# Development and usability evaluation of a multimedia e-learning resource for electrolyte and acid-base disorders

by

**Mogamat Razeen Davids**

Dissertation presented for the degree of

**Doctor of Philosophy**

in the Department of Medicine

Faculty of Medicine and Health Sciences at Stellenbosch University

Supervisors: Professor Mitchell L Halperin and Professor Usuf ME Chikte

March 2015

## DECLARATION

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

This dissertation includes four original papers published in peer-reviewed journals and two unpublished papers. The work also includes, as artefacts, two iterations of the multimedia e-learning resource on which the papers are based. The development of these artefacts, and the development and writing of the papers (published and unpublished) were the principal responsibility of myself and, for each of the cases where this is not the case, a declaration is included in the dissertation indicating the nature and extent of the contributions of co-authors.

Date: .............................

## ABSTRACT

We have developed an innovative multimedia e-learning resource, the Electrolyte Workshop, to provide students and clinicians with instruction and the opportunity for simulated practice in managing electrolyte and acid-base disorders. Our teaching approach is built around relevant physiology and makes use of real cases and storytelling to engage the learner. We have documented the challenges encountered during the development process and have made recommendations for the managing of similar projects.

While there are many factors that must be in place to ensure successful e-learning, this dissertation focuses on an important but under-appreciated factor, namely the usability of the computer interface. Usability describes how easy technology interfaces are to use and is routinely evaluated and optimized in the software development industry. This is not yet the case with e-learning, especially in the area of medical education. Poor usability limits the potential benefit of educational resources, as this means that learners will struggle with the interface as well as with the challenges of the content presented.

A comprehensive usability evaluation of our Electrolyte Workshop was completed. This included testing with typical end-users, where data were collected via standardized questionnaires and by observing and analysing their interactions with our application. We employed heuristic evaluation as an additional approach and assembled a panel of experts to evaluate our application against a set of heuristics, or principles of good interface design.

Many serious usability problems were identified, thus severely limiting the potential educational impact of our Electrolyte Workshop. There was a striking disconnect between the objective measures of usability and self-reported questionnaire data. Our user-testing data make a useful contribution to the debate on how many users are required to find most of the usability problems in an interface. Heuristic evaluation proved to be a very efficient approach. However, both user testing and

heuristic evaluation detected serious problems which were missed with the other method.

These evaluations informed a comprehensive revision of our application and we could then compare the original with an optimized version in a randomized trial. We found large improvements in objective usability measures, which are likely to increase the satisfaction and motivation of learners. There were similar scores on measures of learning. This was not surprising as our participants were all relatively high-knowledge learners and not novices as regards the subject matter.

Our study clearly indicates that the usability evaluation of e-learning resources is critical, and provides an example of how clinician-teachers can improve the usability of the resources they develop. Usability should be evaluated as a routine part of the development and implementation of e-learning materials, modules and programmes. This should start with the earliest versions of the resource, when making changes is easier and less costly. We have demonstrated that a combination of methods should be employed and have highlighted the utility of heuristic evaluation. An iterative approach should be followed, with several cycles of testing and re-design. User testing should always include the study of objective usability measures and not rely only on self-reported measures of user satisfaction.

## OPSOMMING

Ons het ’n innoverende multimediahulpbron vir e-leer, die Electrolyte Workshop, ontwikkel om studente en klinici van ’n onderrighulpmiddel sowel as die geleentheid vir gesimuleerde oefening in die hantering van elektroliet en suur-basis stoornisse te voorsien. Ons onderrigbenadering is gegrond op relevante fisiologie en maak gebruik van werklike gevalle en vertelkuns om die leerder te betrek en te boei. Ons het die uitdagings gedurende die ontwikkelingsproses opgeteken en aanbevelings oor die bestuur van soortgelyke projekte gedoen.

Hoewel suksesvolle e-leer van etlike faktore afhang, konsentreer hierdie verhandeling op ’n belangrike dog onderskatte faktor, naamlik die bruikbaarheid van die rekenaarkoppelvlak. Bruikbaarheid verwys na die gemak waarmee tegnologiekoppelvlakke gebruik kan word, en word gereeld in die sagtewareontwikkelingsbedryf beoordeel en verbeter. Tog is dit nog nie die geval met e-leer nie, veral op die gebied van mediese onderrig. Swak bruikbaarheid beperk die moontlike voordeel van opvoedkundige hulpbronne, aangesien leerders voor die dubbele uitdaging van ’n ingewikkelde koppelvlak én die voorgeskrewe inhoud te staan kom.

’n Omvattende bruikbaarheidsbeoordeling is van die Electrolyte Workshop onderneem. Dit het toetsing met tipiese eindgebruikers ingesluit, waarvoor data met behulp van gestandaardiseerde vraelyste ingesamel en gebruikers se interaksie met die toepassing waargeneem en ontleed is. Ons het heuristiese evaluering as bykomende benadering gebruik en ’n kennerspaneel saamgestel om ons toepassing aan die hand van ’n stel heuristiek, oftewel beginsels van goeie koppelvlakontwerp, te beoordeel.

’n Hele aantal ernstige bruikbaarheidsprobleme is uitgewys, wat die moontlike opvoedkundige impak van die Electrolyte Workshop erg beperk. Daar was merkbare teenstrydigheid tussen die objektiewe bruikbaarheidsmaatstawwe en die selfaangemelde vraelysdata. Ons gebruikerstoetsdata lewer ’n waardevolle bydrae

tot die debat oor hoeveel gebruikers nodig is om die meeste van die bruikbaarheidsprobleme met ’n koppelvlak te ontdek. Heuristiese evaluering was ’n baie doeltreffende benadering. Tog het gebruikerstoetsing op sekere ernstige probleme afgekom wat heuristiese evaluering misgekyk het, en andersom.

Hierdie beoordelings het as grondslag gedien vir ’n omvattende hersiening van die toepassing, waarna ons die oorspronklike weergawe in ’n verewekansigde proef met ’n verbeterde weergawe kon vergelyk. Die objektiewe bruikbaarheidsmaatstawwe het groot verbeterings getoon, wat waarskynlik leerders se tevredenheid en motivering sal verhoog. Leermaatstawwe het soortgelyke tellings opgelewer. Dít was egter te wagte gewees, aangesien die deelnemers almal betreklik ingelig was oor die vakmateriaal, eerder as nuwelinge.

Ons studie het bevestig dat die bruikbaarheidsbeoordeling van e-leerhulpbronne noodsaaklik is, en bied ’n voorbeeld van hoe klinici-opvoeders bruikbaarder hulpbronne kan ontwikkel. Bruikbaarheid behoort as ’n roetinedeel van die ontwikkeling en inwerkingstelling van e-leermateriaal, -modules en -programme beoordeel te word. Dit behoort reeds by die vroegste weergawes van die hulpbron te begin, wanneer dit makliker en goedkoper is om veranderinge aan te bring. Ons het ook getoon dat ’n kombinasie van metodes gebruik behoort te word, en het die nut van heuristiese evaluering beklemtoon. ’n Herhalende benadering moet gevolg word, met etlike siklusse van toetsing en herontwerp. Gebruikerstoetsing behoort altyd die beoordeling van objektiewe bruikbaarheidsmaatstawwe in te sluit, en moenie slegs op selfaangemelde maatstawwe van gebruikerstevredenheid staatmaak nie.

## ACKNOWLEDGEMENTS

I wish to express my appreciation for all the support which enabled me to complete this project successfully:

I owe a huge debt of gratitude to my supervisors, Professors Mitch Halperin and Usuf Chikte, both truly wonderful mentors, for their inspiration, guidance and support.

My PhD project has its origins in the unforgettable year I spent with Mitch Halperin at the Division of Nephrology of St. Michael's Hospital and the University of Toronto. This unique opportunity was made possible by a fellowship award from the International Society of Nephrology.

The talented team of developers at PixelProject who worked on the Electrolyte Workshop, especially Preston and Rahle.

The colleagues and students who generously gave of their time to participate in my usability evaluation studies. Louise Archer and her team in the Clinical Skills Lab who provided the space and computers, and Althea Goosen who very ably coordinated the data collection for our randomised trial.

Justin Harvey from the Centre for Statistical Consultation who assisted with data analysis and, in particular, the Monte Carlo simulation. Karen Grimmer-Somers made a substantial contribution to the paper on the user testing study.

My colleagues in the Division of Nephrology and the Department of Medicine, especially Prof Rafique Moosa and Dr Johan Nel.

All the colleagues who proofread my manuscripts or listened to my presentations, and offered many helpful suggestions.

Susan van Schalkwyk and all the members of our "PhD Support Group".

Funding was received via a grant from the South African Universities Health Sciences Information Technology Consortium, and from Stellenbosch University via two grants from the Fund for Innovation and Research into Learning and Teaching and a Doctoral Fellowship Award from the Faculty of Medicine and Health Sciences. A period of sabbatical leave was granted by Tygerberg Hospital.

Ismail Kola gave me a first taste of research as a lab assistant when I was an undergraduate student at UCT. Anton Doubell was my supervisor on a very exciting MMed study, the first research project I could call my own and one which led to my first publications. The late Stephen Hough supported me throughout, and gave valuable feedback on the earlier drafts of my PhD protocol.

Finally, my family for their love and support: my wife, Gadija, my children, Taahirah and Thaabit, and my parents, Rahma and (the late) Achmat Davids.

# TABLE OF CONTENTS

# CHAPTER 1

# INTRODUCTION

## INTRODUCTION

E-learning is now part of the medical education mainstream [1-3] and has the potential to offer the flexibility of personalized instruction, interaction, collaboration and an engaging, even immersive, learning experience [1]. Learners may be allowed to control the content, sequence, pace and time of learning, and also the medium of delivery. Creative educators are using animation, simulations and virtual 3-D learning environments [4] to create innovative learning resources. Computer-based ‘virtual patients’, for instance, hold particular promise for assisting in the development of clinical reasoning ability [5].

Some authors have advocated the increased use of simulations as an “ethical imperative” [6] and argue that the exposure of patients to the possibility of harm resulting from trainees’ lack of experience can only be justified once approaches that do not put patients at risk have been maximized. Additional advantages of simulations include the ability to provide exposure to uncommon medical conditions and a variety of clinical presentations. Errors can be allowed, and even encouraged, as they provide valuable learning opportunities.

Developing innovative e-learning materials can, however, be very time-consuming and expensive. A 2007 survey of virtual patient development at 108 United States and Canadian medical schools revealed that these computer-based simulations took an average of 16.6 months to complete and that 85% of them cost over $10 000 [7]. It is therefore important to ensure that the time and money invested is justified by the educational impact.

### Improving the educational impact of e-learning in medical education

There are many factors that must be in place to ensure successful e-learning [8]. This study elucidates two critically important but under-appreciated factors. The first is the management of cognitive load and the second, the main focus of the dissertation, is the usability of the computer interface.

The section which follows briefly discusses (i) the development of expertise in clinical problem solving, (ii) cognitive load theory as it pertains to e-learning, (iii) the usability of computer interfaces, and (iv) our teaching approach, which is built on relevant physiology and employs a case-based, narrative format.

### *i. The development of expertise in clinical problem solving*

Medical experts solve most clinical problems by pattern recognition, without resorting to analytical, pathophysiological reasoning [9, 10]. They are able to do this, with good diagnostic accuracy, by drawing on an extensive database of ‘illness scripts’ stored in long-term memory. However, when problems are unusual or complex, the expert can make the shift to analytical reasoning, marshalling extensive relevant basic science knowledge to address the problem [11]. This is often required in disciplines like anaesthesiology, intensive care medicine and nephrology, which are rooted in the basic sciences [11, 12].

Expertise in clinical problem solving is very case-specific [13]. Our challenge is to help students to develop expertise which can be effectively applied when they encounter related but different cases later. Such transfer of expertise is very difficult to achieve [13-15], but can be facilitated by active learning and “deliberate practice” [16] with carefully selected examples [17]. This facilitates the abstraction of the underlying concepts, helping learners to develop a fund of domain-specific knowledge and improving the transfer of clinical reasoning ability from one problem to another. E-learning offers the possibility of fostering deep learning and the transfer of expertise by providing immersive, interactive learning experiences, exposure to multiple cases, and opportunities for deliberate practice. It is important, however, to manage cognitive load and optimize the usability of e-learning resources if we are to realize their full educational potential.

