea, E. (1999). Factors contributing to medication errors: A literature review. Journal of Clinical Nursing, 8, 496-504.

Oulasvirta, A., and Saariluoma, P. (2004). Long-term working memory and interrupting messages in human-computer interaction. Behaviour & Information Technology, 23, 53-64.

Pawlak, W., and Vicente, K. J. (1996). Inducing effective operator control through ecological interface design. International Journal of Human-Computer Studies, 44, 653-688.

Paxton, F., Heaney, D. J., Howie, J. G., Porter, A. M. (1996). A study of interruption rates for practice nurses and GPs. Nursing Standard, 10(43), 33-36.

Segatore, M., Miller, M, and Webber, K. (1994). Medication out of control? The

Canadian Nurse, 90(8), 35-39.

Siegel, S. (1956). Nonparametric statistics for the behavioral sciences. Toronto, ON: McGraw-Hill Book Company.

Storch, N. A. (1992). Does the user interface make interruptions disruptive? A study of interface style and form of interruption (Tech. Report UCRL-JC-108993). Springfield: Lawrence Livermore National Laboratory.

Trafton, J. G., Altmann, E. M., Brock, D. P., and Mintz, F. E. (2003). Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal. International Journal of Human-Computer Studies, 58, 583-603.

Vicente, K. J., Kada-Bekhaled, K., Hillel, G., Cassano, A., and Orser, B. A. (2003). Programming errors contribute to death from patient-controlled analgesia: Case report and estimate of probability. Canadian Journal of Anesthesia, 50, 328-332.

Wakefield, B.J., Wakefield, D.S., Uden-Holman, T., and Blegen, M.A. (1998). Nurses' perceptions of why medication administration errors occur. Medical-Surgical Nursing Journal, 7(1), 39-44.

Walters, J. (1992). Nurses' perceptions of reportable medication errors and factors that contribute to their occurrence. Applied Nursing Research, 5, 86-88.

Wiklund, M., Smith, E., and Baker, M. (2002). Pilot comparison usability test of intravenous infusion pumps. Unpublished report, National Center for Patient Safety Department of Veterans Affairs, Ann Arbor, MI.

Williams, A. (April, 1996). How to avoid mistakes in medicine administration. Nursing Times, 92(13), 40-41.

Zijlstra, F. R. H., Roe, R. A., Leonora, A. B., Krediet, I. (1999). Temporal factors in mental work: Effects of interrupted activities. Journal of Occupational and Organizational Design, 72, 163-185.

APPENDIX 1

Toronto General Hospital
University Health Network

CONSENT FORM

TITLE: A study of the environment in which recovery room nurses work.

INVESTIGATORS: Dr. Kim J. Vicente (1-617-253-5624; kjv@mit.edu)- Faculty Supervisor
Ms. Gillian Hillel (416-978-0881; hillel@mie.utoronto.ca -Investigator

You are being asked to take part in a research study. Before agreeing to participate in this study, it is important that you read and understand the following explanation of the proposed study procedures. The following information describes the purpose, procedures, benefits, discomforts, risks and precautions associated with this study. It also describes your right to refuse to participate or withdraw from the study at any time. In order to decide whether you wish to participate in this research study, you should understand enough about its risks and benefits to be able to make an informed decision. This is known as the informed consent process. Please ask the investigator to explain any words you don't understand before signing this consent form. Make sure all your questions have been answered to your satisfaction before signing this document.

Purpose
You have been asked to participate in a study that is designed to assess the environment that nurses work in. This information will aid in the design and evaluation of interfaces for medical devices that are more user-friendly, reduce human error, and enhance patient safety. The aim is to study the environment in which you work, not evaluate you or your performance.

Procedures
As a participant in this study, the environment in which you perform your regular tasks will be observed.

Risks
There are no risks involved in this study.

Benefits
Information learned from this study may benefit other nurses and patients in the future as it will be used to ensure that interfaces for medical devices are designed to account for the complex working conditions in which they are used. This information will also be used to design more realistic experiments for evaluating interfaces. This may ultimately reduce human error and enhance patient safety.

