Visualizing Magnitude: Graphical Number Representations Help Users Detect Large Number Entry Errors

Judith Borghouts[1,2], Frank Soboczenski[2], Paul Cairns[2], Duncan P. Brumby[1]
[1]UCL Interaction Centre, University College London, London, UK
{judith.borghouts.14, d.brumby}@ucl.ac.uk
[2]Department of Computer Science, University of York, York, UK
{frank.soboczenski, paul.cairns}@york.ac.uk

Nurses frequently have to program infusion pumps to deliver a prescribed quantity of drug over time. Occasional errors are made in the performance of this routine number entry task, resulting in patients receiving the incorrect dose of a drug. While many of these number entry errors are inconsequential, others are not; infusing 100 ml of a drug instead of 10 ml can be fatal. This paper investigates whether a supplementary graphical number representation, depicting the magnitude of a number, can help people detect number entry errors. An experiment was conducted in which 48 participants had to enter numbers from a 'prescription sheet' to a computer interface using a keyboard. The graphical representation was supplementary and was shown both on the 'prescription sheet' and the device interface. Results show that while overall more errors were made when the graphical representation was visible, the graphical representation helped participants to detect larger number entry errors (i.e., those that were out by at least an order of magnitude). This work suggests that a graphical number entry system that visualizes magnitude of number can help people detect serious number entry errors.

INTRODUCTION

Many tasks in healthcare require people to enter numbers shown on one artifact to another (Wiseman, Cox, & Brumby, 2013). For instance, programming an infusion pump involves entering the volume of medication to be given to a patient over a specific period of time. Usually this task can be completed without any problems, but slips do occur, and when they do they can have serious consequences (Westbrook, Woods, Rob, Dunsmuir, & Day, 2010) and contribute towards adverse healthcare events (Cauchi, Gimblett, Thimbleby, Curzon, & Masci, 2012). To give an estimation of the scale of this problem it has been estimated that of the 15 million infusions carried out by members of the UK health service each year, about 700 of these result in adverse events (Vincent, 2010). One potential solution to this problem is to encourage healthcare workers to more carefully check for errors when programming infusion pumps. Alas, this approach is unlikely to mitigate the problem; previous research has consistently shown that people are very poor at detecting number entry errors when they occur (Olsen, 2008; Wiseman, Cox, Brumby, Gould, & O'Carroll, 2013).

An alternative approach to help people spot number entry errors is to improve the design of the device interface that is used. It is well known that graphical representations can be used to convey information better than textual representations (Larkin & Simon, 1987). For infusion pump programming tasks, previous research has shown that visualizations can help understand the numbers that they have entered into the device and also mitigate number entry errors (Gould, Cox, & Brumby, 2013; Thimbleby & Williams, 2013; Tu, Oladimeji, Wiseman, Thimbleby, Niezen, & Cairns, 2014). However, as we shall discuss in more detail below, these studies have tended not to look at the severity of number entry errors made.

Not all number entry errors are alike. For number entry in medicine, the magnitude of the error matters greatly. Infusing 100 ml of a drug instead of 10 ml could be fatal, whereas infusing 10.01 ml is most likely inconsequential. A useful way of categorizing number entry errors in healthcare is to focus on cases in which the entered number is out by at least an order of magnitude from the intended number: an out-by-ten error (Doherty & McDonnell, 2012; Thimbleby & Cairns, 2010). Doherty and McDonnell (2012) studied out-by-ten errors over a period of 5 years at a pediatric hospital and found 252 out-by-ten errors, 196 of which were judged to have resulted in a severe outcome for the patient involved.

Given that people are prone to making number entry errors (Doherty & McDonnell, 2012; Vincent, 2010), and that people are often poor at spotting these errors when they occur (Olsen, 2008; Wiseman, Cox, Brumby, Gould, & O'Carroll, 2013), we consider whether a graphical number representation that visualizes magnitude might make it easier for users to detect out-by-ten errors. In the following section we review recent work that has explored using graphical number representations to help people detect number entry errors. Following this, we describe the results of an experiment that evaluated a novel graphical number representation scheme that was designed to help people detect serious number entry errors.