### *ii. Managing cognitive load*

There is a growing body of evidence supporting the benefit of designing learning materials consistent with Sweller's cognitive load theory [18] and Mayer's cognitive theory of multimedia learning [19]. These theories are based on a model of human cognitive architecture that views learning as involving active processing of information by working memory via separate visual and auditory channels. Figure 1 illustrates the main elements of the cognitive theory of multimedia learning.

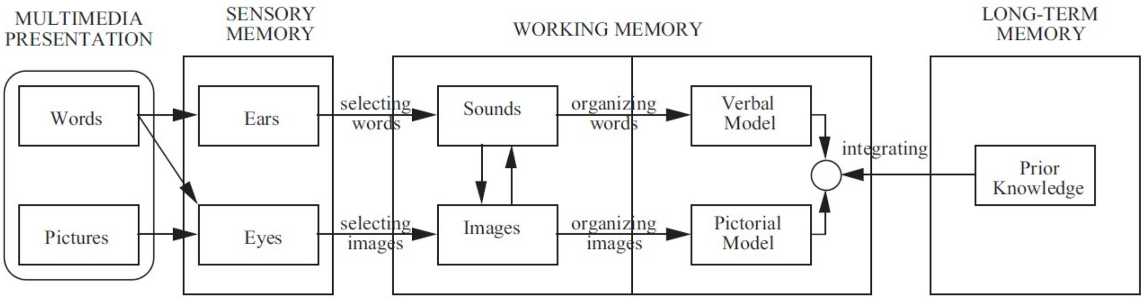



Figure 1. The cognitive theory of multimedia learning. From Mayer, R.E., *The Cambridge Handbook of Multimedia Learning*, 2005, New York: Cambridge University Press, page 37.

Working memory has a very limited capacity. Any load that does not directly contribute to learning is considered extraneous and is likely to impede learning, especially when the material to be learnt is difficult and already has a high intrinsic cognitive load [20, 21]. Difficult content has multiple interacting elements of information which must be assimilated simultaneously for learning to occur [20, 21].

Instructional methods which reduce extraneous cognitive load free up memory and facilitate learning [18, 19]. For example, Mayer [22] has recommended several evidence-based principles to reduce extraneous cognitive load when designing multimedia learning resources. These include the coherence principle, which states that all irrelevant material should be eliminated, the signalling principle, which involves highlighting essential material, and the contiguity principle, which involves placing printed words near the corresponding graphics. Implementation of these principles in the design of learning materials had a significant positive impact on learning in multiple experiments, with medium to large effect sizes (Cohen's *d* of 0.52 to 1.19).

### *iii. Usability*

The International Standard, ISO 9241-11, formally defines usability as the "[e]xtent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [23]. Usability is a concept from the discipline of human-computer interaction that describes how easy user technology interfaces are to use [24]. A user interface should be so intuitive and self-evident that even inexperienced users can accomplish tasks successfully [25]. Design approaches that optimize usability are common in the software development industry, but this is not the case with e-learning, especially in the area of medical education [26].

High usability of e-learning materials is essential in ensuring maximum educational impact [27, 28], especially when the material to be learnt is complex [20, 21]. Poor usability limits the potential benefit [26, 29] by imposing an extraneous cognitive load as users struggle with the interface as well as with the challenges of the content presented.

The two main types of usability evaluation approaches are "usability inspection" and "user testing" [30-32]. Usability inspection involves experts evaluating the application against established design principles [30], while empirical user testing involves typical end-users interacting with the application. Evaluations may be conducted in a wide range of settings, from sophisticated usability laboratories [33] through to informal settings using paper prototypes and think-aloud protocols [34]. Selecting which measures of usability to use is difficult. Some are subjective and others objective; all have their own cost and time requirements, and all examine a particular aspect of usability. The objective measures include parameters such as successful task completion and error rates, while subjective measures include satisfaction, perceived workload, and flow [35].

### *iv. Our teaching approach*

We are developing learning resources to assist students and practicing clinicians in developing expertise in managing electrolyte and acid-base disorders. These are clinical problems which are commonly encountered and which may be life-threatening. This area is highly integrative and quantitative, and it is one that students and clinicians find particularly difficult to master [36]. Our learning resources have been developed in accordance with our teaching approach, which is based on physiology and uses authentic cases and storytelling as a vehicle for instruction.

*Learning around the basic sciences*

Our approach is based on an understanding that learning in this area is most effective when it is built around the relevant basic sciences [11, 12, 36-38]. We use patient data from real cases, and include multiple examples of a range of electrolyte and acid-base disorders. The analysis begins with a focus on the key biochemical abnormality in a particular case, identifying and anticipating threats for the patient, and then analysing the clinical and laboratory data using simple principles of physiology like mass balance and the need for electroneutrality. There is an emphasis on understanding whole-body physiology, deductive reasoning, and quantitative analysis. The intention is to foster deeper learning and the development of sound mental models based on the underlying physiology. Ultimately the aim is more accurate clinical diagnoses and therapies, and better patient outcomes.

*Using narratives: 'Clinical detective stories'*

To increase the engagement of the learner, our learning resources employ an informal, narrative style. This is best illustrated by the series of teaching articles we have published in the "Masterclasses in medicine" section of the *Quarterly Journal of Medicine* (http://qjmed.oxfordjournals.org/). These 'clinical detective stories' use challenging cases to teach our method of analysis. The 'scene of the crime' is the medical ward, the emergency unit or the intensive care unit, and the 'victim' is the patient with a serious electrolyte or acid-base disorder. The 'cast' includes the medical team, ranging from the medical student to the specialist consultant. To play the role of mentor and expert clinical detective, a ‘legend’ from the past has been ‘resurrected’. Professor RA McCance features in the cases of electrolyte and acid-base disorders, Dr RA Phillips for the case of cholera, and Professor Hans Krebs for cases where glucose and energy metabolism is the central theme.

While the medical team contributes recent insights from molecular biology, our expert relies mainly on enduring principles and knowledge which were available before the modern era. Our expert guides and challenges, always returning the focus to basic principles. He helps the team to solve its case and ultimately summarizes the case and ties up the loose ends. At key points in the story, the learner is invited to consider possible diagnoses, anticipate potential complications or decide upon a therapeutic course of action. Drawing the learner into the narrative fits well with the cognitive apprenticeship model which advocates learning in real-world situations and making the thought processes of experts visible to students [39].

The clinical detective story concept for the Masterclasses series of papers was developed by Professors Mitch Halperin, Yehouda Edoute and Razeen Davids during the latter’s fellowship at the University of Toronto in 2000–2001. Appendix 2 includes a selection of these papers, as well as three others illustrating our physiology-based teaching approach.

The value of clinical stories has been summarized by Cox [40] as follows: they recount striking examples that expand our knowledge and expertise in handling similar cases; they are a unit of clinical work; they provide a framework that links all the objective and subjective details around the case; and they explain the influences that determined what diagnostic and management decisions were actually made. Carefully collated case stories can comprise the real-life clinical curriculum.

## Central theme and aims of this research project

While some researchers have found significant learning effects from optimizing usability [41, 42], others have reported improvements in efficiency, satisfaction, or motivation. These latter effects are important in the light of the high dropout rate from e-learning courses [43]. Motivated and self-regulated learners are more likely to persist and succeed in e-learning environments, and optimizing usability can make an important contribution to their satisfaction and motivation.

The central theme of this project is the development and usability evaluation of an innovative multimedia resource for electrolyte and acid-base disorders. The hypothesis is that e-learning resources developed in accordance with best practices as regards user interface design can result in significant improvements in measures of usability, and also in measures of learning.

### ***Outline of the project***

*Aim 1: Building the artefact: an interactive multimedia e-learning resource*

A web-based, interactive, multimedia application has been developed to provide instruction and hands-on experience in managing electrolyte and acid-base disorders. Called the “Electrolyte Workshop”, it consists of case-based tutorials organized into two main sections. In the “WalkThru” section, the concept is ‘look and learn’, analogous to the use of worked-out examples in other disciplines [44]. In the “HandsOn” section, cases are interactive and each includes a treatment simulation where users can select from a menu of therapies to ‘treat’ their patient and receive immediate feedback via animations and text messages. Chapter 2 provides links to the Electrolyte Workshop and describes its submission to an international peer-reviewed repository of health education resources.

*Aim 2: Description of our teaching approach, the development process of the Electrolyte Workshop and an initial usability evaluation*

Chapter 3 explains our teaching approach and describes the development of our Electrolyte Workshop by a team of Flash® developers. We document the challenges encountered during this process and make recommendations for the managing of similar projects. A first evaluation using a standardized user satisfaction questionnaire is also reported.

As our primary target audience is postgraduate trainees and practicing clinicians, we conducted user testing with a group of medical practitioners which included trainees in internal medicine and qualified specialists in internal medicine, nephrology and endocrinology. Morae® usability software (http://www.techsmith.com) was used to facilitate the capture and analysis of information from each testing session. Running in the background, Morae® recorded all participant interactions with our application. This included voice, webcam video of facial expressions, video of all on-screen activity as well as mouse clicks and keyboard activity. Chapter 4 reports on this study.

This chapter also explores the question of the number of participants needed for the usability evaluation of e-learning resources. Establishing the minimum number of users needed is important as it affects the costs and time involved, with usability evaluation more likely to be neglected as resource requirements increase. We conducted a Monte Carlo simulation to determine how many users would be sufficient to test our application and thereby contribute to the debate around this question.

*Aim 4: Usability evaluation by heuristic evaluation*

As teachers who operate in a resource-constrained environment, we investigated whether using one of the inspection methods might be an efficient alternative to testing with end-users. Heuristic evaluation is the most widely used inspection method and involves experts evaluating an interface against a set of generally accepted principles for good design (the heuristics) [45]. We assembled a panel of experts who each applied a set of commonly used heuristics to identify usability problems with our Electrolyte Workshop. The serious problems identified by this method were compared with those previously found by user testing. Chapter 5 reports on this study.

*Aim 5: Revision of the Electrolyte Workshop*

The information gained from the usability evaluations described in Chapters 2, 3 and 4 informed a comprehensive revision of our application. All identified usability problems were addressed and we were then able to compare the original with the optimized version in a randomized trial. Chapter 2 provides hyperlinks to the two versions of the Electrolyte Workshop**.**

*Aim 6: Effect of improving the usability of an e-learning resource: a randomized trial*

Using a randomized trial, we investigated whether addressing the usability problems identified in our Electrolyte Workshop had resulted in measurable improvements in usability and in improvements in learning. Postgraduate trainees in internal medicine and anaesthesiology were randomly assigned to the original or optimized versions. Both subjective and objective measures of usability were studied, and questions which tested recall of information and transfer of problem-solving ability were used as measures of learning. Chapter 6 reports on this study.

## REFERENCES


1. Ellaway R and Masters K. *AMEE Guide 32: e-Learning in medical education Part 1: Learning, teaching and assessment.* Med Teach 2008; **30**(5):455-73.
2. Chumley-Jones HS, Dobbie A and Alford CL. *Web-based learning: sound educational method or hype? A review of the evaluation literature.* Acad Med 2002; **77**(10):S86-93.
3. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ and Montori VM. *Internet-based learning in the health professions: a meta-analysis.* JAMA 2008; **300**(10):1181-96.
4. Hansen MM. *Versatile, immersive, creative and dynamic virtual 3-D healthcare learning environments: a review of the literature.* J Med Internet Res 2008; **10**(3):e26.
5. Cook DA and Triola MM. *Virtual patients: a critical literature review and proposed next steps.* Med Educ 2009; **43**(4):303-11.
6. Ziv A, Wolpe PR, Small SD and Glick S. *Simulation-based medical education: an ethical imperative.* Acad Med 2003; **78**(8):783-8.
7. Huang G, Reynolds R and Candler C. *Virtual patient simulation at U.S. and Canadian medical schools.* Acad Med 2007; **82**(5):446-51.

8. Childs S, Blenkinsopp E, Hall A and Walton G. *Effective e-learning for health professionals and students—barriers and their solutions. A systematic review of the literature—findings from the HeXL project.* Health Info Libr J 2005; **22**:20-32.

9. Patel VL, Groen GJ and Frederiksen CH. *Differences between medical students and doctors in memory for clinical cases.* Med Educ 1986; **20**(1):3-9.

10. Eva KW, Norman GR, Neville AJ, Wood TJ and Brooks LR. *Expert-novice differences in memory: a reformulation.* Teach Learn Med 2002; **14**(4):257-63.

