

Exploring unlikely errors using video games: An example in number entry research

Patrick Oladimeji¹, Harold Thimbleby¹, Paul Curzon², Ioanna Iacovides³, Anna Cox³

¹Future Interaction Technology Lab
Swansea University
Swansea, SA2 8PP, UK

²Queen Mary University
London, E1 4NS, UK

³UCL Interaction Centre
University College London
London, WC1E 6BT, UK

p.oladimeji@swansea.ac.uk
h.thimbleby@swansea.ac.uk

paul.curzon@eecs.qmul.ac.uk

i.iacovides@ucl.ac.uk
anna.cox@ucl.ac.uk

ABSTRACT

A common and important feature of many safety critical interactive devices is number entry. In hospitals, number entry takes the form of setting drug parameters such as doses, volumes, etc. There are several ways a number entry interface can be designed - with different consequences for error and speed. Nurses and healthcare practitioners usually have to interact with different interfaces often under pressure and stress of taking care of patients with different health conditions. Error rates in practice are low, undetected error rates are even lower and obtaining the context in which the errors occur is often incredibly difficult due to poor logging systems in many medical devices and high cost of planning and conducting empirical studies. Laboratory based studies also suffer similar limitations in that, without interventions, error rates are also too low to study. This paper explores the benefits of using a gaming context to study safety critical systems. We argue that a game paradigm provides a way that overcomes many of the problems of studying low error rates in safety critical systems and specifically for number entry in medical contexts.

Author Keywords

Safety critical system design, human error, games, empirical studies.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Safety critical interaction design prioritises safety requirements in the design process. Human error in the use of safety critical systems is an important factor to take into account with respect to requirements. Good interface and interaction

design is critical with respect to the extent to which human error occurs. Meeting these safety requirements requires a good understanding of the scope of potential human error in the use of systems being designed and subsequently improving the design to encourage error detection and better error management.

The design of medical devices has a direct impact on patient safety. According to the NPSA (National Patient Safety Agency) [17], each hospital in England and Wales administers about 7000 medicine doses each day. Intravenous medications are essential for hospitalised patients and the treatment of specific diseases typically requires a mixture of multiple intravenous medications to be administered simultaneously [12]. As a result, a significant portion of drug deliveries are administered intravenously and due to the greater complexity involved in preparing, administering and monitoring intravenous drug deliveries, they pose a higher risk to patient safety in comparison to oral delivery [24, 25, 21, 11].

We are specifically concerned with medical device applications such as those used in infusion therapy. Setting up an infusion pump correctly involves entering numbers corresponding to the drug dose, rate, volume to be infused (VTBI) or time. Number entry interfaces can be designed in a number of different ways and two interfaces with exactly the same user interface elements can be implemented to work very differently such executing the same sequence of user actions on both interfaces produces completely different results [5]. It is therefore extremely important that HCI studies are conducted to ensure the consequences of different interface designs and features with respect to human error is understood. However, error rates in actual use of systems are generally extremely low. For example, Vicente et al. [23] estimated the rate of errors related to number entry tasks, such as those arising from programming infusion devices to be in the range of 1 in 33,000 to 1 in 338,800 for an example device. This makes it difficult to conduct traditional lab based experiments to study human error in a way that gives significant results within a realistic time frame. Given the potential severity of the consequences when they do occur, even extremely low error rates in safety critical systems should be addressed where possible.

Furthermore studying such errors in context as opposed to

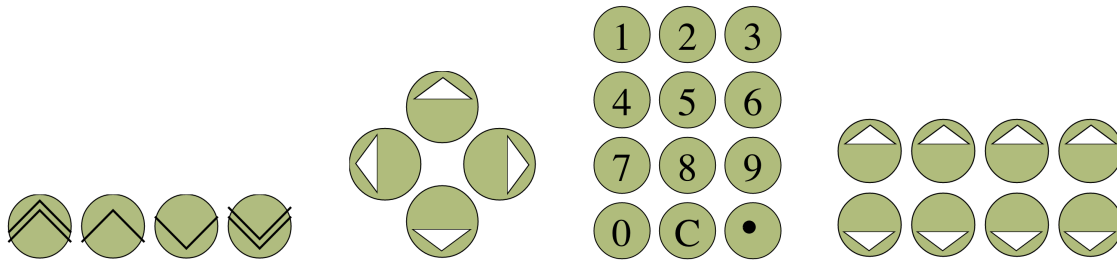


Figure 1. The different types of number entry interfaces encountered in the game.

in the lab can be both difficult and expensive. Errors that are reported are often the result of active failures that emerge at the sharp end of Reasons swiss cheese model [20] rather than the underlying causes of those failures. Consequently, it is more difficult to ask questions or reason about the latent errors that occur in a safety critical system.

