Tailoring Number Entry Interfaces To The Task of Programming Medical Infusion Pumps

Sarah Wiseman, Duncan P. Brumby, Anna Cox and Orla Hennessy
UCL Interaction Centre, University College London
{s.wiseman, brumby}@cs.ucl.ac.uk, {anna.cox, o.hennessy.11}@ucl.ac.uk

Medical devices are often used to administer medication to patients. This task usually requires a caregiver to enter specific numerical values into a device. In such safety-critical domains, it is vital that this task can be done quickly and accurately. We consider whether tailoring the interface to make it easier for commonly entered numbers to be inputted makes this task faster and less error-prone. To evaluate this idea we take data from infusion pumps programmed on the ward and make adaptations to three existing interfaces to make the task easier (by adding buttons or altering the effects of interaction). The results of a lab-based experiment show that tailoring the interface in this way can significantly reduce the number of key presses that are required to complete the task. We also present findings regarding the process of tailoring interfaces for more general device design.

INTRODUCTION

Many interfaces used in the medical domain require people to enter numbers. For instance, patient record systems require users to enter patient date of birth information. When it comes to the giving of treatment, many medical devices must be programmed to give a specific quantity of medication over time. Infusion pumps in particular require medical workers to enter the volume of medication to be given and the rate at which it should be given over time. This job is made challenging because there are a variety of different makes and models of infusion pump in circulation, meaning that there are many different kinds of number entry interface to be used and learnt. Despite the range of existing interface designs, it appears that none have been designed with the specific task of entering the numbers for an infusion treatment in mind (Wiseman, Cox, & Brumby, 2013). Improving number entry tasks in the medical domain is a pressing challenge, as human error when entering numbers has been attributed as the underlying cause of many unnecessary deaths each year (Kohn, Corrigan, & Donaldson, 2000). Hence medical devices may benefit from the development of tailored interfaces to support more accurate number entry.

If we consider contexts outside the medical domain there are examples of where interfaces have been adapted to better fit the task they are being used for. The aim of this has often been to improve the speed and accuracy of interactions. For instance, since the advent of touch screens it has been possible to adapt typing interfaces so that the screen dynamically changes on the application they are being used in. Take the standard QWERTY keyboard. Here the keys can be adapted for the tasks of entering URLs or email addresses by adding buttons such as '.com' or '@' to the main keyboard. This is motivated by the frequency of these strings and allows users to complete their task of typing a URL using fewer key presses.

In this paper, we consider whether adapting interfaces may have benefits in safety critical domains where incorrect number entry can lead to serious consequences. There are a number of cases of loss of life in the medical domain due to inaccurate number entry (Institute for Safe Medication Practices Canada, 2007; Vicente, Kada-Bekhaled, Hillel, Cassano, & Orser, 2003). Here we investigate how medical interfaces can be adapted to better meet the needs of the task. Such design interventions have already been considered with regard to taking medical notes. The swiftkeyhealthcare.com virtual keyboard uses information about the most commonly input medical words to make the process of note taking a quicker and more accurate task. Rather than looking at how text-typing interfaces can be adapted, we will instead look at how information about common numbers can make number entry faster and more accurate. We explore whether adapting number entry interfaces to match the task can help to reduce the number of key presses required for a task.

Related Work

Current number entry error research highlights the wide variety of causes of error. One study reports a taxonomy of 13 error types and possible causes (Wiseman, Cairns, & Cox, 2011). Slip errors do not just occur at the point when a user is typing the number. Wiseman et al. highlight the fact that errors can be caused by misreading a number, misremembering it, or mistyping it. The error most closely associated with the design of the interface itself is the motor slip error which occurs when a number is read and remembered correctly but is not successfully typed. Research has been conducted into the possible prevention of some errors within the medical domain. Guidelines have been produced by the Institute of Safe Medical Practice to help prevent errors from occurring on the ward. Building on these guidelines, researchers have shown that syntactic rules can be used to help to prevent some types of number entry errors from occurring. For example, preventing numbers entered with more than one decimal point or avoiding trailing zeros (Thimbleby & Cairns, 2010). Others have looked at the effect of interface design on the key bounce error; that is, where an intended single key press is interpreted as a double key press (Oladimeji, Thimbleby, & Cox, 2011). Although previous work shows how errors may be reduced, the focus is often on syntactical errors (such as typing two decimal points in a number) or mechanical errors (caused by the interface incorrectly recording the user's actions). Catching these errors only represents a small sample of possible error types. Many incorrectly typed numbers are otherwise syntactically correct and so would not be caught using this system (e.g., typing 789 instead of 798).