Toronto General Hospital
University Health Network

Confidentiality
All information obtained during the study will be held in strict confidence. You will be identified with a study number only. No names or identifying information will be used in any publication or presentations. No information identifying you will be transferred outside the investigators in this study.

Participation
Your participation in this study is voluntary. Your can choose not to participate or you may withdraw at any time.

Questions
If you have any general questions about the study, please call the principal investigator, Gillian Hillel, at 416-978-088 1, or e-mail her at hillel@mie.utoronto.ca.

Consent
I have had the opportunity to discuss this study and my questions have been answered to my satisfaction. I consent to take part in the study with the understanding I may withdraw at any time. I have received a signed copy of this consent form. I voluntarily consent to participate in this study.

Nurse's Name (Please Print)	Signature	Date

I confirm that I have explained the nature and purpose of the study to the subject named above. I have answered all questions.

Investigator's Name	Signature	Date

APPENDIX 2

University Health Network

Doctor's Order Sheet

ANAESTHESIA/ACUTE PAIN SERVICE
PATIENT CONTROLLED ANALGESIA
(PCA) ORDERS

Addressograph

ALLERGIES
NO KNOWN ALLERGIES ☐

DATE AND TIME			SIGNATU

1. While on PCA device, patient is to receive No further supplemental Narcotics unless approved by the Anaesthesia/Acute Pain Service.

2. **PCA DRUG:**
   Morphine 2.5 mg/mL
   Meperidine _____ mg/mL
   Other: _____ mg/mL

3. **PUMP SETTINGS**
   Dose 1.5 mg to _____ mg
   Initial Lockout Interval 5 min
   Four hour limit 30 mg
   Continuous dose 1 mg/h.

4. **MONITORING:**
   Respiratory Rate, Sedation Score q 2 h x 24hr, then q 4 h. Record on PCA Flow Sheet.

5. **TREATMENT OF SIDE EFFECTS:**
   Have Naloxone (Narcan) 0.4 mg/mL vial readily available at Nursing Station.

   NAUSEA/VOMITING:
   Dimenhydrinate (Gravol) 25-50 mg IV/IM q 3 - 4 h prn.
   IV dose to be infused in 15 - 30 min

   ITCHINESS:
   Diphenhydramine (Benadryl) 25-50 mg IV/IM q 3 - 4 h prn.

a) Respiratory rate less than 10/min
b) Blood Pressure less than 90 mmHg systolic
c) Pulse less than 50/min
d) Sedation Score of 3 (somnolent, difficult to rouse)
e) Unsatisfactory analgesia
f) If four hour limit of drug dose is reached before 4 hours has elapsed.

8. In an emergency, if NO response after calling the above Beeper #, call Anaesthesia Resident through locating.

9. If side effects of slow respiratory rate, hypotension or somnolence occur, STOP PCA

10. RN will check and verify PCA setting once per shift

11. When tolerating fluids well, D/C PCA then start oral analgesics with: _____

PAGER NUMBERS:	GD	WD
Staff Anaesthetist	(416) 667-6095	(416) 667-3557
Nursing Coordinator	(416) 667-6035	(416) 667-3765
Anaesthetist Resident	(416) 664-3490	(416) 589-3618

_____ Physician's Signature

_____ Date

FORM 2126 (Rev.15/1/2001)