These issues are particularly pertinent in our work studying the extent to which different kinds of number entry interfaces support the prevention of human error. The adoption of a gaming paradigm from a citizen science context may provide a way to overcome these problems. Video games are particularly valued for their richness and complexity as well as the possibility to customise game play to fulfil a scientific research requirement [4]. A gaming context can offer a more engaging setting where the player is motivated to achieve a goal.

Games have been used as research tools for a variety of purposes e.g. using Space Fortress to investigate skill acquisition [9]; examining post-completion errors [3, 1]; considering gender differences in virtual and real worlds [7] and comparing forms of interaction [2, 16]. The more recent trend towards crowdsourcing as a way to create and collate scientific data (e.g., iSpot [6]) and “games with a scientific purpose” such as Foldit [10], indicate how video games can be used an effective way to recruit large numbers of people and engage them in scientific problems. Foldit is a multiplayer online game where players manipulate protein structures represented as visual puzzles in order to unlock the secrets of protein folding (see [13], for a comparison of the algorithms produced by Foldit players and scientists).

Using games as a research tool therefore has the benefits of being able to reach a wider user base thus increasing the probability of user error occurrence. We argue that such a gaming context in particular has potential to provide a paradigm for investigating the kinds of number entry errors that can occur when using a medical device.

Number Entry Interfaces

Number entry is an essential process of many interactive systems requiring users to specify numeric values required by an application. The application then uses this value to fulfil a higher level goal the user intends to achieve. The accuracy of this process (number entry) is particularly essential in the design of safety critical systems since the consequences of error could be really high. For example, out by 10 (or tenfold)

errors have been well reported in literature and have been regarded as a serious risk to patient safety [14, 8].

Unlike text entry, the numeric values specified in a number entry interface have limited scope for error correction. Whereas in text entry, there is a set alphabet whose elements are concatenated according to rules of a specific language to form valid words in that language, numeric quantities specified as numbers usually have a specific meaning and are valid in most combinations. Occasionally, syntax errors do occur in number entry and those can be blocked and alerted to the user [22]. The next sections describe the initial design of a game intended to fill this role.

DESIGN OF GAME

The design of this game adheres to the four defining traits of a game defined by McGonigal [15], i.e., the game has a clear goal, rules, feedback system and voluntary participation. Our game is a simulation based game where the player assumes the role of a nurse on a busy ward whose primary goal is to save all the virtual patients encountered in the game.

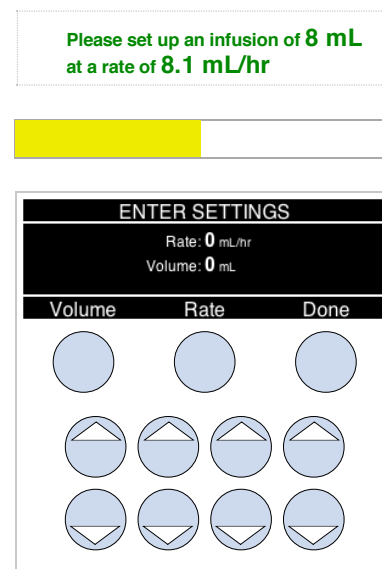


Figure 3. The view of the infusion pump when a player is entering the settings required by the patients.

The player is expected to do this by correctly setting up the infusion pump that corresponds to each patient. An infusion