11. Norman GR, Trott AL, Brooks LR and Smith EKM. *Cognitive differences in clinical reasoning related to postgraduate training.* Teach Learn Med 1994; **6**:114-20.

12. Norman G. *The essential role of basic science in medical education: the perspective from psychology.* Clin Invest Med 2000; **23**(1):47-51.

13. Eva KW, Neville AJ and Norman GR. *Exploring the etiology of content specificity: factors influencing analogic transfer and problem solving.* Acad Med 1998; **73**(10):S1-5.

14. Elstein AS, Shulman LS and Sprafka SA. *Medical problem solving: an analysis of clinical reasoning*. Cambridge, MA: Harvard University Press 1978.

15. Patel VL and Cranton PA. *Transfer of student learning in medical education.* J Med Educ 1983; **58**(2):126-35.

16. Guest CB, Regehr G and Tiberius RG. *The life long challenge of expertise.* Med Educ 2001; **35**(1):78-81.

17. Patel VL, Glaser R and Arocha JF. *Cognition and expertise: acquisition of medical competence.* Clin Invest Med 2000; **23**(4):256-60.

18. Sweller J, van Merrienboer JJG and Paas FGWC. *Cognitive architecture and instructional design.* Educ Psychol Rev 1998; **10**(3):251-96.

19. Mayer RE. *Multimedia learning*. Cambridge, UK: Cambridge University Press 2001.

20. Sweller J. *Cognitive load theory, learning difficulty, and instructional design.* Learn Instr 1994; **4**(4):295-312.

21. Sweller J and Chandler P. *Why some material is difficult to learn.* Cogn Instr 1994; **12**(3):185-233.

22. Mayer RE. *Applying the science of learning to medical education.* Med Educ 2010; **44**(6):543-9.

23. Abran A, Khelifi A, Suryn W and Seffah A. *Usability meanings and interpretations in ISO standards.* Software Qual J 2003; **11**(4):325-38.

24. Nielsen J. (2012) *Usability 101: Introduction to usability.* [online] Nielsen Norman Group. Retrieved from http://www.nngroup.com/articles/usability-101-introduction-to-usability/. [Accessed 31 August 2014].

25. Krug S. *Don't make me think! A common sense approach to Web usability* (2nd ed.). Berkeley, CA: New Riders 2006.

26. Sandars J. *The importance of usability testing to allow e-learning to reach its potential for medical education.* Educ Prim Care 2010; **21**(1):6-8.

27. Boling E and Sousa G. *Interface design issues in the future of business training.* Bus Horiz 1993; **36**(6):54.

28. Brinck T, Gergle D and Wood SD. *Usability for the Web: designing Web sites that work* (1st ed.). San Francisco, CA: Morgan Kaufmann Publishers 2002.

29. Zaharias P. *Usability in the context of e-learning.* Int J Technol Human Interact 2009; **5**(4):37-59.

30. Nielsen J and Mack RL. *Usability inspection methods*. New York: Wiley 1994.

31. Jones MG and Farquhar JD. *User interface design for web-based instruction.* In Khan BH, Ed. *Web-based instruction*. Englewood Cliffs, NJ: Educational Technology Publications 1997; 239-44.

32. Dumas JS and Salzman MC. *Usability assessment methods.* Rev Hum Factors Ergon 2006; **2**(1):109-40.

33. Nielsen J. *Usability laboratories.* Behav Inform Technol 1994; **13**(1-2):3-8.

34. Snyder C. *Paper prototyping: the fast and easy way to design and refine user interfaces*. San Diego, CA: Morgan Kaufmann 2003.

35. Hornbæk K. *Current practice in measuring usability: challenges to usability studies and research.* Int J Hum-Comput St 2006; **64**(2):79-102.

36. Dawson-Saunders B, Feltovich PJ, Coulson RL and Steward DE. *A survey of medical school teachers to identify basic biomedical concepts medical students should understand.* Acad Med 1990; **65**(7):448-54.

37. Woods NN, Brooks LR and Norman GR. *The value of basic science in clinical diagnosis: creating coherence among signs and symptoms.* Med Educ 2005; **39**(1):107-12.

38. Woods NN, Neville AJ, Levinson AJ, Howey EH, Oczkowski WJ and Norman GR. *The value of basic science in clinical diagnosis.* Acad Med 2006; **81**(10):S124-7.

39. Hmelo CE. *Problem-based learning: Effects on the early acquisition of cognitive skill in medicine.* J Learn Sci 1998; **7**(2):173-208.

40. Cox K. *Stories as case knowledge: case knowledge as stories.* Med Educ 2001; **35**(9):862-6.

41. Avouris NM, Dimitracopoulou A, Daskalaki S, Tselios NK. *Evaluation of distance-learning environments: impact of usability on student performance.* International Journal of Educational Telecommunications 2001; **7**(4):355-378.

42. Meiselwitz G, Sadera W. *Investigating the connection between usability and learning outcomes in online learning environments.* Journal of Online Learning and Teaching 2008; **4**(2):9.

43. Zaharias P, Poylymenakou A. *Developing a usability evaluation method for e-learning applications: beyond functional usability.* Int J Hum-Comput Interact 2009; **25**(1):75-98.

44. Renkl A. *The worked-out examples principle in multimedia learning.* In Mayer RE, Ed. *The Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press 2005; 229-45.

45. Nielsen J. (1995) *10 Usability heuristics for user interface design.* [online] Nielsen Norman Group. Retrieved from http://www.nngroup.com/articles/ten-usability-heuristics/. [Accessed 31 July 2013].

# CHAPTER 2

# THE E-LEARNING ARTEFACT: THE ELECTROLYTE WORKSHOP

## THE E-LEARNING ARTEFACT: THE ELECTROLYTE WORKSHOP

Our Electrolyte Workshop is freely accessible and can be used by educators as a reusable learning object for the teaching of electrolyte and acid-base disorders. We intend to develop additional content, and ultimately present multiple examples of each type of disorder, allowing learners to encounter key concepts and the same physiology-based approach in a variety of contexts.

The original version is available at http:// http://www.learnphysiology.org/sim1/ and the revised version which was optimised for usability is at http://www.learnphysiology.org/sim2/.

Teachers who are concerned about the usability of the resources they develop are welcome to examine these artefacts to better appreciate the impact of poor usability and see how this has been addressed in the subsequent redesign. This example may also be helpful in making the case for evaluating usability and motivating for resources to be allocated for this important part of the development process.

We have submitted the Electrolyte Workshop to MedEdPORTAL Publications (https://www.mededportal.org/), a repository of health education resources, where it is undergoing peer review. The Instructor Guide which forms part of this submission follows on the next five pages.

# Instructor guide for the Electrolyte Workshop - a multimedia e-learning resource for electrolyte and acid-base disorders

## List of resource files

**Source files:**
ElectrolyteWorkshop2014.zip – containing folders with files of the original (Sim1) and the revised (Sim2) versions.

**Instructor guide:**
InstructorGuide.pdf – this document.

**Supplementary publications:**
These papers describe the development and evaluation of the Electrolyte Workshop in some detail and can be accessed from the journals' websites.

- Davids MR, Chikte UME and Halperin ML. Development and evaluation of a multimedia e-learning resource for electrolyte and acid-base disorders. Advances in Physiology Education 2011; 35(3):295-306.
- Davids MR, Chikte U, Grimmer-Somers K and Halperin ML. Usability testing of a multimedia e-learning resource for electrolyte and acid-base disorders. British Journal of Educational Technology 2014; 45(2):367–381.
- Davids MR, Chikte UME and Halperin ML. An efficient approach to improve the usability of e-learning resources: the role of heuristic evaluation. Advances in Physiology Education 2013; 37:242-8.
- Davids MR, Chikte UME and Halperin ML. Effect of improving the usability of an e-learning resource: a randomized trial. Advances in Physiology Education 2014; 38(2):155-60.

## When, how, and the order in which to use each resource file

Use the revised/optimised version (Sim2) of the Electrolyte Workshop as the learning resource. This is a multimedia application built in Flash which provides instruction on electrolyte and acid-base disorders, and the opportunity for deliberate practice through an interactive treatment simulation.

The Electrolyte Workshop can be accessed by double-clicking the **index.html** file in the **Sim2 folder**. Once the application is launched the navigation is easy and self-evident.

Start with the WalkThru section and the case of acute hyponatraemia related to the use of the drug 'ecstasy'. The teaching approach in this section of the application is 'look-and-learn', analogous to the use of worked examples in other disciplines. It demonstrates how an expert would interpret the patient data and embark on therapy. Animations illustrate changes in body fluid compartment sizes, brain size, blood pressure, and plasma sodium concentrations.

The case in the HandsOn section can be used next. This is a case of chronic hyponatraemia due to Addison's disease. After viewing the introductory slides which describes the case and the treatment goals, the learner reaches the interactive treatment simulation where he/she is able to select from a menu of therapies to 'treat' the patient. Feedback is provided via animations and text messages.

The glossary provides definitions/explanations for terms which may be unfamiliar and can be accessed from its tab at the top right of the screen or from hyperlinks on the slides.

## The purpose/goal of the resource (including educational objectives)

The Electrolyte Workshop can be considered a 'reusable learning object' and used as supplementary material in teaching electrolyte and acid-base disorders (see A and B below). It could also be used by educators as a real-world example to teach about usability and make the case for routinely evaluating e-learning resources at their institutions (see C below). Usability describes how easy technology interfaces are to use and is routinely optimized in the software development industry. This is seldom done with e-learning in medical education. Poor usability limits the potential benefit of educational resources, as learners struggle with the interface as well as with the challenges of the content presented.

**Educational objectives:**

**A. Case 1 (WalkThru case of acute hyponatraemia) - LEARNERS should be able to:**

1. Describe the major body fluid compartments with respect to their volumes.
2. Describe the development of acute hyponatraemia due to water overload and the dangers to the patient.
3. Manage acute, symptomatic hyponatraemia using hypertonic saline. Calculate the correct dose to use once a target plasma sodium concentration has been selected.
4. Appreciate the importance of water absorption in the GI tract as a reason for a continued fall in plasma sodium concentration after admission to hospital.

**B. Case 2 (HandsOn case of chronic hyponatraemia with treatment simulation) - LEARNERS should be able to:**

1. Describe the development of chronic hyponatraemia due to Addison's disease and the dangers to the patient.
2. Use simple blood and urine tests to help in the assessment of extracellular fluid volume.
3. Appreciate the importance of urine chemistry in the diagnosis of electrolyte disorders.
4. Use the simulation to practice the treatment of hyponatraemia and, in particular, develop confidence in the accurate prescription of fluid therapy.

**C. Versions 1 (original) and 2 (revised) of the Electrolyte Workshop, together with associated publications - EDUCATORS should be able to:**

1. Explain the concept of usability and how critical it is to learning.
2. Compare the two versions and appreciate what makes for good and bad usability.
3. Identify usability problems in their own learning resources that could impact negatively on learning.
4. Improve the development processes of their e-learning resources to follow an iterative approach with routine evaluation of usability.

## The conceptual background (why and how it was created)

We are clinician-teachers with an interest in developing learning resources to assist students and practicing clinicians in developing expertise in managing electrolyte and acid-base disorders. These challenging clinical problems are common and may be life threatening. Our teaching approach is based on the relevant physiology and uses authentic cases and storytelling as a vehicle for instruction. The analysis begins with a focus on the key biochemical abnormality in a particular case, identifying and anticipating threats for the patient, and then analysing the clinical and laboratory data using simple principles of physiology like mass balance and the need for electroneutrality. There is an emphasis on deductive reasoning and quantitative analysis.