University Health Network

Doctor's Order Sheet

ANAESTHESIA/ACUTE PAIN SERVICE
PATIENT CONTROLLED ANALGESIA
(PCA) ORDERS

Addressograph

ALLERGIES
NO KNOWN ALLERGIES ☐

DATE AND TIME			SIGNATU
	1. While on PCA device, patient is to receive No further supplemental Narcotics unless approved by the Anaesthesia/Acute Pain Service.		
	2. PCA DRUG: Morphine 2 mg/mL Meperidine _____ mg/mL Other: _____ mg/mL		
	3. PUMP SETTINGS Dose 1 mg to _____ mg Initial Lockout Interval 6 min Four hour limit 25 mg Continuous dose _____ mg/h.		
	4. MONITORING: Respiratory Rate, Sedation Score q 2 h x 24hr, then q 4 h. Record on PCA Flow Sheet.		
	5. TREATMENT OF SIDE EFFECTS: Have Naloxone (Narcan) 0.4 mg/mL vial readily available at Nursing Station. NAUSEA/VOMITING: Dimenhydrinate (Gravol) 25-50 mg IV/IM q 3 - 4 h prn. IV dose to be infused in 15 - 30 min ITCHINESS: Diphenhydramine (Benadryl) 25-50 mg IV/IM q 3 - 4 h prn.		
	[illegible] :0/1 ml b) Blood Pressure less than 90 mmHg systolic c) Pulse less than 50/min d) Sedation Score of 3 (somnolent, difficult to rouse) e) Unsatisfactory analgesia f) If four hour limit of drug dose is reached before 4 hours has elapsed.		
	8. In an emergency, if NO response after calling the above Beeper #, call Anaesthesia Resident through locating.		
	9. If side effects of slow respiratory rate, hypotension or somnolence occur, STOP PCA [illegible] Acute Pain Service		
	10. RN will check and verify PCA setting once per shift		
	11. When tolerating fluids well, D/C PCA then start oral analgesics with: _____		
	PAGER NUMBERS: GD WD Staff Anaesthetist (416) 667-6095 (416) 667-3557 Nursing Coordinator (416) 667-6035 (416) 667-3765 Anaesthetist Resident (416) 864-3490 (416) 589-3618		
	_____ Physician's Signature _____ Date		

FORM 2126 (Rev. 15/1/2001)

University Health Network

Doctor's Order Sheet

ANAESTHESIA/ACUTE PAIN SERVICE
PATIENT CONTROLLED ANALGESIA
(PCA) ORDERS

Addressograph

ALLERGIES
NO KNOWN ALLERGIES ☐

DATE AND TIME		SIGNATURE

1. While on PCA device, patient is to receive No further supplemental Narcotics unless approved by the Anaesthesia/Acute Pain Service.

2. **PCA DRUG:**
   Morphine 1.5 mg/mL
   Meperidine _____ mg/mL
   Other: _____ mg/mL

3. **PUMP SETTINGS**
   Dose _____ mg to _____ mg
   Initial Lockout Interval _____ min
   Four hour limit 20 mg
   Continuous dose 1.5 mg/h.

4. **MONITORING:**
   Respiratory Rate, Sedation Score q 2 h x 24hr, then q 4 h. Record on PCA Flow Sheet.

5. **TREATMENT OF SIDE EFFECTS:**
   Have Naloxone (Narcan) 0.4 mg/mL vial readily available at Nursing Station.

   NAUSEA/VOMITING:
   Dimenhydrinate (Gravol) 25-50 mg IV/IM q 3 - 4 h prn.
   IV dose to be infused in 15 - 30 min

   ITCHINESS:
   Diphenhydramine (Benadryl) 25-50 mg IV/IM q 3 - 4 h prn.
   IV dose to be infused in 15 - 30 min

   b) Blood Pressure less than 90 mmHg systolic
   c) Pulse less than 50/min
   d) Sedation Score of 3 (somnolent, difficult to rouse)
   e) Unsatisfactory analgesia
   f) If four hour limit of drug dose is reached before 4 hours has elapsed.