Related work

The influence of information representations on task performance has been studied widely in the field of Human-Computer Interaction. Here we focus our review on a few recent studies that have evaluated different forms of graphical number representations designed to support number entry in the context of medical infusion pump programming (Gould et al., 2013; Thimbleby & Williams, 2013; Tu et al., 2014).

Thimbleby and Williams (2013) studied nomograms to facilitate calculations on infusion pump parameters, such as medication volume and rate of infusion. As can be seen in Figure 1a, a nomogram consists of a number of scales, with one scale per infusion pump parameter. If the values of two parameters are known, a line can be drawn through these two points on the scales which then intersects with the third scale at a point corresponding to the correct value for that parameter.

Gould et al. (2013) developed a graphical representation to help people visualize the rate of infusion in an infusion pump programming task. In the graphical representation, shown in Figure 1b, the rate of infusion can be inferred from the gradient of the line (connecting duration of infusion and volume to be infused). Gould et al. asked participants to do simple calculations using this graphical representation and compared this to a simple textual description of the numerical values involved (i.e., the rate, duration, and volume to be infused). Results showed that people were faster and equally accurate in solving problems with the graphical representation.

The results of Thimbleby and Williams (2013) and Gould et al. (2013) show that graphical representations can help people make sense of numbers. However, these studies focused on calculation tasks in which the graphical representation was used as an external support aid to perform a mental calculation. While calculation tasks are important, we are primarily interested here in number entry tasks.

A recent study by Tu et al. (2014) evaluated the benefits of a graphical number representation to help people detect number entry errors. The graphical number representation developed by Tu et al. (2014) is shown in Figure 1c. The representation consists of a circle, where digits (0-9) and the decimal point each have a specific position on the circle, and an arrowed line connecting the positions shows the order of characters for a number. In Figure 1c, the positions of 6, the decimal point and 8 are connected, displaying the number 6.8. Tu et al. found that participants made fewer errors when using this graphical representation, but at the expense of slower task completion times, suggesting a speed-accuracy trade-off. A further limitation of Tu et al.'s representation is that it does not have a direct mapping with the semantics of the number, so the meaning of the representation has to be learned by participants. Given this concern, we were motivated to develop a visualization technique for representing the magnitude of a number that is easy to understand and use.

Visualizing magnitude of number

In developing our representation for visualizing the magnitude of a number, we took inspiration from Norman's (1994) Naturalness Principle. Norman argues that information representations should be designed to match the properties of what they are intending to represent. For instance, the physical properties of representation can be used to represent the size of the number (Larkin & Simon, 1987). An advantage of such a proportional representation is that it changes an abstract symbol like the Arabic notation into a perceptual one that is much easier to interpret and compare. This can make it far easier for the user to perceive large differences in magnitude, which would be harder to detect in a textual representation.