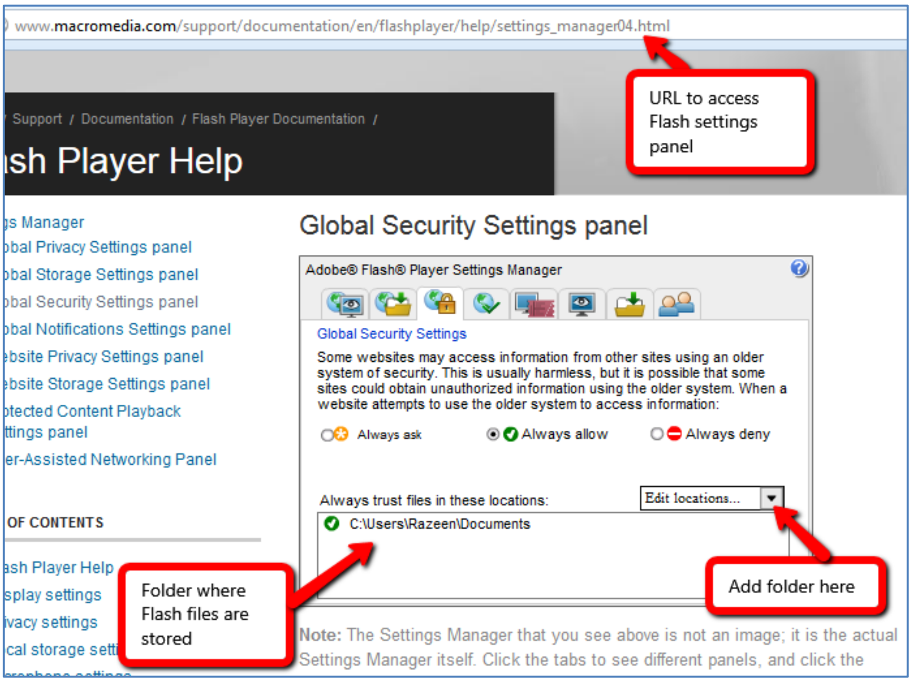
A Web-based resource is immediately available to a wide audience and the use of multimedia and interaction promotes engagement and active learning. Ultimately, the intention is to foster deeper learning and produce better patient outcomes.

## Practical implementation advice

The Electrolyte Workshop can be considered a 'reusable learning object' and used as supplementary material in teaching electrolyte and acid-base disorders.

### 1. Running the application off a local machine (your own hard drive).

As the Electrolyte Workshop is developed in Flash it will need the **Flash Global Security Settings** to be updated before it will run off a local machine. See screenshots below. These can be accessed at http://www.macromedia.com/support/documentation/en/flashplayer/help/settings_manager04.html. Add the folder where the Electrolyte Workshop files are located as a trusted location by selecting the "Always allow" radio button and then navigating to the folder where you have unzipped the files.

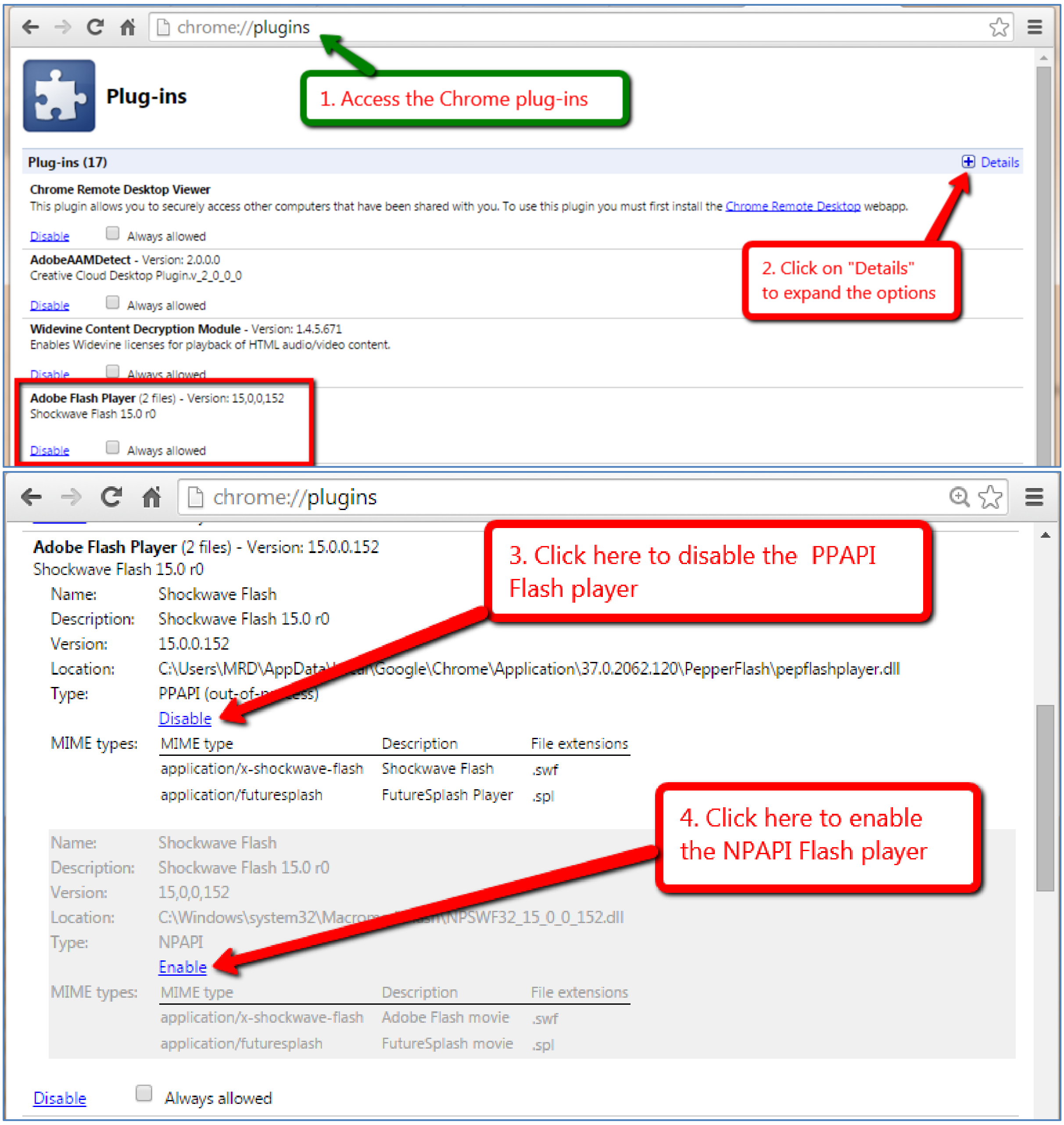



The application can then be launched by **double-clicking on the index.html** files in the Sim1 and Sim2 folders.

**2. Running the application off your hard drive using Google Chrome as your browser.**

The Chrome browser has its own internal Flash player and does not use the Flash Global Security Settings in the normal way. You may not be able to select a “trusted folder” as above so that you can launch the application in Chrome. The solution below disables that internal Flash player and enables the normal Adobe plug-in as follows:

enter 'chrome://plugins/' in the Chrome url (address) bar
click the 'Details' [+] button at the right
deactivate the 'PPAPI' Flashplayer
activate the 'NPAPI' Flashplayer




Once you have done this you can update the Flash Global Security Settings as in point 1 above and then run the application.

**3. Running the application without a browser using the free Adobe Flash Player.**

An alternative solution is to avoid using a browser and directly run the Shockwave Flash file in each folder. Simply double-click on the **ElectrolyteWorkshop.swf** file. This requires that the free Adobe Flash Player be installed.

**4. Running the application from our website (not the peer-reviewed version).**

This is a further option but is not the peer-reviewed version as hosted by MedEdPORTAL. When accessed via the Internet there are no special requirements. The URL's are http://www.learnphysiology.org/sim1/ for the original version and http://www.learnphysiology.org/sim2/ for the fully revised version.

**How has it been successfully deployed?**

LEARNERS: I have used the Electrolyte Workshop at Stellenbosch University to supplement our teaching to undergraduate medical students and to postgraduate trainees in internal medicine and nephrology. A link to the website is provided in their course documentation and I briefly show them the application in class. They are encouraged to explore the application on their own to reinforce the material covered in class.

EDUCATORS: We have conducted a comprehensive evaluation of the application, focusing on usability. This informed an extensive revision and the two versions were then compared in a randomized trial. The relevant publications are listed above. At our own university, and elsewhere, I have delivered presentations using these two versions to illustrate the effect of poor usability on the educational impact of e-learning resources and to make the case for an iterative development approach with the routine evaluation of usability.

**Limitations of the resource and ideas for improving/expanding it**

Content is currently limited to these two cases. It represents the beginnings of a resource that can offer a rich learning experience and assist students and colleagues to acquire expertise in the challenging area of electrolyte and acid-base disorders. We need to continue content development, so that we present multiple examples of each type of disorder, allowing our users to encounter key concepts and the same physiology-based approach in a variety of contexts.

There is a move away from Flash and towards HTML5 which offers an open source structure, multimedia support without the need for browser plug-ins, better power efficiency as regards device battery life, and better accessibility of content to search engines. For the future development of our multimedia e-learning resources we are exploring authoring tools which can publish to multiple formats including HTML5 and Flash, as well as to a format optimized for the iPad.

# CHAPTER 3

# DEVELOPMENT OF THE ELECTROLYTE WORKSHOP

## The underlying teaching approach, the development process and an initial usability evaluation via user satisfaction questionnaire

Published paper:

Davids MR, Chikte UME, Halperin ML. Development and evaluation of a multimedia e-learning resource for electrolyte and acid-base disorders. Advances in Physiology Education 2011;35(3):295-306.

# Development and evaluation of a multimedia e-learning resource for electrolyte and acid-base disorders

**Mogamat Razeen Davids,[1] Usuf M. E. Chikte,[2] and Mitchell L. Halperin[3]**

[1]*Division of Nephrology and Department of Medicine, Stellenbosch University and Tygerberg Hospital, Cape Town;* [2]*Department of Interdisciplinary Health Sciences, Stellenbosch University, Cape Town, South Africa; and* [3]*Division of Nephrology, St Michael's Hospital and University of Toronto, Toronto, Ontario, Canada*

Received 2 December 2010; Accepted 9 May 2011

**Davids MR, Chikte UME, Halperin ML.** Development and evaluation of a multimedia e-learning resource for electrolyte and acid-base disorders. *Adv Physiol Educ* 35: 295–306, 2011; doi:10.1152/advan.00127.2010.—This article reports on the development and evaluation of a Web-based application that provides instruction and hands-on practice in managing electrolyte and acid-base disorders. Our teaching approach, which focuses on concepts rather than details, encourages quantitative analysis and a logical problem-solving approach. Identifying any dangers to the patient is a vital first step. Concepts such as an "appropriate response" to a given perturbation and the need for electroneutrality in body fluids are used repeatedly. Our Electrolyte Workshop was developed using Flash and followed an iterative design process. Two case-based tutorials were built in this first phase, with one tutorial including an interactive treatment simulation. Users select from a menu of therapies and see the impact of their choices on the patient. Appropriate text messages are displayed, and changes in body compartment sizes, brain size, and plasma sodium concentrations are illustrated via Flash animation. Challenges encountered included a shortage of skilled Flash developers, budgetary constraints, and challenges in communication between the authors and the developers. The application was evaluated via user testing by residents and specialists in internal medicine. Satisfaction was measured with a questionnaire based on the System Usability Scale. The mean System Usability Scale score was 78.4 ± 13.8, indicating a good level of usability. Participants rated the content as being scientifically sound; they liked the teaching approach and felt that concepts were conveyed clearly. They indicated that the application held their interest, that it increased their understanding of hyponatremia, and that they would recommend this learning resource to others.

clinical problem solving; Flash; wireframing; prototypes; software development; System Usability Scale; hyponatremia; usability

ELECTROLYTE AND ACID-BASE DISORDERS are clinical problems that are common and may be life threatening. This area is highly integrative and quantitative, and it is one that students and clinicians find particularly difficult to master (9).

Medical experts solve most clinical problems using pattern recognition, drawing on a large domain-specific database of schemata or "illness scripts" (15, 22, 41). When an unusual or complex situation is encountered, however, the expert can draw on extensive relevant basic science knowledge and apply it to the problem (38). This is often required in disciplines such as anesthesiology, nephrology, and intensive care medicine, where much of the clinical reasoning involves the application of physiology (37). Electrolyte and acid-base disorders are typical examples common to these disciplines where an understanding of physiology is central to correct diagnosis and treatment.

As teachers, our challenge is to help students and clinicians develop an expertise in clinical problem solving that can be effectively applied when they encounter related, but different, problems. This transfer of expertise is difficult to achieve (10, 14, 39). It can be facilitated by active learning and "deliberate practice" with carefully selected examples (12, 13, 40, 42). This helps to develop a fund of domain-specific knowledge and facilitates the abstraction of underlying concepts and the transfer of clinical reasoning ability from one problem to another.