8. In an emergency, if NO response after calling the above Beeper #, call Anaesthesia Resident through locating.

9. If side effects of slow respiratory rate, hypotension or somnolence occur, STOP PCA

10. RN will check and verify PCA setting once per shift

11. When tolerating fluids well, D/C PCA then start oral analgesics with: _____

PAGER NUMBERS:	GD	WD
Staff Anaesthetist	(416) 667-6095	(416) 667-3557
Nursing Coordinator	(416) 667-6035	(416) 667-3766
Anaesthetist Resident	(416) 864-3490	(416) 589-3618

_____ Physician's Signature

_____ Date

FORM 2126 (Rev.18/1/2001)

University Health Network

Toronto General Hospital Toronto Western Hospital Princess Margaret Hospital

Doctor's Order Sheet

ANAESTHESIA/ACUTE PAIN SERVICE
EPIDURAL/SPINAL ANALGESIA ORDERS

Addressograph

PLEASE USE BLACK OR BLUE BALLPOINT PEN, PRESS FIRMLY

ALLERGIES:
NO KNOWN ALLERGIES ☐
KNOWN ALLERGIES (Specify) ☐

ORDERED	POSITION	TAKEN

1. While on Epidural Analgesic, the patient is to receive **NO** further supplement Narcotics or other CNS depressants unless approved by the Anaesthesia/Acute Pain Service.

2. **SINGLE DOSE OPIOID INJECTION:**
   ☐ Epidural ☐ Spinal/Intrathecal.
   Operating Room Dose: Drug _____ mg Time: _____

3a. **DRUG FOR CONTINUING EPIDURAL ANALGESIA:**
   Choose ☑ one
   ☐ Bupivacaine 0.1% and Fentanyl 4 mcg/mL in 250 mL normal saline.
   ☐ Bupivacaine 0.1% and Hydromorphone 0.015 mg/mL in 250 mL normal saline.
   ☑ Ropivacaine 0.2% in 200 mL normal saline.
   ☐ Continuous Infusion
   Rate _____ mL/h.
   ☑ **Epidural PCA**
   Rate 5 mL/h.
   Bolus Dose 2 mL.
   Bolus Lockout 20 min.
   4 hour limit 50 mL.

4. **TREATMENT OF SIDE EFFECTS:**
   Dimenhydrinate (Gravol®) 25-50 mg IV/IM q 3 - 4 h prn for nausea/vomiting.
   Diphenhydramine (Benadryl®) 25-50 mg IV/IM q 3 - 4 h prn for pruritus.
   For urinary retention: In and Out catheterization per protocol.
5. Naloxone 1 mL (0.4 mg/mL ampoule to be readily available at Nursing Station.
6. Maintain IV access for 12 hours after Epidural discontinued.

University Health Network

Toronto General Hospital Toronto Western Hospital Princess Margaret Hospital

Doctor's Order Sheet

ANAESTHESIA/ACUTE PAIN SERVICE
EPIDURAL/SPINAL ANALGESIA ORDERS

Addressograph

PLEASE USE BLACK OR BLUE BALLPOINT PEN, PRESS FIRMLY

ALLERGIES:
NO KNOWN ALLERGIES ☐
KNOWN ALLERGIES (Specify) ☐

ORDERED		POSITION	TAKEN	
	1. While on Epidural Analgesic, the patient is to receive **NO** further supplement Narcotics or other CNS depressants unless approved by the Anaesthesia/Acute Pain Service.			
	2. **SINGLE DOSE OPIOID INJECTION:** ☐ Epidural ☐ Spinal/Intrathecal. **Operating Room Dose:** Drug _____ mg Time: _____			
	3a. **DRUG FOR CONTINUING EPIDURAL ANALGESIA:** **Choose ☑ one** ☐ Bupivacaine 0.1% and Fentanyl 4 mcg/mL in 250 mL normal saline. ☐ Bupivacaine 0.1% and Hydromorphone 0.015 mg/mL in 250 mL normal saline. ☑ Ropivacaine 0.2% in 200 mL normal saline. ☑ Continuous Infusion Rate 5 mL/h. ☐ **Epidural PCA** Rate _____ mL/h. Bolus Dose _____ mL. Bolus Lockout _____ min. 4 hour limit _____ mL.			
	4. **TREATMENT OF SIDE EFFECTS:** Dimenhydrinate (Gravol®) 25-50 mg IV/IM q 3 - 4 h prn for nausea/vomiting. Diphenhydramine (Benadryl®) 25-50 mg IV/IM q 3 - 4 h prn for pruritus. **For urinary retention:** In and Out catheterization per protocol.			
	5. Naloxone 1 mL (0.4 mg/mL ampoule to be readily available at Nursing Station.			
	6. Maintain IV access for 12 hours after Epidural discontinued.			