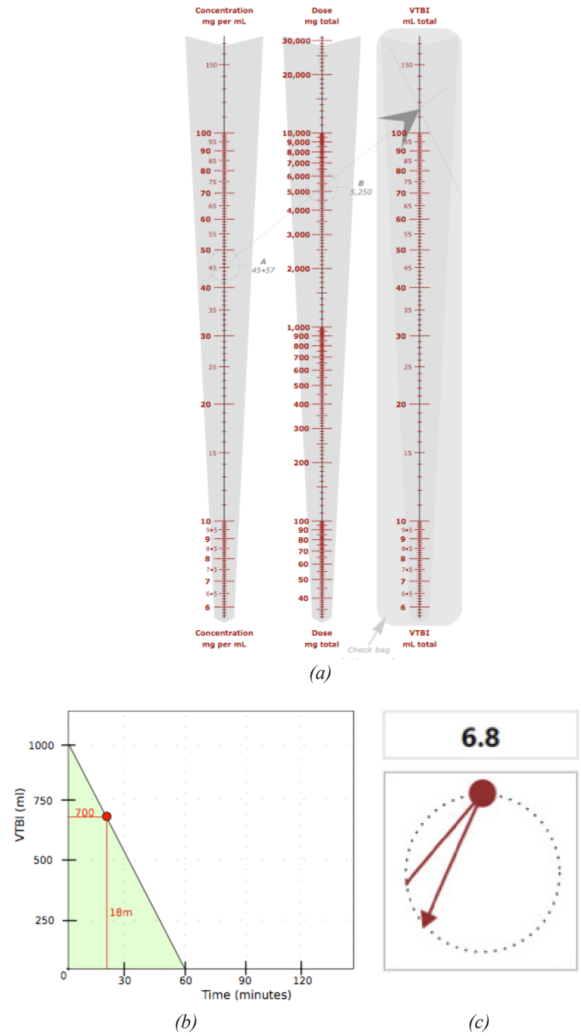

Figure 1: Graphical number representations from previous studies. (a) The nomograms used by Thimbleby and Williams (2013) with one scale for each infusion pump parameter. (b) Gould et al.'s (2013) chart with two parameters plotted on the axes. (c) The representation used by Tu et al. (2014), where each digit has a specific position on the circle, and the red arrow connects the positions. See text for an explanation of the representations.

We developed a novel representation for visualizing the magnitude of a number. This representation was designed to help people detect severe out-by-ten errors. As can be seen in Figure 2, our graphical representation consists of a block whose size is proportional to the number. The width of the block represents the order of magnitude, and the height of the block represents the overall value of the number. So, for instance, the blocks of values 90 and 900 are equally high, but differ in width. Similarly, 100 and 900 have an equal block width, but differ in height (see Figure 2). On the display, the block size of the entered number was updated with each keystroke that the participant made. The digits were positioned on top of the block, so the participant did not need to make many eye movements to perceive both representations. Another reason why a block with two dimensions was used rather than a bar or line, which only has one dimension, was to make more efficient use of space, and to be able to show a wider range of numbers. If one dimension had been used, only a limited set of possible values could have been shown on the screen or the representation would have been scaled in a way that would have made it hard to see differences in length.

Overview of study

We conducted a study in which participants used our representation for visualizing the magnitude of a number when performing a number entry task. Participants were asked to enter a collection of numbers that were either presented in a graphical representation (as shown in Figure 2) or in a standard textual format. We consider whether our representation system helped participants detect number entry errors. We assume that the addition of the visualized number block changes the abstract task of comparing digits and interpreting their magnitude to a perceptual one of comparing block sizes. Small differences in size might be hard to detect but larger differences will be easier to perceive. The expectation is that this kind of representation will make it easier to detect large severe errors, such as out-by-ten errors.

A secondary aim of our study was to investigate how the cost of accessing the to-be-entered number affects the errors people make, and if this is different for textual and graphical representations. Previous studies showed that if it takes more effort to access the information needed for a task, people try to encode the information better in memory, and this deeper encoding leads to fewer errors being made on that task (Back, Cox, & Brumby, 2012; Morgan, Patrick, Waldron, King, & Patrick, 2009; Soboczenski, Cairns, & Cox, 2013). Back et al. (2012) manipulated Information Access Cost (IAC) in programming infusion pumps, and found that when a prescription form showing the to-be-entered numbers was placed further away from the input device, participants memorized the numbers in chunks and as a result made fewer errors. Soboczenski et al. (2013) encouraged deeper encoding by presenting numbers in a transcription task in a less legible font color, and found this also resulted in fewer errors. The current study varied IAC by either placing the numbers next to the computer or further away. We speculate that it might be easier for participants to adopt a more memory-intensive strategy with a graphical number representation because this representation would allow a number's magnitude to be directly perceived.

METHOD

Participants

Forty-eight participants (26 male) completed the experiment. They were all students at a university and came from a variety of disciplines and nationalities. Ages ranged from 19 to 39 with a mean age of 24.69 (SD = 3.8).