Increasingly, medical educators are engaging learners using animations to illustrate dynamic processes and simulations to provide the opportunity to interact with clinical problems. Well-known examples include Chopra's operating room simulator and "Harvey," the cardiology patient simulator (7, 21, 44). A receptive atmosphere now exists for the increased use of simulations following large studies describing preventable injuries to patients as a result of medical error (4, 20, 27, 28). Some authors view the use of simulations as an "ethical imperative" and argue that harm, or exposure to the possibility of harm, to patients in the course of training or resulting from trainees' lack of experience can only be justified once approaches that do not put patients at risk have been maximized (62). Additional advantages of simulations include the ability to provide exposure to uncommon conditions and a variety of clinical presentations. Errors can be allowed, and even encouraged, as they provide valuable learning opportunities. Trainees can then have their first encounters with real patients having already attained higher levels of confidence and proficiency.

Developing good simulations and other e-learning materials can be resource intensive, and it is therefore important to ensure that the time and money invested is justified by the educational impact. The usability of computer interfaces has a major impact on the effectiveness of e-learning but is an aspect that has not been sufficiently emphasized in the medical education literature (26, 33, 45). The concept derives from the field of human-computer interaction and describes how easy a technology interface is to use. The terms "fitness for purpose" and "quality of use" (2) are also used to describe usability. The International Organization for Standardization, in ISO 9241-11 (25), defines usability as the "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use."

Address for reprint requests and other correspondence: M. R. Davids, Division of Nephrology and Dept. of Medicine, Stellenbosch Univ. and Tygerberg Hospital, PO Box 19063, Tygerberg 7505, Cape Town, South Africa (e-mail: mrd@sun.ac.za).

High usability of e-learning materials is an essential element in ensuring maximum educational impact (3, 5), especially when dealing with complex subject matter (49, 50). Poor design can impair learning, as the user has to struggle with challenging content as well as with the technology interface. Reducing such an extraneous cognitive load can lead to large gains in learning efficiency (49); thus, optimizing the usability of learning resources seems essential. A recent review of internet-based medical education (58) reported that learners are more likely to be engaged when the technology is easy to use.

The routine evaluation of usability is well established in the software development industry. An iterative design approach is followed and involves the creation of prototypes, testing them, and making improvements based on the test results (32). This cycle is repeated until performance and usability goals are met, and only then will the application be shipped to the marketplace. This approach is seldom used in the development of e-learning resources, especially in the area of medical education (45).

The two main types of evaluation techniques are empirical user testing, which involves typical end users using the application, and usability inspection, which involves experts evaluating the application against established metrics or design principles (5). Evaluations are conducted in a wide range of settings, from sophisticated usability laboratories (34) through to informal settings using paper prototypes and think-aloud protocols (47). Many tools are available to assist with the collection, analysis, and reporting of usability data (24). These range from self-administered questionnaires (6) to specialized software for recording and analyzing user testing sessions. Selecting which measures of usability to use is difficult. Some are subjective and others objective, all have their own cost and time requirements, and all examine a particular aspect of usability. Among the objective measures are parameters such as successful task completion, completion times, and error rates, whereas subjective measures include aspects such as satisfaction, perceived workload, fun, aesthetics, and flow (23). Nielsen and colleagues (35, 36, 54) have popularized simpler and less expensive methods, pointing out that any usability testing is better than no testing at all, and demonstrating that four to five users are sufficient for each cycle of testing.

We have developed a Web-based multimedia resource to help students and clinicians acquire expertise in the diagnosis and treatment of electrolyte and acid-base disorders. Colleagues in disciplines such as internal medicine, pediatrics, nephrology, endocrinology, emergency medicine, and intensive care medicine are included in our target user population. Our "Electrolyte Workshop" provides instruction and allows for hands-on practice in treating electrolyte disorders via a highly interactive simulation. This e-learning application can be freely accessed at http://www.learnphysiology.org/sim1/. An evaluation of the Electrolyte Workshop was conducted with a group of residents and specialists in internal medicine to determine its level of usability. User satisfaction was measured via a questionnaire based on the widely used and validated System Usability Scale (SUS) (6). The SUS is a low-cost option that is easily administered and scored and is therefore an especially useful tool in resource-constrained environments.

In this article, we briefly describe our general teaching approach and report on the development of our Electrolyte Workshop. We discuss the challenges encountered, outline the resources required, and highlight the lessons learned. Recommendations are made for managing the development of similar projects. We then report on our evaluation of user satisfaction involving a group of specialists and postgraduate trainees in internal medicine.

## METHODS AND RESULTS

Ethical approval for the project was granted by the Health Research Ethics Committee of Stellenbosch University (approval no. N08/05/158).

### *The Underlying Teaching Approach*

Our teaching approach is based on an understanding that learning in this area is most effective when it is built around the relevant basic sciences (9, 37, 38, 59, 60). Building sound mental models based on physiological principles is likely to aid knowledge retention and retrieval, especially when novel or complex problems are encountered and when pattern recognition might not be effective (38). We emphasize an understanding of integrative whole body physiology, deductive reasoning, a focus on concepts rather than details, and quantitative analysis. We emphasize principles of control and point out to our students that if one recognizes the function of a metabolic process, it is often possible to deduce the nature of its likely controls (46). For example, if one views buffering as a way to protect enzymes, receptors, and transporters from an acid load, then the importance of "directing" protons to the bicarbonate buffer system becomes clear. Students are then better able to deduce that this process can be optimized by effective removal of the resultant $CO_2$ through hyperventilation and the provision of an adequate circulation.

We use clinical and laboratory data from real cases in our teaching. Each case starts with the identification of a key abnormality, usually the most abnormal electrolyte or acid-base parameter. This is usually self-evident and often the reason for the consultation. When there is more than one major abnormality present, any one can be selected as the starting point and the analysis followed through to the end. Thereafter, the next abnormality is dealt with in a similar way.

The first step is always to identify and address immediate threats for the patient and to anticipate dangers that may develop later. These often arise as unintended consequences of therapy. The primary imperative is always to "save the patient!"

We continue with the diagnostic workup once urgent therapeutic issues have been addressed. Patient data are analyzed and interpreted focusing on the actual versus expected (appropriate) response to the particular perturbation. We repeatedly emphasize the concept of "an appropriate response" and the ability to recognize it in a given situation. This requires an understanding of the relevant physiology. Students are asked what an appropriate response would be and then what clinical or laboratory data are required to identify whether their patient is responding appropriately, for example, "How should your patient's kidneys respond to hypernatremia?" and "What data would tell you that antidiuretic hormone is acting on the kidneys?"

A problem-solving process that is simple and logical is encouraged. Basic principles such as mass balance and the need for electroneutrality in body fluid compartments are used repeatedly. Other frequently used concepts are those of "driving forces and permeability," which are the elements that determine whether water or solutes will move across cell membranes. Through this line of inquiry and educational approach the student arrives at a functional diagnosis, e.g., "fast sodium absorption disease" in a patient with hypokalemia and hypertension. This is followed by making a structural or anatomic diagnosis, e.g., "overactive epithelial sodium channel disease," and is sometimes followed by assigning a specific diagnostic label, e.g., "Liddle's syndrome." We emphasize the systematic analytic process and not the arrival at the correct diagnostic label. To discourage students from taking shortcuts and jumping to possible final diagnoses too quickly (i.e., guessing!), we often ask them to interpret a set of clinical and laboratory data, specifying that "a diagnostic label is not required."

A quantitative analytic approach is always promoted. For example, "In this 60-kg female with a plasma sodium concentration of 130 mmol/l where we've estimated the extracellular fluid volume to be contracted by ~10%, what would be the magnitude of her sodium deficit?" followed by "What volume of 0.9% saline would be required to correct a sodium deficit of 230 mmol?" Since exact answers are seldom required at the bedside, students are urged to round off numbers and develop their skills at estimating rather than resorting to using a calculator. Another example comes from a case of cholera (61) where severe extracellular fluid contraction completely masked metabolic acidosis. An important teaching point was that one has to consider the content, and not just the concentration, of plasma bicarbonate when assessing acid-base status.

Our teaching approach provided the background for the development of the multimedia, Web-based learning resource that is discussed below. Here, we describe our e-learning application, the development process involved, and the challenges encountered and suggest some recommendations for managing similar projects.

*Development of the Electrolyte Workshop*

*Description of the application.* Flash from Adobe Systems (www.adobe.com) was used for the development of our Electrolyte Workshop, a case-based multimedia application illustrating and simulating the pathophysiology, diagnosis, and treatment of a variety of electrolyte and acid-base disorders. Flash can provide an engaging user experience by producing interactive content that can include pictures, sound, and video. It is well suited to creating rich content for the Web because its files are relatively small.

The Electrolyte Workshop (Fig. 1) consists of two main sections, each currently containing one case-based tutorial on hyponatremia. Eventually we plan to cover all the common electrolyte and acid-base disorders, with several examples of each. In the WalkThru section, the case consists of a series of 14 slides that presents the clinical problem of acute hyponatremia related to ingesting the drug Ecstasy. Through words and pictures, we demonstrate how an expert would interpret the patient data and embark on treatment. Flash animation is used to illustrate and emphasize important changes in body compartment sizes, brain size, blood pressure, and plasma sodium concentrations (PNa). The pace at which information is presented is controlled by the user as s/he navigates from one slide

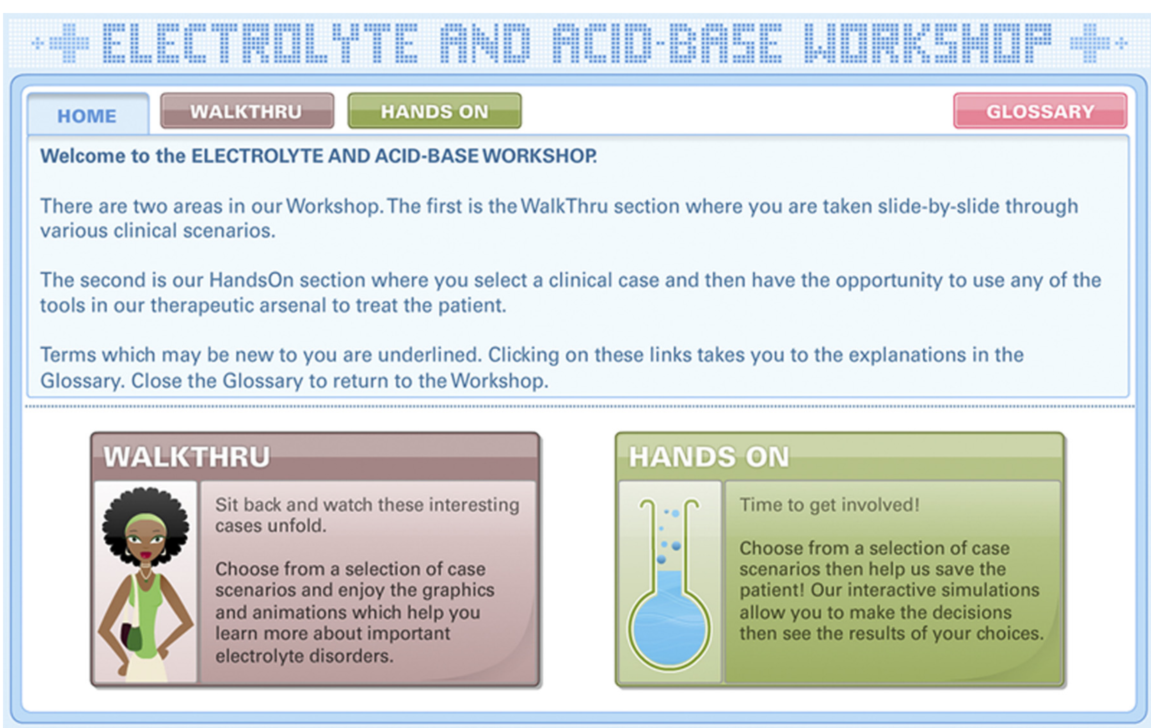



Fig. 1. The Electrolyte Workshop. In the WalkThru section of the application, the user navigates through case scenarios to learn how an expert would analyze the data and embark on treatment. Animation is used to illustrate changes in body compartment sizes, brain size, blood pressure, and plasma sodium concentrations. In the interactive HandsOn section, a "treatment console" allows users to practice managing the patient. The glossary provides explanations for terms that may be unfamiliar. Hyperlinks in the text of the case scenarios link to the appropriate glossary entries.

to the next. One of the interesting aspects highlighted in this case is that of hidden dangers related to stomach contents. Water ingested shortly before admission, and present in the lumen of the gastrointestinal tract, still has to be absorbed and may further aggravate the severe hyponatremia.