University Health Network

Toronto General Hospital Toronto Western Hospital Princess Margaret Hospital

Doctor's Order Sheet

ANAESTHESIA/ACUTE PAIN SERVICE
EPIDURAL/SPINAL ANALGESIA ORDERS

Addressograph

PLEASE USE BLACK OR BLUE BALLPOINT PEN, PRESS FIRMLY

ALLERGIES:
NO KNOWN ALLERGIES ☐
KNOWN ALLERGIES (Specify) ☐

ORDERED		POSITION	TAKEN
	1. While on Epidural Analgesic, the patient is to receive **NO** further supplement Narcotics or other CNS depressants unless approved by the Anaesthesia/Acute Pain Service.		
	2. **SINGLE DOSE OPIOID INJECTION:** ☐ Epidural ☐ Spinal/Intrathecal. **Operating Room Dose:** Drug _____ mg Time: _____		
	3a. **DRUG FOR CONTINUING EPIDURAL ANALGESIA:** **Choose ☑ one** ☐ Bupivacaine 0.1% and Fentanyl 4 mcg/mL in 250 mL normal saline. ☐ Bupivacaine 0.1% and Hydromorphone 0.015 mg/mL in 250 mL normal saline. ☑ Ropivacaine 0.2% in 200 mL normal saline. ☐ Continuous Infusion Rate _____ mL/h. ☑ Epidural PCA Rate _____ mL/h. Bolus Dose 4.5 mL. Bolus Lockout 25 min. 4 hour limit 40 mL.		
	4. **TREATMENT OF SIDE EFFECTS:** Dimenhydrinate (Gravol®) 25-50 mg IV/IM q 3 - 4 h prn for nausea/vomiting. Diphenhydramine (Benadryl®) 25-50 mg IV/IM q 3 - 4 h prn for pruritus. **For urinary retention:** In and Out catheterization per protocol.		
	5. Naloxone 1 mL (0.4 mg/mL ampoule to be readily available at Nursing Station.		
	6. Maintain IV access for 12 hours after Epidural discontinued.		

APPENDIX 3

PCA User Interface Study Information Sheet

Dear Participant,

Thank you for participating in this study. Please carefully read this explanation of the study and the tasks that you will be asked to perform. Feel free to ask the experimenter any questions that you may have before you begin the experiment.

Goal:

The goal of this study is to determine the usability of two interfaces for a patient-controlled analgesia device. Nurses in a hospital recovery room would typically program this device from settings prescribed on a doctor's order sheet, so that patients in pain can receive anesthetic. This experiment is part of a Masters thesis, which is being conducted under the supervision of Dr. Kim Vicente.

Description:

You will be participating in two sessions, which will be conducted on two different days. Today is your first session. Each session will take 1-2 hours. You will perform six tasks involving one of the interfaces in the first session, and six tasks involving the other interface in the second session. Each task may or may not also involve answering questions related to information in a patient's chart (during which you will be audiotaped). The sessions will be organized as follows:

1. Explanation of study and signing of consent form (1^{st} session only).
2. Filling in of demographics questionnaire (1^{st} session only).
3. Training on the interface.
4. Instruction on how to fill in a mental workload questionnaire (1^{st} session only).
5. 5 min of practice on the interface.
6. 2 min of looking through a patient's chart.
7. Performing the six tasks and filling out the mental workload questionnaire after each.
8. Debriefing & interview to address any questions participant has, participant's comments about the interface and/or the tasks, etc.
9. Payment & signing of receipt (2^{nd} session only).