Other research has focused on the effects of different number entry interfaces (Oladimeji et al., 2011). This research compared two interface designs used on infusion pumps, one familiar, and one unfamiliar to users. It was found that on the less familiar interface, users were more likely to check for errors in the numbers they had entered. This suggests that different interface types may affect how likely people are to make number entry errors. Additional research has shown that prompting users to check for errors is not necessarily the most effective way to reduce the occurrence of erroneous numbers being entered into devices (Wiseman, Cox, Brumby, Gould, & O'Carroll, 2013). Instead, this research showed that a more effective way to reduce error was to have people enter an additional 'checksum' number. However, having people enter this additional and redundant checksum number took time and this might impact negatively on how quickly the task can be completed in operational environments.

Another way to reduce the likelihood of error when entering information is to reduce the number of key presses required to enter the same data. One way to achieve this is to shorten some key press sequences by providing short cut keys to common values. For instance, financial calculators often have a dedicated '.00' button for entering round monetary amounts. Similarly, many shop checkout systems will have buttons denoting common currency denominations. A specialized example of number entry adaptation occurs in optician software that is used to enter the prescription of glasses. Such software can provide additional buttons to the interface for the most common prescription endings: .00, .25, .50 and .75.

One of the requirements for developing interfaces that are adapted to the task they are used for is to understand which numbers are commonly entered. In order to make such adaptations to the infusion pump number entry system, information has to be known about which numbers are most often used in infusions. Wiseman, Cox and Brumby (2013) have looked at the numbers used when programming infusion pumps, these numbers were taken from logs of infusion pumps used on five types of wards. They found clear and stable patterns in the numbers that were entered into the pumps across various hospital wards. In particular, it was noticed that over 50% of all infusions will use either the number 1000, 500 or 100, and that very few used the decimal point.

With such strong patterns emerging, it is clear that medical number entry interfaces could be tailored towards the task of medical infusions, by making it easier to enter the most common values. To evaluate this idea we tailor three existing interfaces (by adding keys or altering the way in which button presses affect the number being entered). We evaluate these redesigned interfaces in a study in which participants were asked to enter a series of medical prescriptions, based on the numbers in the existing real world data set of infusion settings. We expect to see that tailoring reduces the number of key presses needed to program an interface, thus reducing the chances for error whilst maintaining interaction speed.

METHOD

Participants. Thirty participants (13 female) from the UCL Psychology subject pool took part in the study. The age of participants ranged from 18 to 43 (M=26.4 years, SD=7.0 years). None of the participants were experienced in using medical devices.

Design. A mixed 3x2 (*interface type* x *tailoring*) design was used, where interface type was the between-subjects factor. The three interface types that were evaluated were: keypad, chevron and dial. For each interface participants used a standard and a tailored version.

The dependent variables in the experiment included the time taken to enter a number and complete each trial, along with timing data for inter-key press timings (where appropriate). Key press logs allowed for error data to be collected (whenever the number entered by the participants did not match the number on the prescription) including the frequency and type of error. User preference was recorded at the end of the trials.

Materials. Participants interacted with a physical prototype of a representative infusion pump (Figure 3). The interface of the prototype could be easily changed between a number pad, a chevron design and a dial. We explain each in turn.

The number pad interface (Figure 1a) has input buttons for the numbers 0 – 9 along with buttons for a clear function and a decimal point. This interface resembles that which is common on calculators and mobile phones with each press of a digit key adding that digit to the display. Moreover, this entry interface can be found on many infusion pumps.

The chevron interface (Figure 2) has four buttons for incrementing and decrementing a number shown on a separate screen. Pressing the upwards-pointing chevrons would increase the number being displayed: by a large amount if the double chevron is pressed, and a small amount if the single chevron is pressed. Interacting in the opposite way would decrease the number. The chevron design is also common on infusion pumps.

For the dial interface, the participant simply turns a knob to increase (turning clockwise) or decrease (turning anti-clockwise) a number on the screen. Pressing the dial in allows for smaller increments and decrements to be made, either smaller integers or numbers after the decimal place. While this interface is used less often on medical devices, it is nonetheless a familiar interface used in many other contexts, e.g., to control the volume of a music system.

As well as exploring different types of input device, we also explore the style of interface: standard vs. tailored. The tailored keypad interface had additional buttons for the most common numbers (e.g., '100', '500' and '1000' buttons, see Figure 1b). The chevron and dial interfaces were adapted by changing the value of the increments and decrements. For the standard interface of the chevron and dial, the large increment and decrement value was set at 10. On the tailored interface, the large increment and decrement value was set to 25 to better match the types of numbers being input (because 61.5% of all numbers have been found to be multiples of 25, Wiseman, Cox, & Brumby, 2013). This means that these two interfaces looked the same in both the standard and tailored condition, but had different effects upon the number on screen.