The HandsOn section is also self-paced but more interactive. The case is that of chronic hyponatremia due to Addison's disease. It starts off with five "lead-in" slides that describe the clinical problem and highlight the key issues such as the adrenal hormone deficiencies and the contacted extracellular fluid volume. Users are then provided with a "treatment console" (Fig. 2, *A* and *B*), a highly interactive simulation where they are able to select from a menu of therapies (and dosages) and apply their treatment. The main issue here is the danger of too rapid correction of the chronically low PNa, which may result in serious neurological damage. The available therapies include a selection of intravenous fluids, from water through to 3% saline. It also includes drugs (a vasopressin analog and cortisol in this case) and sodium or potassium salts. More than one treatment can be administered; this happens sequentially and not simultaneously, so that feedback can be given after each step via on-screen text and animations. The animations illustrate the effects of treatment choices by showing changes in body fluid compartment volumes, brain size, blood pressure, and PNa. The text messages indicate the success of the interventions applied, for example, "Your patient developed osmotic demyelination and died from serious neurological damage. This was caused by a too-rapid rise in PNa! Try again?" After successful completion of the treatment simulation, the case concludes with a final slide of "take home messages."

The two cases were chosen as they provide striking examples of acute and chronic hyponatremia, respectively. Additional examples will be added later so that so that students encounter the same important principles in a variety of contexts. In both sections we use the teaching approach described above and try to foster a better grasp of the underlying physiology principles (Table 1). The quantitative aspect is stressed consistently to reinforce the importance of accuracy in the assessment of the disorders and in prescribing treatment. The HandsOn simulation provides practice to help users develop a better feel for treating these conditions and, in particular, the ability to use intravenous fluids correctly and confidently.

An informal, conversational style was deliberately chosen for the presentation of the cases. Users are addressed directly using words such as "you" and "your." Mayer and colleagues (29) have shown that this "personalization effect" can improve the engagement of users and promote active learning. As undergraduate and postgraduate health science students are the main target audience, we created a trendy young character called Suzie (Fig. 1, *bottom left*) as the patient in the case scenarios, also with the intention of increasing students' engagement through the compelling need to save her from harm!

To provide support, especially for novice users, there is a glossary (Fig. 3), which can be accessed from the main navigation menu or via text hyperlinks in the cases. Terms that might be unfamiliar or require further explanation are underlined to indicate a hyperlink to the glossary.

*The development process.* We contracted a team of Flash developers with extensive experience in Web design and animation but with no background in the biomedical sciences. After initial discussions, we constructed "wireframes" or "storyboards" to communicate our ideas for the two case scenarios. Microsoft PowerPoint was used for this purpose.

The iterative development process that followed is shown in Figs. 2 and 4. From each set of PowerPoint slides, the developers created static screenshots (.jpeg files) to reflect the different screens or slides in the case. We reviewed these and compiled a list of changes, which were then implemented by the developers. Two iterations were required during this phase. The application was designed to fit into an area of $800 \times 600$ pixels to accommodate users with smaller computer screens. A grayscale version was produced initially, and color was added once all changes had been agreed upon.

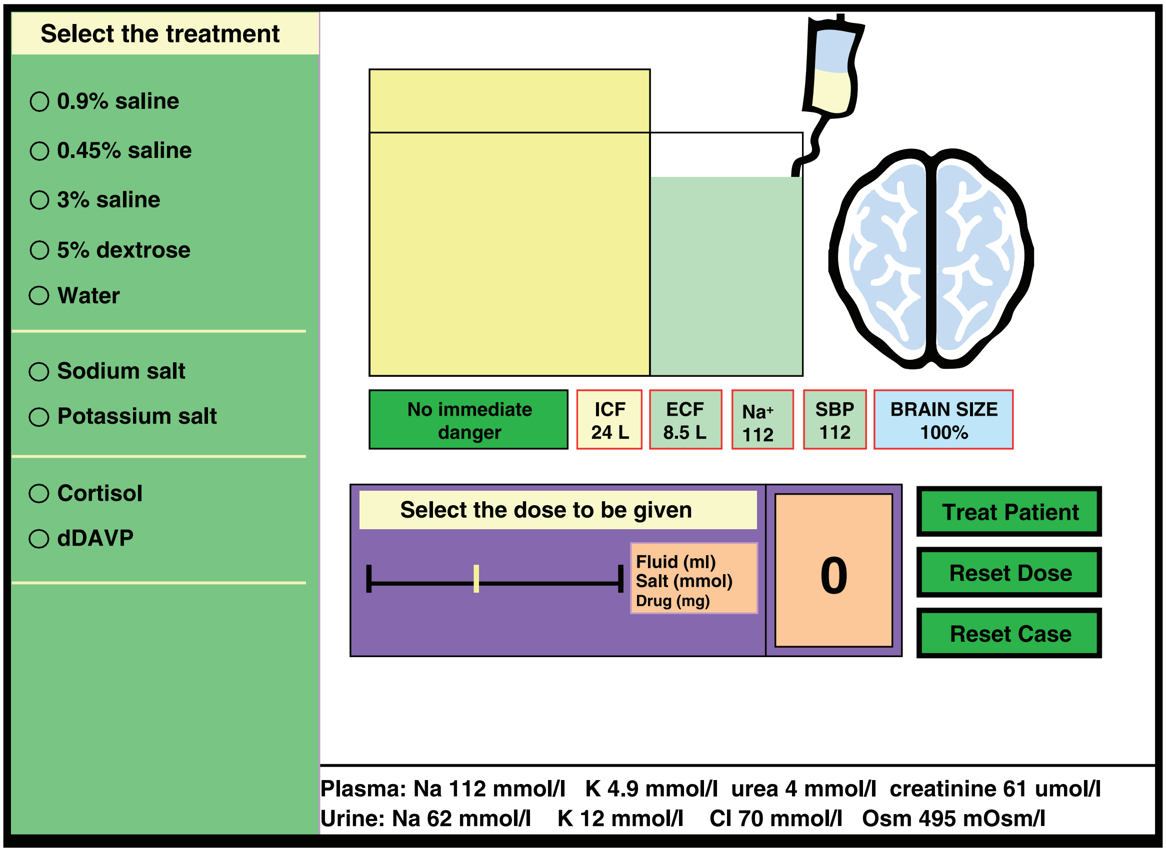
An animated, interactive version was then built using Flash. The ActionScript programming language, which is similar to JavaScript, provided interactivity and controlled the simulation (56). The algorithms and formulae embedded in the ActionScript code were provided by the authors. It calculated the effects of user-selected therapies and then controlled the display of appropriate text messages and changes to graphic elements. For example, in response to the administration of a particular volume of hypotonic fluid, the brain of the patient would swell by a precisely calculated amount, body fluid compartment volumes would increase, and PNa would decrease accordingly. A text message would then be displayed based on the resultant PNa, for example, "PNa has fallen even further! This means that brain cells are swelling. Would you like to try something else?"

After three cycles of development and reviewing and revising the application, the fully animated, interactive version was completed.

*Challenges.* Several challenges were encountered during the development process (Table 2). The shortage of skilled Flash developers has been mentioned. A limited budget was another constraint (see below for a discussion of the costs involved). With respect to the actual development of the simulation, the main challenge was the difficulty in communicating to the developers exactly what was required. Their feedback was that our PowerPoint wireframes were not detailed or accurate enough. As this was a novel project for all concerned, it was difficult to fix all the specifications for the work at the start. Different expectations of the number of iterations that would be allowed became apparent, as did differences in the list of features that were included in the original costing. These issues were amicably resolved through discussion.

Fig. 2. *A* and *B*: the treatment simulation illustrates the iterative development process. This simulation of a patient with chronic hyponatremia offers the user a selection of treatments and dosages and displays important patient parameters, including brain size, fluid compartment volumes, and plasma sodium concentration. Additional feedback is provided by way of text messages. *A*: a wireframe was constructed by the authors using PowerPoint, with each slide representing a "screen" of the application. dDAVP, desmopressin; ICF, intracellular fluid; ECF, extracellular fluid; SBP, systolic blood pressure. *B*: the final, live version of the treatment simulation, which was built using Flash. The images, values, and text messages are dynamic, changing in response to user input.

A


Select the treatment
0.9% saline
0.45% saline
3% saline
5% dextrose
Water
Sodium salt
Potassium salt
Cortisol
dDAVP
No immediate danger
ICF 24 L
ECF 8.5 L
Na+ 112
SBP 112
BRAIN SIZE 100%
Select the dose to be given
Fluid (ml)
Salt (mmol)
Drug (mg)
0
Treat Patient
Reset Dose
Reset Case
Plasma: Na 112 mmol/l K 4.9 mmol/l urea 4 mmol/l creatinine 61 umol/l
Urine: Na 62 mmol/l K 12 mmol/l Cl 70 mmol/l Osm 495 mOsm/l

B

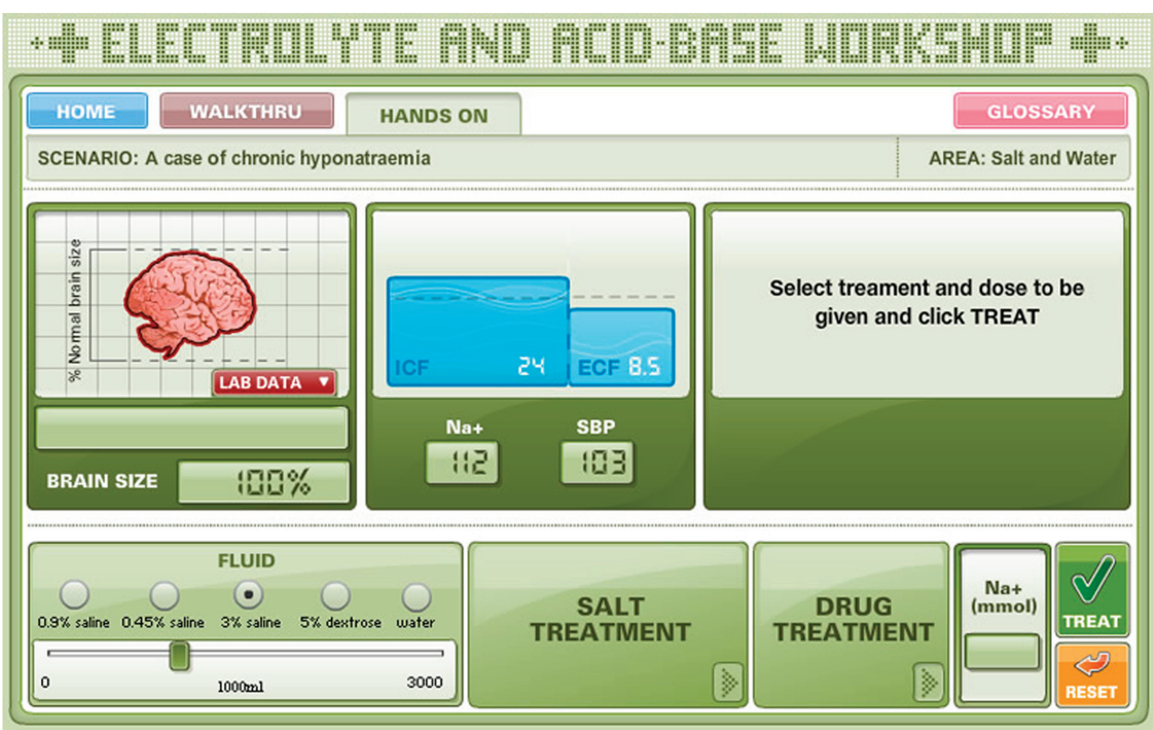

ELECTROLYTE AND ACID-BASE WORKSHOP
HOME
WALKTHRU
HANDS ON
GLOSSARY
SCENARIO: A case of chronic hyponatraemia
AREA: Salt and Water
% Normal brain size
LAB DATA
BRAIN SIZE
100%
ICF 24
ECF 8.5
Na+
112
SBP
103
Select treament and dose to be given and click TREAT
FLUID
0.9% saline
0.45% saline
3% saline
5% dextrose
water
0
1000ml
3000
SALT TREATMENT
DRUG TREATMENT
Na+ (mmol)
TREAT
RESET

Table 1. *Examples of physiology concepts taught in the cases*

| Concept | Comment |
|---|---|
| The PNa is the ratio of $Na^+$ to water in the ECF compartment. | Hyponatremia may be due to a loss of $Na^+$ and/or a net gain of water. |
| Effective osmolality (reflected by PNa) determines the size of the ICF compartment. | Cells are swollen with hyponatremia and shrunken with hypernatremia; brain cells are the most important in this regard. |
| Acute hyponatremia may develop when there is intake of EFW and vasopressin prevents its renal excretion. | Both a source of EFW and vasopressin are needed to develop and sustain acute hyponatremia. |
| A smaller muscle mass results in a greater degree of hyponatremia if a given volume of EFW is retained. | Brain cell swelling is more likely in individuals who are smaller or cachectic if EFW is retained. |
| Water in the stomach can be absorbed after admission and further lower PNa. | Take a good history of fluids ingested and monitor PNa and neurological status carefully. |
| Examine the mass balances for both water and $Na^+$ to predict the change in PNa. | This simple but very useful tool is also called a "tonicity balance." |
| $Na^+$ is the major ECF osmole, and $Na^+$ content therefore determines ECF volume. | The ECF volume should be contracted if hyponatremia is due to a $Na^+$ deficit. |
| $Na^+$ excretion should be minimal when ECF and effective arterial volume are low. | Any $Na^+$ excretion is excessive with a low effective arterial volume. |
| Water excretion may impaired when there is a low effective arterial volume. | This is due to the reduced glomerular filtration rate with low filtrate delivery to the distal nephron and from the nonosmotic release of vasopressin. |
| The major danger with chronic hyponatremia is osmotic demyelination from a too-rapid rise in PNa. | The risk is greater in patients with hypokalemia or malnutrition. |

PNa, plasma $Na^+$ concentration; ECF, extracellular fluid; ICF, intracellular fluid; EFW, electrolyte-free water.