Special instructions:

It is important that you work as quickly and accurately as possible when programming the interfaces and when answering the questions related to information in a patient's chart. This means that you should work as fast as you can, while making as few mistakes as possible. We have endeavored to make the tasks that you perform in this study realistic of tasks that nurses perform in the hospital recovery room every day. When you are programming the interfaces, imagine that you are doing this for a patient who is currently in a lot of pain and needs pain killer medication as soon as possible. When you are answering the questions, imagine that a busy physician needs information immediately on a patient, and the patient's health is dependant on the speed and accuracy with which you answer the questions.

Compensation:

At the conclusion of your second session, you will be paid $20/hr for both sessions.

APPENDIX 5

Demographics Questionnaire – Nursing Students

Initials:

Degree you are pursuing:

Year of study:

Year of graduation:

Have you ever programmed an infusion pump before?

Have you ever programmed a PCA device (IV or Epidural) before?

Have you ever received any training on how to program a PCA device (IV or Epidural) before?

Have you ever seen a patient's chart before?

Have you ever read a patient's chart before?

Have you ever had to extract information from a patient's chart before?

Demographics Questionnaire – Experienced Nurses

Initials:

Degrees/Diplomas held:

Year of graduation from nursing program:

Length of nursing program (years):

How many years in total have you been practicing as a nurse?

How many years have you been working in the PACU?

How many hours per week do you work in the PACU?

How many shifts do you work per week?

What was the date of your last shift in the PACU?

How many minutes of training have you received on the PCA device?

Have you ever programmed the PCA device?

How many times per day do you program the PCA device?

How many minutes of training have you received on the Epidural PCA device?

Have you ever programmed the Epidural PCA device?

How many times per day do you program the Epidural PCA device?

APPENDIX 6

NASA-TLX Questionnaire Instructions (adapted from Hart and Staveland, 1988)

The purpose of this questionnaire is to examine the "workload" you experienced in the **interface programming task** you just performed. Workload is a difficult concept to define precisely, but a simple one to understand generally. The factors that influence your experience of workload may come from the task itself, your feelings about your own performance, how much effort you put in, or the stress and frustration you felt. The workload contributed by different task elements may change as you get more familiar with a task, perform easier or harder versions of it, or move from one task to another. Physical components of workload are relatively easy to conceptualize and evaluate. However, the mental components of workload may be more difficult to measure.

A set of six rating scales was developed for you to use in evaluating your experiences during the interface programming tasks. Please read the descriptions of the scales carefully. If you have a question about any of the scales, please ask me about it. It is extremely important that they be clear to you. You may keep the descriptions with you for reference while you fill out the questionnaire.

Mental Demand – How much mental and perceptual activity was required (i.e. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

Physical Demand – How much physical activity was required (i.e. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Temporal Demand – How much time pressure did you feel due to the rate pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

Effort – How hard did you have to work (mentally and physically) to accomplish your level of performance?

Performance – How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

Frustration Level – How insecure, discouraged, irritated, stressed and annoyed versus secure, relaxed and complacent did you feel during the task?

APPENDIX 7

Questions Asked in Interrupted Trials

Chart A:

- What surgery did the patient undergo?
- What sedation was the patient on?
- How much heparin is the patient on?

- Why did the patient undergo the surgery?
- Which inotropes did the patient require?
- What is the patient's hemoglobin?

- Where is the patient's IV located?
- What antihypertension medication is the patient taking?
- If the patient's blood sugar is 8.0mmol/L, how much humulin R insulin can you give him?

- How long has the patient been diabetic?
- What is the PTT?
- How much IV fluid did the patient receive?

- What was the patient's blood loss?
- What are the IV fluid replacements on transport to the floor?
- Why wasn't an epidural put in?

- How old is the patient?
- What antibiotic did the patient receive on route to the OR?
- How much fluid on discharge to the unit is the patient's IV set to maintain?

Additional questions:

- Was a CBC drawn?
- Was and ECG done?
- Was a chest x-ray done?
- What is the potassium level?
- What is the temperature?
- How much benadryl was ordered?
- Does the patient have any surgical drains?
- What is the patient's height?
- Was a blood gas drawn?
- What is the patient's weight?
- What is the patient's blood pressure?
- What was the O2 saturation post op?
- What diet is the patient on?
- What is the BMI?
- What previous operations has the patient had?