The prototype was mounted to a stand at a height that allowed participants to operate the device from a seated position. The numbers that the participants were to enter were presented in pairs as a prescription of volume and rate. These were displayed on a nearby laptop computer. During the experiment participants were able to clear the number they had entered and begin again if they wished, for instance if they made an error.

As for the numbers to be entered, we strived to use values that were ecologically relevant to the task that our infusion pump interfaces were designed for. To this end we sampled numbers from a data set collected from the logs of infusion pumps used on various hospital wards. The sample maintained the ratio of "common" and "uncommon" numbers (Wiseman, Cox, & Brumby, 2013).

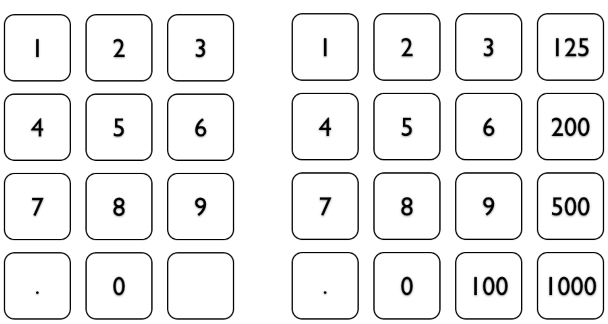

Figure 1 Standard and tailored number pad layouts used in the experiment

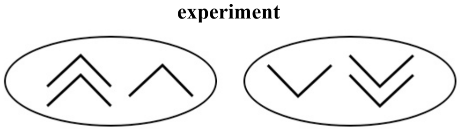

Figure 2 Chevron interface used in the experiment

Figure 3 Prototype infusion pump set up with the tailored number pad interface (as illustrated)

Procedure. Participants were given an initial training period in which they entered 10 numbers to allow them to become familiar with the interface. In the main experimental phase, participants entered 60 pairs of numbers on both the standard interface and the tailored interface (the order of which was counterbalanced across participants). After each pair of numbers (volume and rate) had been entered, a new prescription was displayed on the laptop screen. After participants had entered each of the 60 number pairs they were given a 5 minute break whilst the interface was changed from standard to tailored, or vice versa. Critically, the same number set was used in both the standard and tailored condition, with the order of presentation being randomized within each condition. Participants were told to either enter the numbers quickly or accurately (half in each condition). Participants were also told that if they made an error and keyed the wrong digit that they could correct it by using a clear button before entering it.

RESULTS

We analyse the time to completion, the number of key presses and number of errors for the six interfaces. For statistical analysis, effects were judged significant at a .05 significance level.

Time

We first consider the time taken to enter each pair of numbers (i.e., to complete the prescription task). Results show that participants were significantly faster at entering (pairs of) numbers using the keypad interface (M=5.57s, SD=1.45s) than when using the dial interface (M=7.81s, SD=0.84s) or the chevron interface (M=12.37s, SD=2.12s), $F(2, 27) = 49.41$, $p<.001$. Participants were also significantly faster at entering numbers using the tailored interface (M=8.14s, SD=2.90s) compared to the standard interface (M=9.03s, SD=3.71s), $F(1, 27) = 20.43$, $p<.001$. However, the interface type x tailoring interaction was also significant, $F(2, 27) = 14.56$, $p<.001$. Follow up tests of this significant interaction show that the speedup benefit provided by tailoring the interface was limited to the chevron interface, $F(1, 27) = 47.30$, $p<.001$; there was no significant benefit of tailoring the interface for either the keypad, $p<.17$, or the dial interface, $p<.64$.

Key Presses

To better understand how participants were interacting with each of the interfaces, we consider the total number of key presses made to enter a given number. Results show that participants made significantly fewer key presses when entering a number using the keypad interface (*M*=3.3, *SD*=0.1) than when using the chevron interface (*M*=12.6, *SD*=0.7) or the dial interface (*M*=15.4, *SD*=0.8), $F(2, 27) = 1105.24$, $p<.001$. Participants also made significantly fewer key presses when entering a number using the tailored interface (*M*=8.7, *SD*=4.4) compared to the standard interface (*M*=12.2, *SD*=6.2), $F(1, 27) = 420.37$, $p<.001$. The interface type x tailoring interaction was also significant, $F(2, 27) = 62.41$, $p<.001$. However, follow up tests of this interaction show that the benefit provided by tailoring the interface, in terms of fewer key presses being needed to enter a number, extended to all interface types (all *p*'s < 0.05).