For future projects, the developers have requested that detailed and accurate screen-by-screen wireframes be provided up front. This should include all content and all algorithms and formulae with relevant text messages needed for user feedback in the interactive parts of the application. It should be clearly stated which parameters need to be tracked by the application. Finally, it must be specified at the start which elements need to be editable by the client. This would establish the full extent of the development required. Not only would this be essential for accurate cost assessment, but it provides the basis for a written agreement from which the project can be managed and eventually signed off.

*Resources required.* There is a shortage of skilled Flash developers in South Africa, and more than a year was spent finding a suitable team who were willing and able to execute the project within the available budget. We focused on finding

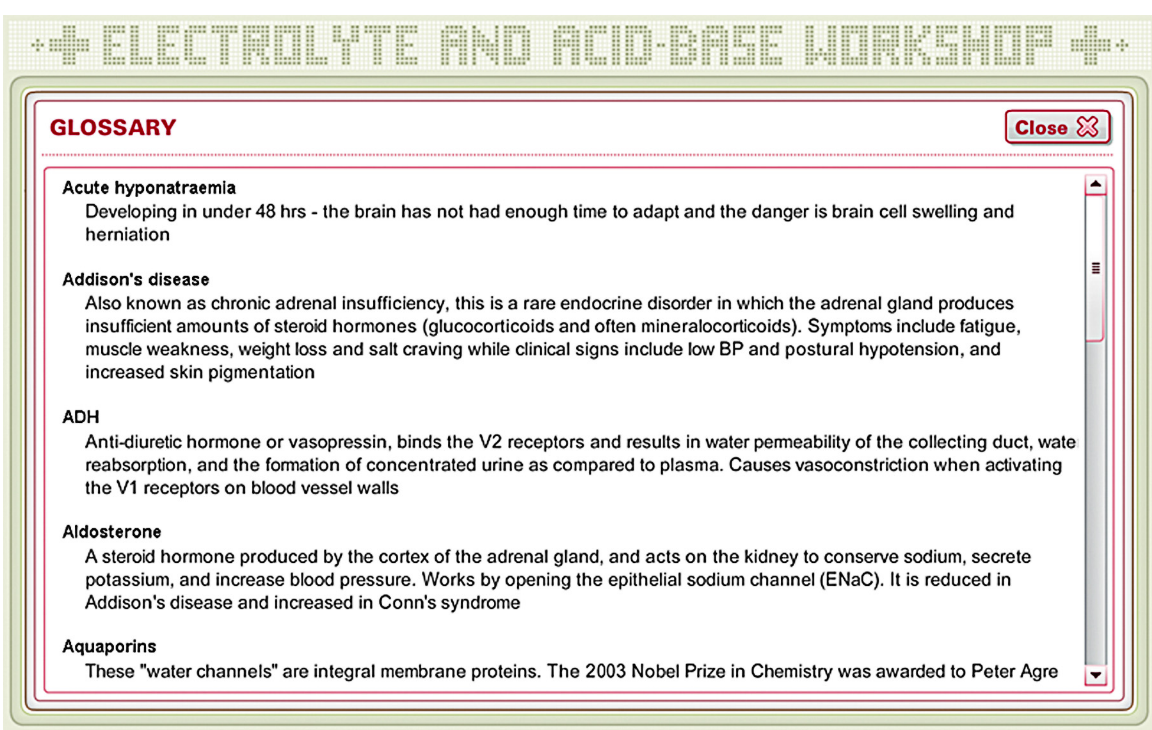



Fig. 3. The glossary. The glossary provides help with terms that may be unfamiliar or need further explanation. It can be accessed via hyperlinks within the text of the cases or from the main navigation tab at the top of the screen.

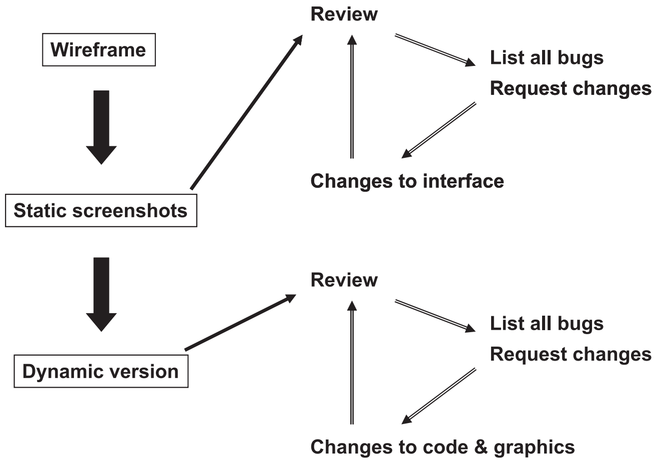



Fig. 4. The development process. From a PowerPoint wireframe, which was supplied by the authors, the developers constructed static screenshots of the application using .jpeg files. These were revised in several iterations until agreement was reached on the look and functionality of the interface. The coding and animation were then added, and, after several more cycles of testing and debugging, the live application was finalized.

developers in the greater Cape Town area so that face-to-face meetings could be held as required.

The development team comprised of one programmer and one designer. The work was completed over a period of 7 mo. Development costs amounted to approximately $5,300 (R40,000; $1 = R7.60), a greatly reduced rate offered because of the academic, not for profit nature of the project and because the interest of the developers was piqued by the project. The usual rate for commercial Flash development in South Africa is $60–85/h (R450–650/h), and a fixed-price estimate for a project such as this would usually be around $12,500 (R95,000).

*Evaluation of the Electrolyte Workshop*

*Participants and testing procedures.* User testing was conducted at two South African academic departments of medicine with 10 residents and 6 specialists in internal medicine, nephrology, and endocrinology. This group is typical of our target user population because their disciplines involve the management of electrolyte and acid-base disorders. Although five users have often been reported to be sufficient to test usability (35), we engaged a larger number of participants to improve the usability error detection rate (16, 53) and to allow us to consider the influence of participants' knowledge and experience. The application was presented via two 15-in. laptop computers, each equipped with a mouse. Participants received written instructions that included information about the purpose of the application and of their involvement, which was to help us improve the application. Their tasks were to view the home page, familiarize themselves with the different sections of the application, and then work through the cases in the WalkThru and HandsOn sections, trying different options in the treatment simulation. They were also asked to look at the glossary. No time limits were set, and participants worked at their own pace.

*Data collection.* Participants each completed a two-page questionnaire at the end of their session. The first part consisted of the 10-item SUS developed by Brooke (6). SUS yields a single number representing a composite measure of the overall usability of the system being studied. The scores have a potential range of 0 to 100, with a score of 70 or greater regarded as acceptable (1). SUS is widely used by usability professionals and is reliable, freely distributed, easy to administer, and easy to score (1, 53). It can be used to provide a point estimate measure of usability and customer satisfaction, compare different tasks within the same interface, compare different versions of a system, and compare competing systems or interface technologies (1).

We used SUS with minor adaptations (Table 3), replacing all occurrences of the word "system" with "application" and changing the word "cumbersome" in *item 8* to "cumbersome/clumsy." Other authors have recommended replacing "cumbersome" with "awkward" to improve understanding, especially when the survey involves non-native English speakers, as was the case with some of our participants (1, 17).

The next section of the questionnaire (Table 4) included 11 statements about various aspects of the application, such as the

Table 2. *Challenges in the development process*

| | Client Experiences | Developer Experiences |
|---|---|---|
| Communication | Very challenging communicating ideas about a complex simulation to developers with no background in biomedical science. The importance of the simulation being both qualitatively and quantitatively accurate was not initially appreciated by the developers. | Difficulties in the client effectively communicating the logic, knowledge, and understanding of all the possible consequences: ". . .exist in your or your students' minds. . .," "needs to be worked out on paper. . ." Only at the end did it become apparent that we are needing to be tracking and reporting on more than one thing, i.e., ECF, PNa, etc.; this was not apparent only from the [PowerPoint] we were given." |
| Expectations that the application would be editable | We expected the application to serve as a template, with the addition of further cases being easier and less costly. We expected to be able to edit and add to the glossary as well as the text on the various slides. | The client expectations of the end product and deliverables versus the available budget was too high. ". . .Expectation that this workshop would be fully editable is not realistic given the existing budget." |
| Project scope and iterations expected before completion | We viewed this as an iterative process where the final product was not clear in our minds at the start. We expected to go through many iterations until the application was "perfect." | Additional features and content were added after the initial discussions [and cost assessment], e.g., the glossary. |

Shown are the challenges as perceived by the authors ("client") and the development team. Both parties were inexperienced with software development projects of this nature. Authors' comments or interpretation are in brackets.

Table 3. *Summary of the SUS results*

| Modified SUS Statements* | Strongly Disagree | | Disagree | | Neutral | | Agree | | Strongly agree | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % |
| *1*. I would like to use this application often if more cases are added. | | | | | | | 4 | 25 | 12 | 75 |
| *2*. I found the application complex. | 4 | 25 | 7 | 43.75 | 2 | 12.5 | 2 | 12.5 | 1 | 6.25 |
| *3*. I thought the application was easy to use. | | | 1 | 6.25 | 2 | 12.5 | 4 | 25 | 9 | 56.25 |
| *4*. I need the support of an expert to be able to use this application. | 10 | 62.5 | 4 | 25 | 2 | 12.5 | | | | |
| *5*. I found the various parts of the application well integrated. | | | 1 | 6.25 | 2 | 12.5 | 8 | 50 | 5 | 31.25 |
| *6*. I thought there was too much inconsistency in the application. | 6 | 37.5 | 8 | 50 | 1 | 6.25 | 1 | 6.25 | | |
| *7*. I would imagine that most of my colleagues would learn to use this application very quickly. | | | | | 1 | 6.25 | 6 | 37.5 | 9 | 56.25 |
| *8*. I found the application cumbersome/clumsy to use. | 9 | 56.25 | 4 | 25 | 1 | 6.25 | 1 | 6.25 | 1 | 6.25 |
| *9*. I felt very confident using the application. | 1 | 6.25 | 1 | 6.25 | 4 | 25 | 5 | 31.25 | 5 | 31.25 |
| *10*. I'll need to learn a lot of things before I could use this application. | 7 | 43.75 | 2 | 12.5 | 1 | 6.25 | 3 | 18.75 | 3 | 18.75 |

*n* = no. of particpants, with *n* = 16 participants in total. SUS, System Usability Survey. Statements were scored with a five-point Likert scale, where *1* = strongly disagree, *3* = neutral, and *5* = strongly agree. *For clarity, all occurrences of the word "system" were replaced with "application," and "cumbersome" in *item 8* was changed to "cumbersome/clumsy."

soundness of the content, the ease of navigation, and whether participants would recommend the resource to others. Participants indicated their level of agreement on a five-point Likert scale. This was followed by questions on the treatment simulation (Table 4) and a question on the length of each case (Table 5). Participants were also asked about the suitability of the application as a learning resource for groups ranging from specialists to medical and nursing students (Table 6).