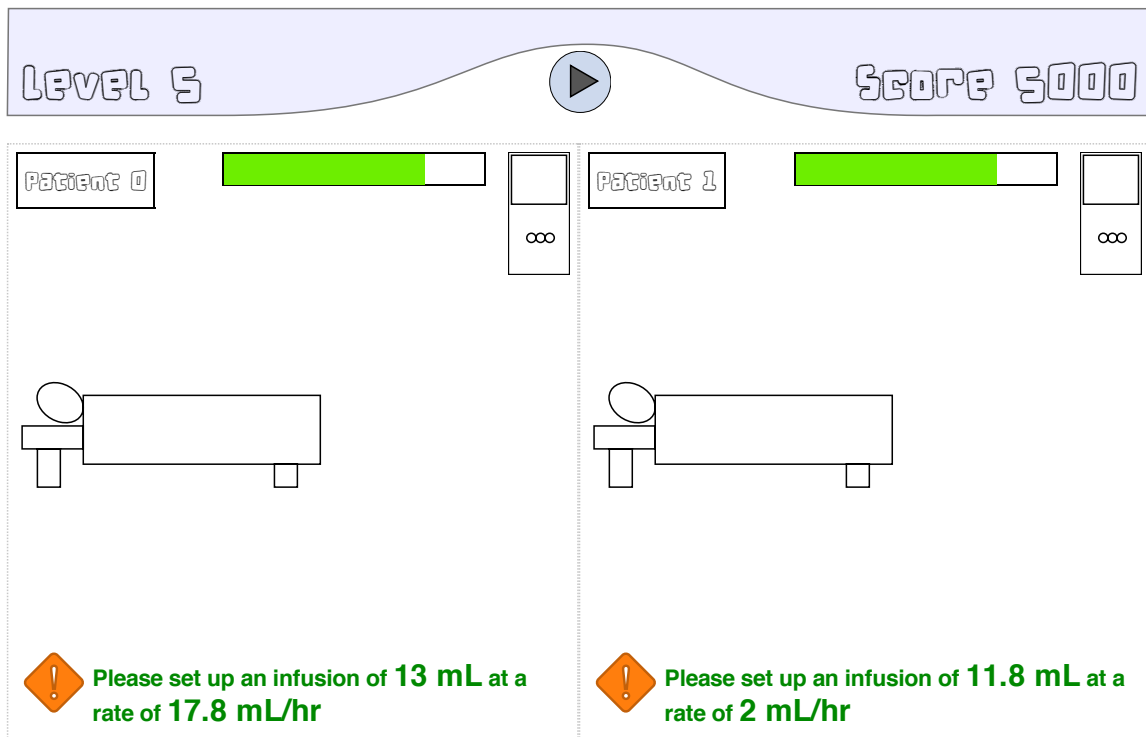


Figure 2. A view of the game at a level with two virtual patients.

pump is set up by entering the correct rate and volume required by the patient for an infusion. The required settings are displayed in an instruction for each virtual patient.

Human experiments sometimes use a microworld to simulate aspects of a real task. Here we introduce specific gaming elements such as points, high scores, levels and a progress bar that shows how much time is left to enter the required settings. These elements make it a game rather than just a task and they increase the possibility that players enjoy the process of playing. This of course changes the intrinsic aim - in a microworld it would be to actually save patients whereas in the gaming context it is score higher points and move up the levels.

The game is designed to explore the effect of task and interface style on performance (both speed and accuracy). Two forms of number entry tasks are present in the game. At the beginning of each level, the player will need to enter new values into the pump. This implies changing the pump parameters from zero to the requested value. If more than one setting is required by the same virtual patient within a level, then entry of that setting requires changing the pump parameters from a non-zero number to the requested value.

The player progresses through different levels up to level 14. As the game level increases, the game complexity also increases. The complexity of the game is defined by the following parameters. The first is the type of interface used. We ranked the complexity of the interfaces based on GOMS estimates on task completion times and the relative speed ranking for number entry interfaces presented in [18]. Another com-

plexity factor is the number of patients on the virtual ward. This starts from one virtual patient to a maximum of four virtual patients. We also vary the number of different interfaces on the ward for levels with more than one virtual patient.

The infusion pump used in the game is minimalist, containing a display, 3 fixed soft buttons for context sensitive menus and a number entry interface. There are four different number entry interfaces used in the game as shown in Figure 1. Each interface works differently and each is in use on commercial medical devices around the world. The different variations of number entry interfaces allow for corresponding different styles of infusion pumps.

The game has two main views: an overview and a focused view. The overview shows all the virtual patients being managed by the player. This view contains information such as the health level of each patient, the instruction about the settings (i.e., rate and volume) required for the patients infusion. The number of patients starts at one and increases to a maximum of four as the player progresses through the levels of the game. The other view is a focused view where players may interact with the pump corresponding to the patient they intend to attend to. If nothing is done while in the overview mode, the health levels of all patients who have not yet received a correct dose progressively depletes. When in the focused view mode, only the health of the patient in focus progressively depletes. All patients out of view remain stable until they are in focus. A patients health improves when they have received the correct dose displayed in the instruction through their infusion pump. The game is over when all patients health have been completely depleted.