Going beyond these data it is also possible to determine how many key presses or dial rotations would be made on each of the interfaces using "optimum" strategies. This was calculated by finding the minimum number of key presses required to enter each of the numbers in the target number set used in this experiment. Comparing the participants' key presses to these estimates of the optimal strategy allows us to determine how well participants had learned to use the interface in its standard and tailored form. Table 1 shows the optimal performance compared to the actual performance on each of the interfaces. It can be seen that participants were performing very close to optimal on the keypad, but were making many more key presses than were necessary for the chevron and dial interface. It can also be seen that the tailored interface was able to make reductions in the number of key presses required for all interfaces, even on the keypad where participants had to use one fewer key press per number.

Table 1. Mean number of key presses needed to enter the numbers using an "optimal" strategy compared the mean number of key presses used by participants in the experiment

	Standard Interface		Tailored Interface	
	Actual M (SD)	*Optimal*	*Actual M (SD)*	*Optimal*
Keypad	3.7 (0.1)	3.7	2.9 (0.1)	2.9
Chevron	15.0 (0.7)	14.3	10.2 (0.7)	9.5
Dial	17.8 (1.1)	14.3	13.1 (1.1)	9.5

Errors

We next consider number entry errors. Errors were defined as any number that was entered which did not match the target number on the prescription. In general participants made very few errors – the mean error-rate across all conditions was 0.99% (*SD*=1.07%). Error-rates were marginally higher when participants used the keypad interface (*M*=1.42%, *SD*=1.52%) compared to the chevron (*M*=0.87%, *SD*=0.67%) or the dial interface (*M*=0.67%, *SD*=0.77%). However, this difference was not statistically significant, $p=0.89$. There was also no difference in error-rate between the standard (*M*= 1.00%, *SD*=1.36%) and tailored interface (*M*= 0.97%, *SD*=1.05%), $p=0.28$. The interaction was also non-significant, $p=.21$.

An analysis of error *types* was conducted using a number entry error taxonomy (Wiseman, Cairns, & Cox, 2011). As can be seen in Figure 4 there are different error profiles for the standard interfaces compared to the tailored interfaces. In particular, the tailored interfaces were less susceptible to missing digit errors. If we look at just those errors which resulted in a number being entered that was out by at least a factor of 10 (arguably the most severe errors in a hospital ward), it can be seen that missing digit and added digit errors make up 75% of all errors on the standard interfaces but only 42.86% of all errors on the tailored interfaces.

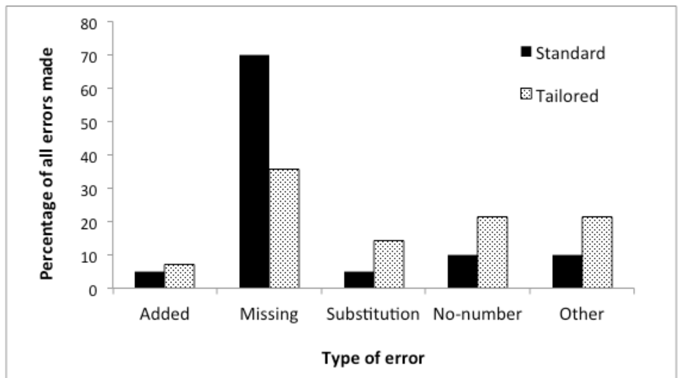

Figure 4 Types of error made by participants in the experiment on both interfaces. Note the drop in digit missing errors on the tailored interfaces. 'No-number' errors occurred when a participant skipped entering a number on that trial

User Preference

Finally we asked participants for their reaction to the tailoring of interfaces – the feedback was varied. Tailoring was preferred for the keypad (8 out of 10 participants preferred the tailored version), but was widely rejected for the dial interface (9 out of 10 participants preferred the standard version). For the chevron interface, participants did not have a clear preference (6 out of 10 participants preferred the tailored interface).

We also asked participants how familiar they were with the interface that they had used. Nearly all participants in the keypad interface condition and the dial interface condition had experience using similar interfaces. However, participants in the chevron condition had the highest number of participants (30%) with no previous experience of that interface.