- Was an arterial line put in?
- When will the patient be discharged?

Chart B:

- What surgery did the patient undergo?
- What medications does the patient take at home?
- How much insulin did the patient receive?

- Why did the patient undergo the surgery?
- Did the patient have an arterial line?
- What is the patient's hemoglobin?

- What respiratory ailment has been in the patient's history?
- What diet is the patient on?
- What is the patient's body mass index?

- When is the patient's follow-up appointment?
- What is the patient's platelet count?
- How much does the patient weigh?

- Does the patient have her gall bladder?
- When will the patient be discharged?
- What is the patient's blood type?

- How old is the patient?
- What previous operations has the patient had?
- How much O2 support is the patient on?

Additional questions:

- What is the patient's height?
- What is the patient's blood pressure?
- What is the patient's temperature?
- Was a CBC drawn?
- Was a blood gas drawn?
- What was their O2 saturation post op?
- Is the patient diabetic?
- How long has the patient been diabetic?

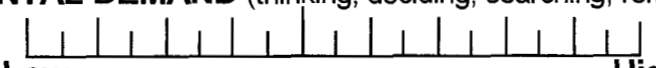
APPENDIX 8

NASA-TLX Questionnaire (adapted from Hart and Staveland, 1988)

Initials:
Trial #:

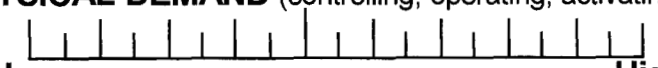
Rate the trial by marking each scale at the point which matches your experience. Each line has two endpoint descriptors to help describe the scale. Please consider your responses to these scales carefully.

MENTAL DEMAND (thinking, deciding, searching, remembering)



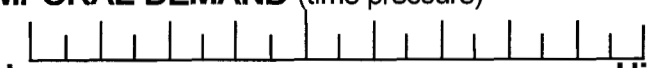
Low (easy, simple) **High** (demanding, complex)

PHYSICAL DEMAND (controlling, operating, activating)



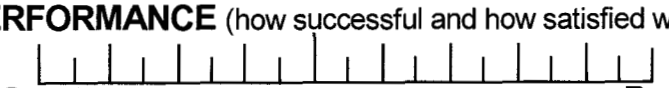
Low (easy, restful) **High** (demanding, laborious)

TEMPORAL DEMAND (time pressure)



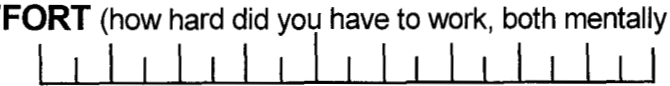
Low (leisurely) **High** (frantic)

PERFORMANCE (how successful and how satisfied were you with performing this task?)



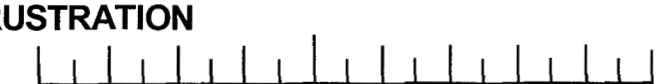
Good **Poor**

EFFORT (how hard did you have to work, both mentally and physically?)



Low **High**

FRUSTRATION



Low (gratified, complacent) **High** (discouraged, annoyed)

Section B: Comments (Use back of page if required)

Instructions:

For each pair of factors, circle the factor that contributed more to the workload for the task you just performed.

1) Physical Demand / Mental Demand

2) Temporal Demand / Mental Demand

3) Performance / Mental Demand

4) Frustration Level / Mental Demand

5) Effort / Mental Demand

6) Temporal Demand / Physical Demand

7) Performance / Physical Demand

8) Frustration Level / Physical Demand

9) Effort / Physical Demand

10) Temporal Demand / Performance

11) Temporal Demand / Frustration Level

12) Temporal Demand / Effort

13) Performance / Frustration Level

14) Performance / Effort

15) Effort / Frustration Level