In an attempt to promote accuracy and speed, players score points for all correct entries and are rewarded with extra points that are proportional to how quickly they enter the numbers.

The game can be accessed at <http://is.gd/stpdemo>.

DISCUSSION

The problem of studying low frequency errors, particularly in safety critical contexts could be addressed by using a gaming context to run experiments on a number of participants several order of magnitudes larger than would typically be recruited for lab studies. This brings about a number of advantages and disadvantages as we discuss below.

There can be no control over the context in which the game is played and as a result, it might be difficult to ascertain the factors that affect the results obtained. Very little of these variables e.g., anonymous information about the users device, can be measured and these need to be taken into account during analyses of the results.

Due to the gaming nature of the study, it is possible that the results may only apply to the playful context of the game which may be different to a work context. We will argue that the objective of the study, i.e., to understand the performance of several number entry interfaces, is unlikely to be affected by the underlying goal of the user. We believe the details of the motivation for performing a task as ubiquitous as number entry is unlikely to affect how people enter numbers in reality.

On the other hand, due to the potential increase in number of participants, the statistical power of the experiment increases meaning that the effects of variables in the design of the experiment are more likely to be found if there are any. The potential diversity in the sample of participants taking part in the experiment also means that results obtained would be more generalisable than those from a typical lab experiment.

If the game is well designed and thus engaging, then experimental participants may essentially forget they are taking part in a study. Consequently, they may behave more naturally and so make the results more natural. In contexts with low error rates such as that which we are studying, obstacles that aim to increase the error rate may be more naturally placed in a gaming context. These include placing very tight time constraints on higher levels of the game, increasing the numbers of patients that need attention or increasing the number of different interfaces used to enter numbers at any single level as we have done in the design of our game.

The aim of the number entry game is to understand the various performance differences (both speed and accuracy) between different number entry interfaces and explore the space of possible errors given different styles of interfaces. User actions when entering the settings of the infusion pumps will be logged during the game. All data logged will be anonymous and stored securely for the duration of the project. Data from this game will be collected over the next 6 months and analysed to provide insights into minimising errors and optimising speed for number entry interfaces in a safety critical context.

To increase the number of players of the game and the chances of players returning to play the game, we have introduced a social aspect to the game which allows players share their highscores on social media websites such as facebook and twitter.

Specifying numbers correctly is a safety critical task that requires precision and very often speed. The advent of computing and different forms of interactive devices have brought about a diverse variety of ways for specifying numbers to interactive systems. In practice, number entry errors rates are very low usually between 1-5% [19], hence studying them requires a more aggressive experimental design to provoke errors or larger scale experiment than what is currently found in typical user studies. We have proposed the use of games to attract a large user base and we presented an example of a game designed for this purpose.

ACKNOWLEDGEMENTS

Funded as part of CHI+MED (see www.chi-med.ac.uk): Multidisciplinary Computer- Human Interaction research for the design and safe use of interactive medical devices project, EPSRC Grant Number EP/G059063/1. Special thanks to Abigail Cauchi, Rimvydas Rukšenis and Sarah Wiseman for design suggestions in early versions of the game and Frank Soboczenski for useful feedback.

REFERENCES

1. Back, J., Cheng, W., Dann, R., Curzon, P., and Blandford, A. Does being motivated to avoid procedural errors influence their systematicity? In *People and Computers XX Engage*, N. Bryan-Kinns, A. Blandford, P. Curzon, and L. Nigay, Eds. Springer London, 2007, 151–157.
2. Bianchi-Berthouze, N., Kim, W., and Patel, D. Does body movement engage you more in digital game play? and why? In *Affective Computing and Intelligent Interaction*, A. Paiva, R. Prada, and R. Picard, Eds., vol. 4738 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2007, 102–113.
3. Byrne, M. D., and Bovair, S. A working memory model of a common procedural error. *Cognitive Science* 21, 1 (Jan. 1997), 31–61.
4. Calvillo-Gmez, E., Gow, J., and Cairns, P. Introduction to special issue: Video games as research instruments. *Entertainment Computing* 2, 1 (2011), 1–2.
5. Cauchi, A. Differential formal analysis: evaluating safer 5-key number entry user interface designs. In *Proceedings of the 4th ACM SIGCHI symposium on Engineering interactive computing systems*, EICS '12, ACM (New York, NY, USA, 2012), 317320.
6. Clow, D., and Makriyannis, E. iSpot analysed: participatory learning and reputation. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, LAK '11, ACM (New York, NY, USA, 2011), 3443.