Design

A 2x2 (number representation type x IAC) between-participants design was used. Numbers were represented in only a textual manner, or in a textual and graphical manner. In the low IAC condition, the numbers were placed next to the laptop where the numbers had to be entered. In the high IAC condition, the numbers were situated approximately 70 cm away from the laptop and turned 180 degrees from the position of the front of the screen.

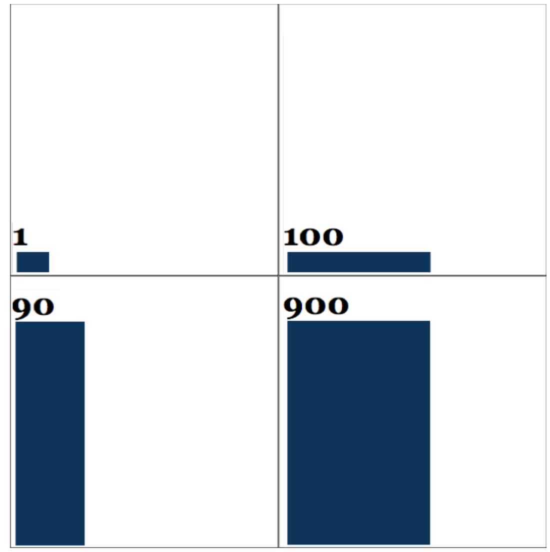

Figure 2: Four examples to illustrate how numbers with differing magnitudes would be displayed with the graphical representation used in this study.

The dependent variables were task completion time and number of errors made. In addition, we also considered the proportion of errors that were out-by-ten errors. In this study, out-by-ten errors comprise all errors that are out by at least an order of magnitude, so out-by-100 or out-by-1000 errors are also included.

Materials

The experiment was conducted on a 13-inch MacBook Pro laptop, and participants had to enter numbers using an external number pad. The app NumPad Remote was used to simulate a touch screen number pad with a calculator layout and was run on an iPod Touch. The numbers to enter were shown on paper cards, and had the same representation as how they were presented on the computer screen. Participants had to enter 50 numbers. These numbers were of varying length, being made up of either two, three, four, five or six digits each (e.g., 10.43 was a four digit number). Most of the numbers were decimal numbers. All participants had to enter the same set of numbers, in the same order, and there were 10 numbers of each length. The entered numbers and task time were automatically recorded.

Procedure

Participants were welcomed and informed about the number entry task. In the graphical condition, the block was explained and its meaning was illustrated with some examples. Participants were given two numbers to enter as practice to get familiar with the set-up. Once they were ready, the time would start and participants had to enter 50 numbers as fast as they could. This instruction was added to elicit errors and ensure there was a sufficient amount of errors to study. A stopwatch was situated next to the number pad to increase time pressure, but this was not used to record the time. Overall, the experiment took approximately 5 minutes with a low IAC, and 10 minutes with a high IAC.

RESULTS

For the analysis of errors, we first consider the total number of trials in which a transcription error was made (over the 50 trials completed). For the effect of number representation, participants made significantly fewer errors when using the textual-only representation (M = 2.1, SD = 2.2) than the textual+graphical representation (M = 3.8, SD = 3.6), $F(1, 44) = 4.33$, p =.04, η^2 =.09. Participants also made significantly fewer errors in the low IAC condition (M = 1.9, SD = 2.5) than in the high IAC condition (M = 4.0, SD = 3.3), $F(1, 44) = 6.19$, p =.02, η^2 = .12. The interaction was not significant, $F < 1$.

While the above analysis shows that participants made more errors when using the textual+graphical representation, this analysis does not give any indication of the magnitude of number entry errors made. One of the assumed benefits of the graphical representation is that it should allow participants to easily detect larger errors. We therefore conducted an additional analysis where we considered the proportion of number entry errors that are out by at least an order of magnitude (out-by-ten errors). We found that participants were less likely to make out-by-ten errors when using the textual+graphical representation (8 of 91 errors) than when using the textual-only representation (14 of 50 errors), $\chi^2(1)$ = 9.04, p = .003. There was no such difference in the likelihood that participants made out-by-ten errors dependent on IAC condition, $\chi^2 < 1$.