DISCUSSION

This work was conducted in order to find out whether tailoring interfaces, by making them better suited the to the tasks completed on them, could make interaction with them faster and less error prone. The results of the conducted experiment show that interfaces do benefit from tailoring: with reductions in the number of key presses needed to program a prescription and, with some interface types, a reduction in programming time. There was however no benefit of tailoring the interface for reducing the likelihood of error: error rates in this study were consistently low regardless of interface type or tailoring. This shows that tailoring an interface can, in some instances, offer an increase in speed, without the usual associated cost of increased error due to lower accuracy.

The reason that a time improvement was seen on only on the chevron interface, and not on the dial or keypad interface may be due to familiarity of the interface itself. Most participants in the experiment had used a keypad and dial interface prior to the experiment, whereas most of those who used the chevron interface had not. This meant that an equal amount of learning was required for both the standard and tailored chevron interface. This was not the case for the other two: only the tailored interface of the keypad and dial was new to participants, requiring them to learn how to interact with them. This explanation may account for why there was a reduction in key presses but no associated reduction in programming time. In future work, the experiment should be run for longer, to allow participants to fully learn the new interfaces.

The familiarity with the interfaces can also give us more guidance for when tailoring interfaces more generally. The tailoring was generally seen as a benefit to the keypad and chevron by the participants, but not to the dial interface. The tailoring occurred in two different ways in the experiment, either by adding visually obvious extra buttons (keypad) or by changing the way the interface affected the number on screen (chevron and dial). It seems that participants liked the tailoring when it was visually salient, but had mixed opinions when it was hidden. The preference for tailoring on the unfamiliar chevron interface suggests that because the participants had no preconceived notions about how that interface worked, they were willing to see the benefits of the tailored version, as both were just as new to the participants. However, when the dial interface was tailored, it went against what the participants expected from this familiar *looking* (but not *acting*) interface. Thus we have a guideline to consider when tailoring interfaces – if the interface is familiar to participants, make the tailoring visually obvious to avoid negative feelings towards the interface from participants; hiding adaptations on familiar interfaces can leave users confused, regardless of the beneficial effects of reducing the number of key presses.

One limitation of this study was the participants' inexperience in the medical domain. Although the task used an infusion pump-like device and realistic data, participants were not medical professionals, and therefore did not have the same level of prior experience. It may also be useful to investigate the effects of manipulating motivation in this study in the future. There were no obvious penalties for entering numbers incorrectly in this experiment, which is obviously at odds with the real world application of this to the medical domain. Making errors more costly to the participants could possibly alter the number and type of errors made. However, this study has been a preliminary step, showing that tailoring can have some benefits and may be worth exploring further in the medical domain.

CONCLUSION

Interfaces can be effectively tailored to meet the needs of the task with benefits including reduced interaction time and fewer key presses without effecting error rates. Tailoring interfaces does not therefore offer a solution to reducing user error but may help with making the task easier. User preference for interface adaptation is also strongly affected by the familiarity of the interface. When tailoring a familiar interface, users prefer adaptations to be visually salient. In the future, tailoring may make it possible to create an interface that users prefer, is faster, and reduces key presses and thus the chance for error.

ACKNOWLEDGEMENTS

This work was supported by the UK Engineering and Physical Sciences Research Council [EP/G059063/1].

REFERENCES

Institute for Safe Medication Practices Canada. (2007). *Fluorouracil Incident Root Cause Analysis*.

Kohn, L., Corrigan, J., & Donaldson, M. (2000). *To err is human: building a safer health system*.

Oladimeji, P., Thimbleby, H., & Cox, A. (2011). Number entry interfaces and their effects on error detection. *Human-Computer Interaction–INTERACT 2011*, 178–185.

Thimbleby, H., & Cairns, P. (2010). Reducing number entry errors: solving a widespread, serious problem. *Journal of the Royal Society, Interface*, *7*(51), 1429–39. doi:10.1098/rsif.2010.0112

Vicente, K. J., Kada-Bekhaled, K., Hillel, G., Cassano, A., & Orser, B. A. (2003). Programming errors contribute to death from patient-controlled analgesia: case report and estimate of probability. *Canadian journal of anaesthesia*, *50*(4), 328–32. doi:10.1007/BF03021027

Wiseman, S., Cairns, P., & Cox, A. L. (2011). A taxonomy of number entry error. *BCS Conference on HCI* (pp. 187–196).

Wiseman, S., Cox, A. L., & Brumby, D. P. (2013). Designing Devices With the Task in Mind: Which Numbers Are Really Used in Hospitals? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *55*(1), 61–74. doi:10.1177/0018720812471988

Wiseman, S., Cox, A. L., Brumby, D. P., Gould, S. J. J., & O'Carroll, S. (2013). Using Checksums to Detect Number Entry Error. *CHI 2013*.