7. de Castell, S., Campbell, S., Jenson, J., Shipulina, O., Cimen, A., and Taylor, N. The eyes have it: Measuring spatial orientation in virtual worlds to explain gender differences in real ones. In *CHI 2010 Workshop on Video Games as Research Instruments* (Oct. 2010).
8. Doherty, C., and Donnell, C. M. Tenfold medication errors: 5 years experience at a university-affiliated pediatric hospital. *Pediatrics* (Apr. 2012).
9. Donchin, E. Video games as research tools: The space fortress game. *Behavior Research Methods* 27, 2 (1995), 217–223.
10. Good, B. M., and Su, A. I. Games with a scientific purpose. *Genome Biology* 12, 12 (2011), 135.
11. Kaushal, R., Bates, D. W., Landrigan, C., McKenna, K. J., Clapp, M. D., Federico, F., and Goldmann, D. A. Medication errors and adverse drug events in pediatric inpatients. *JAMA: the journal of the American Medical Association* 285, 16 (Apr. 2001), 2114–2120. PMID: 11311101.
12. Keohane, C. A., Hayes, J., Saniuk, C., Rothschild, J. M., and Bates, D. W. Intravenous medication safety and smart infusion systems: lessons learned and future opportunities. *Journal of infusion nursing: the official publication of the Infusion Nurses Society* 28, 5 (Oct. 2005), 321–328. PMID: 16205498.
13. Khatib, F., Cooper, S., Tyka, M. D., Xu, K., Makedon, I., Popovi, Z., Baker, D., and Players, F. Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences* (Nov. 2011).
14. Lesar, T. S. Tenfold medication dose prescribing errors. *The Annals of Pharmacotherapy* 36, 12 (Dec. 2002), 1833–1839.
15. McGonigal, J. *Reality Is Broken: Why Games Make Us Better and How They Can Change the World*. Penguin Group, The, 2011.
16. McMahan, R. P., Ragan, E. D., Leal, A., Beaton, R. J., and Bowman, D. A. Considerations for the use of commercial video games in controlled experiments. *Entertainment Computing* 2, 1 (2011), 3–9.
17. National Patient Safety Agency. Safety in doses: medication safety incidents in the NHS. Tech. rep., National Patient Safety Agency, 2007.
18. Oladimeji, P. Towards safer number entry in interactive medical systems. In *Proceedings of the 4th ACM SIGCHI symposium on Engineering interactive computing systems*, EICS '12, ACM (New York, NY, USA, 2012), 329332.
19. Oladimeji, P., Thimbleby, H., and Cox, A. Number entry interfaces and their effects on error detection. In *Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction - Volume Part IV*, INTERACT'11, Springer-Verlag (Berlin, Heidelberg, 2011), 178185. conference.
20. Reason, J. *Human Error*. Cambridge University Press, 1990.
21. Taxis, K., and Barber, N. Ethnographic study of incidence and severity of intravenous drug errors. *BMJ* 326, 7391 (Mar. 2003), 684.
22. Thimbleby, H., and Cairns, P. Reducing number entry errors: solving a widespread, serious problem. *Journal of the Royal Society Interface* 7, 51 (Oct. 2010), 1429–1439.
23. Vicente, K. J., Kada-Bekhaled, K., Hillel, G., Cassano, A., and Orser, B. A. Programming errors contribute to death from patient-controlled analgesia: case report and estimate of probability. *Canadian Journal of Anesthesia/Journal canadien d'anesthésie* 50, 4 (2003), 328332.
24. Westbrook, J. I., Rob, M. I., Woods, A., and Parry, D. Errors in the administration of intravenous medications in hospital and the role of correct procedures and nurse experience. *BMJ Quality & Safety* (June 2011).
25. Wirtz, V., Taxis, K., and Barber, N. D. An observational study of intravenous medication errors in the united kingdom and in germany. *Pharmacy World & Science: PWS* 25, 3 (June 2003), 104–111.