Finally, we consider the average time participants took to enter a number (i.e., complete a trial). There was no significant effect of representation on number entry time, F<1. But as expected, participants were significantly faster at entering a number when in the low IAC condition (M = 4.58, SD = 0.84 s) than when in the high IAC condition (M = 9.97, SD = 2.16 s), $F(1, 44) = 125.00$, $p < .001$, $\eta^2 = 0.74$. There was also no significant interaction, F<1.

GENERAL DISCUSSION

This paper investigated the influence of a graphical number representation on errors in a number entry task. We took transcribing numbers in an infusion pump as the specific scenario. Results show that people made more errors in the graphical condition. However, a closer look at the type of errors reveals an interesting finding. Even though, overall, fewer errors were made in the textual condition, almost twice as many out-by-ten errors were made.

This suggests that participants made use of the visualized number block in the graphical representation, which gives people a sense of the number's magnitude, but the precise amount cannot be easily read from this representation alone. This would explain why more errors were made overall, but why these errors were smaller in severity than the ones made in the textual condition.

The finding that a graphical representation can improve some aspect of number entry is consistent with previous work (Gould et al., 2013; Thimbleby & Williams, 2013). However, these studies focused on a calculation task and the representations were designed to show the relation between multiple numbers. A novel finding of the current study is that a graphical representation can also be useful when making sense of the magnitude of one number.

Tu et al. (2014) used a graphical number representation to support visual checking in number entry, however their representation did not have a direct mapping with the semantics of the number and people took a longer time to use and perhaps understand this type of representation. In the current study, there was no difference in task completion time between representations, which means even though people had to process more information with both digits and a block, this did not slow participants down, and the difference in errors was not due to a speed-accuracy trade-off.

The results suggest that participants were using the representation to check their input when entering numbers. This finding is in contrast to earlier work that found that people do not look to check their number entries, even when they are explicitly asked to do so (Olsen, 2008; Wiseman, Cox, Brumby, Gould, & O'Carroll, 2013). However, in these previous studies only textual representations were used. It could be that a graphical representation changes the way people check their entries.

Our results suggest that a graphical representation helps people in perceiving or checking the magnitude of a number. This can be of great importance in situations where the magnitude matters, such as when entering a medication dose.

Although the finding that out-by-ten errors were reduced is promising, it is not ideal that more errors were made in total. Future work should extend the graphical design and look if the design can be refined in such a way that smaller differences can be noticed as well. An example is to include an x- and y-axis to the block from which the exact value can be read. The exact value could already be read in this study since the digits were placed on top of the block, but perhaps this was not seen as part of the block and people primarily used the block itself.

One possible application of the current graphical representation could be that people are first asked to enter a number textually, and after that are presented with a graphical representation as an extra check. This would prevent people solely using the graphical representation and would encourage them to use the more precise textual representation first. The second representation would then hopefully help them detect large errors that are hard to see in a textual representation, but would become clear with a graphical representation. However, this suggestion assumes that participants primarily used the graphical representation in checking and there is no data available to support this claim. We argue that our findings suggest the different representations influence errors, but it is unclear how exactly people use the representation. Future work is needed to see if and when people look at the representations to check their entries.

A limitation of this study was that participants were not nurses, and that the number entry task was not performed on an infusion pump prototype. Nurses' prior experience using infusion pumps with a textual representation may influence their initial performance on different types of representations. Nevertheless, this study has clearly shown the effect of a graphical representation on reducing out-by-ten errors, and may be worthwhile examining in the medical domain where an out-by-ten error is a common and dangerous error (Doherty & McDonnell, 2012; Thimbleby & Cairns, 2010).

The motivation of this study was to reduce out-by-ten errors in hospitals, but its findings could also be useful for other number entry applications where a magnitude has to be entered, such as financial amounts.

Finally, a secondary aim of the study was to investigate the influence of IAC on errors. Previous research had shown a positive effect of a high IAC on reducing errors in number entry (Soboczenski et al., 2013) and programming infusion pumps (Back et al., 2012). In this study, no positive effect was found which might have been due to the set-up of the experiment. Participants were instructed to complete the task as fast as possible, which may have interfered with people's effort to memorize the information and made the high IAC condition harder and erroneous. Further work is needed to investigate the influence of IAC, and to what extent manipulating this variable is applicable in a medical context.

CONCLUSION

This work is an important step in understanding how the type of number representation can influence the type of errors during a number entry task. The study contributes to the design of number entry interfaces, by showing that a graphical number representation can considerably reduce out-by-ten errors. A potential direction for future research would be a refinement of the graphical representation design, to prevent that a reduction in severe errors will not lead to an increase in overall errors. Furthermore, in order to determine the applicability of a graphical representation in a medical setting, future studies should involve medical staff.

REFERENCES

Back, J., Cox, A., & Brumby, D.P. (2012). Choosing to interleave: human error and information access cost. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems, CHI '12* (pp. 1651–1654). ACM Press.

Cauchi, A., Gimblett, A., Thimbleby, H., Curzon, P., & Masci, P. (2012). Safer "5-key" number entry user interfaces using differential formal analysis. In *Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers, BCS-HCI '12* (pp. 29–38). British Computer Society.

Doherty, C. & Mc Donnell, C. (2012). Tenfold medication errors: 5 years' experience at a university-affiliated pediatric hospital. *Pediatrics*, *129*(5), 916–924.

Gould, S., Cox, A., & Brumby, D.P. (2013). Using graphical representations to support the calculation of infusion parameters. In *Human-Computer Interaction – INTERACT 2013, Lecture Notes in Computer Science, 8120,* 721-728. Springer.

Larkin, J. & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 99*, 65–99.

Morgan, P.L., Patrick, J., Waldron, S.M., King, S.L., & Patrick, T. (2009). Improving memory after interruption: exploiting soft constraints and manipulating information access cost. *Journal of Experimental Psychology: Applied*, *15*(4), 291–306.

Norman, D.A. (1994). *Things That Make Us Smart.* Basic Books.

Olsen, K.A. (2008). The $100,000 keying error. *Computer*, *41*(4), 108–107.

Soboczenski, F., Cairns, P., & Cox, A.L. (2013). Increasing accuracy by decreasing presentation quality in transcription tasks. In *Human-Computer Interaction – INTERACT 2013, Lecture Notes in Computer Science, 8118,* 380-394. Springer.

Thimbleby, H. & Cairns, P. (2010). Reducing number entry errors: solving a widespread, serious problem*. Journal of the Royal Society, Interface*, *7*(51), 1429–1439.

Thimbleby, H. & Williams, D. (2013). Using nomograms to reduce harm from clinical calculations. In *2013 IEEE International Conference on Healthcare Informatics* (pp. 461–470). IEEE.

Tu, H., Oladimeji, P., Wiseman, S., Thimbleby, H., Niezen, G., & Cairns, P. (2014). Employing number-based graphical representations to enhance the effects of visual check on entry error detection. *International Symposium on Interaction Design and Human Factors (IDHF 2014),* Kochi, Japan.

Vincent, C. (2010). What interactive medical device manufacturers need. *BCS Interfaces,* 84, 14–15.

Westbrook, J. I., Woods, A., Rob, M. I., Dunsmuir, W. T. M., & Day, R. O. (2010). Association of interruptions with an increased risk and severity of medication administration errors. *Archives of Internal Medicine, 170*, 683–690.

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, 55,* 61-74.

Wiseman, S., Cox, A.L., Brumby, D.P., Gould, S.J.J., & O'Carroll, S. (2013). Using checksums to detect number entry error. In *Proceedings of the 2013 Conference on Human Factors in Computing Systems, CHI '13* (pp. 2403–2406). ACM Press.