Participants were then asked to rate their own level of computer literacy (Table 7). Finally, they were asked to comment on things they liked about the application, anything they did not like or which could be improved, and for final suggestions or comments.

*Results of the evaluation.* The results of the SUS are shown in Table 3. The mean score was 78.4 ± 13.8 (range: 45–100). There were no differences between senior (specialists, $n = 6$) and more junior (residents, $n = 10$) colleagues. Mean scores were 82.1 ± 10.5 and 76.3 ± 15.6 ($P = 0.477$) in these two groups, respectively.

On analysis of individual questionnaire items, senior clinicians expressed a greater degree of confidence in using the application ($P = 0.037$), but there were no other differences between the two groups. User satisfaction with various aspects of the Electrolyte Workshop is shown in Table 4. Participants rated the content as being scientifically sound (15 of 16 participants agreed); they liked the clinical detective story approach (14 of 16 participants), the emphasis on key concepts (14 of 16 participants), and felt that these concepts were conveyed clearly (14 of 16 participants). They indicated that the application held their interest (14 of 16 participants), that it increased their understanding of the topic (14 of 16 participants), and that they would recommend this learning resource to others (15 of 16 participants).

A few participants felt that the glossary was not useful (5 of 16 participants) and that navigation was difficult (3 of 16 participants). The treatment simulation was experienced as realistic by 9 of 16 participants (5 participants were neutral) and increased 8 participants' confidence for managing similar

Table 4. *User satisfaction with the Electrolyte Workshop*

| Statements | Strongly Disagree | | Disagree | | Neutral | | Agree | | Strongly Agree | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % |
| Content scientifically sound | | | | | 1 | 6.25 | 5 | 31.25 | 10 | 62.5 |
| Glossary was not useful | 6 | 37.5 | 3 | 18.75 | 2 | 12.5 | 4 | 25 | 1 | 6.25 |
| Liked "clinical detective story" | | | 1 | 6.25 | 1 | 6.25 | 7 | 43.75 | 7 | 43.75 |
| Animations distracted, unhelpful | 8 | 50 | 6 | 37.5 | 1 | 6.25 | 1 | 6.25 | | |
| Key concepts put over clearly | | | | | 2 | 12.5 | 6 | 37.5 | 8 | 50 |
| Did not like the character (Suzie) | 9 | 56.25 | 2 | 12.5 | 5 | 31.25 | | | | |
| Liked emphasis on key principles | | | | | 2 | 12.5 | 6 | 37.5 | 8 | 50 |
| Increased my understanding | | | 2 | 12.5 | | | 10 | 62.5 | 4 | 25 |
| Navigation was difficult | 7 | 43.75 | 5 | 31.25 | 1 | 6.25 | 3 | 18.75 | | |
| Failed to hold my interest | 9 | 56.25 | 5 | 31.25 | 1 | 6.25 | 1 | 6.25 | | |
| Would recommend to others | | | | | 1 | 6.25 | 5 | 31.25 | 10 | 62.5 |
| Treatment simulation | | | | | | | | | | |
| Very realistic | | | 2 | 12.5 | 5 | 31.25 | 4 | 25 | 5 | 31.25 |
| Increased my confidence | 1 | 6.25 | 1 | 6.25 | 6 | 37.5 | 4 | 25 | 4 | 25 |
| Difficult to use | 8 | 50 | 2 | 12.5 | 4 | 25 | | | 2 | 12.5 |
| Too far removed from the real world | 7 | 43.75 | 2 | 12.5 | 6 | 37.5 | 1 | 6.25 | | |

*n* = no. of particpants, with *n* = 16 participants in total. Statements were scored with a five-point Likert scale, where *1* = strongly disagree, *3* = neutral, and *5* = strongly agree.

Table 5. *Participant evaluation of the case length*

| | Very Long | | Long | | Just Right | | Short | | Very Short | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % |
| Case length | | | 1 | 6.25 | 12 | 75 | 3 | 18.75 | | |

*n* = no. of particpants, with *n* = 16 participants in total.

cases (6 participants were neutral). It was considered difficult to use by two participants and too far removed from the real world to be useful by one participant (6 participants were neutral).

Participants considered the application suitable for residents, subspecialty trainees, and specialists. It was also considered suitable for medical students but not for nursing students. Half of the participants considered it suitable for renal and intensive care unit nurses. Of the three participants who indicated that their level of computer literacy was weak, none found the application difficult to use.

Comments from the open-ended questions are shown in Table 8. Participants liked the interactive nature of the application, the real-life cases, and the overall design. Most of the negative comments related to the treatment simulation. These included being limited to apply only one treatment at a time, struggling with the slider control, inadequate guidance and feedback from the application, and an inability to navigate back to the lead-in slides once in the treatment simulation.

## DISCUSSION

A multimedia application was successfully developed to provide instruction in the area of electrolyte and acid-base disorders. It is a reusable learning object that is sharable and that can easily be incorporated into learning management systems. Two case scenarios were built in this first phase to explore the feasibility and optimal design of the application.

While learning from authentic, complex, and ill-defined problems is encouraged by constructivist approaches and facilitates the transfer of expertise, novices may be overwhelmed and demotivated if problems are too difficult (52). The WalkThru section of the application was therefore designed to facilitate learning using worked-out examples (43, 51). This allows students to move from focusing on finding solutions to appreciating the underlying principles or rules. Offering "step-by-step" guidance in this section promoted the development of expertise by "making the thinking of experts visible," in line with the cognitive apprenticeship model (8). This model involves a focus on teaching the processes, the cognitive and metacognitive skills, by which experts solve complex problems.

Deliberate practice in a specific domain is important in the acquisition of expertise (11), and the HandsOn section was designed around this principle. The treatment console is an example of a deterministic simulation, where a given action by the user in a particular situation always produces the same result. This predictability helps novices build confidence in their ability to apply treatment correctly and accurately. In the future, we intend to cater for more experienced users by including probabilistic simulations to model the uncertainty and unpredictability that is always part of managing real patients (19).

Although the initial development included only two case scenarios, eventually multiple examples of each type of disorder will be presented, so that students encounter key concepts and the same step-by-step physiology-based approach in a variety of contexts. Sound physiological principles should provide the framework around which their new knowledge and schemata are built. This facilitates knowledge retrieval and application and also improves the accuracy of the nonanalytic (pattern recognition) components of the clinical reasoning process (60). The provision of a range of cases will also cater for users with different levels of expertise, with more advanced users free to skip some of the simpler WalkThru cases and tackle the more challenging HandsOn cases directly.

Regarding our choice of development platform, we decided on Flash to create a resource that was visually appealing and interactive, increasing the likelihood that students would be engaged and motivated to use it. It also needed to be delivered via the internet so as to be accessible to a wide audience. Learning objects created with Flash can be used with a variety of learning management systems. They can be accessed via any Web browser using the free Adobe Flash Player plug-in, avoiding the problems of cross-browser and cross-platform incompatibility. They will run on almost all personal computers as well as an increasing number of mobile phones and other devices (30). The delivery of rich content with relatively small file sizes is possible because Flash uses vector graphics, which are represented by mathematical formulae, rather than bitmap graphics with their larger file sizes.

Mastering Flash involves a very steep learning curve, however, and thus it is not a realistic option for most educators. Skilled developers are in short supply and are expensive. Since starting the project, we have become aware of many tools that allow nonexperts to build e-learning courses without special programming skills. These range from simple PowerPoint-to-Flash converters through to high-end applications with prebuilt templates, interactions, quizzes, branching logic, and the like. The eLearning Guild (www.elearningguild.com) has compiled a useful report on authoring and development tools (55) that could serve as a starting point for readers interested in developing their

Table 6. *Participant evaluation of the suitability of the application*

| | Subspecialty Registrars | | Registrars | | Specialists | | Medical Students | | Renal and/or Intensive Care Unit Nurses | | Nursing Students | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % |
| Suitability | 15 | 93.75 | 16 | 100 | 14 | 87.4 | 13 | 81.25 | 8 | 50 | 3 | 18.75 |

*n* = no. of particpants, with *n* = 16 participants in total.

Table 7. *Participant evaluation of their own computer literacy*

| | Very Weak | | Weak | | Adequate | | Good | | Very Good | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % |
| My computer literacy | | | 3 | 18.75 | 8 | 50 | 3 | 18.75 | 2 | 12.5 |

*n* = no. of particpants, with *n* = 16 participants in total.

own materials. In the future, we plan to use Flash developers more selectively, thus reducing our dependency on expensive, scarce skills. For a project like this one, for example, we could use PowerPoint to develop many of the slides and the simpler animations ourselves. The more complex animations and the treatment simulation would then be the only parts we would need to "outsource," and these components would then simply be inserted on the appropriate slide of the case. The entire PowerPoint presentation could then easily be converted to Flash using a variety of applications such as Articulate (www.articulate.com), iSpring (www.ispringfree.com), Adobe Presenter (www.adobe.com/products/presenter), and others.

Those e-learning applications that have to be custom built by independent contractors need to produce the desired end product within time and budgetary restrictions and ideally allow for easy expansion and maintenance. Unfortunately, this is the case with only a minority of software development projects. Research by the Standish Group (48) has found that the most common outcome is that projects are late, over budget, or have less than the required features. Many projects are never completed or used, and only around one-third are delivered on time, within budget, and with the required features. For small projects such as ours, we offer some recommendations to increase the chances of a successful outcome: the use of wireframes and prototyping, following an iterative development process, and drawing up a formal written agreement.

### *Recommendations*

*Wireframes and prototypes.* Wireframes and prototypes are used to represent the structure and functionality of a website or desktop application (57) and are constructed early on in the development process, before any artwork or coding is undertaken. They provide a basis for communication between clients and developers, helping to define the functionality of each page (or screen) and the positioning of elements such as navigation menus and search fields. It is important to reach agreement on the user interface and functionality right at the beginning of the project. Once these specifications have been defined, an accurate cost assessment can be done. These critical early steps help to reduce the risks and costs of the software development process. Simple wireframes can be created using paper prototyping (47) or easy-to-use software such as Microsoft PowerPoint, Balsamiq Mockups (www.balsamiq.com), MockupScreens (www.mockupscreens.com), or Pencil Project (www.evolus.vn/pencil), an add-on for the Firefox browser. Higher-

Table 8. *Comments from the open-ended questions*

| Postitive Comments | Negative Comments and Suggestions for Improvement |
|---|---|
| Immediate feedback. All the information immediately available. Realistic clinical scenario. | Add SI units [normal values] to numbers used. |
| A new way of teaching such a difficult topic. | Would have liked to be able to go back to clinical details after reaching treatment console i.e. navigation difficult. |
| Interactivity. Format kept my interest. | Did not like treatment options and approach to them. [only 1 treatment at a time, etc.] |
| User friendly. Colourful, fun. | Treatment console - not enough options, difficult to understand (is fluid given per hr or per 24 hr?), after cortisol administration it gives the option of more treatment - which ones? do they mean fluid? No feedback afterwards from program. Correct answer not given. What is purpose of the "cortisol block"? A bit stressful! |
| Step by step approach. Good explanation; topic well covered. | I would like to be able to use combinations of therapy. One can't go back except to start from the beginning of the case once one is in the treatment console. One of the response messages was not complete |
| I enjoyed the opportunity to change the different treatments and see how it affects the patient. | Technical aspects - cannot click directly onto treatment scroll bar [and pick dose], must scroll [drag slider]. Select radio button - did not know what the radio button was. |
| Very practical and allows a realistic experience with electrolytes and experimentation, seeing how different modalities affect the patient. | I was not aware there was [only] one case scenario; I thought there were more coming so I sat for a long time thinking more were coming. Suggest informing user how many cases there are. |
| It simulated real case scenario very well. | Glossary can be emphasized - I missed this until the end. |
| Visual learning - very applicable. | |
| Relevant character with real-life problem (especially WalkThru case). Extremely well designed with really "funky" layout and illustrations. | |
| I think this is a major leap forward in the training of electrolyte and acid-base disorders. Thank you. | |
| Much improved way of teaching. | |
| Potentially good program especially because hyponatraemia treatment can be very tricky. | |
| A good, practical refreshing model to learn and become more comfortable with electrolyte problems. | |
| Program easy to use and understand - can benefit medical personnel with electrolyte and fluid management. | |
| Thanks, I would definitely use this as a learning aid if there were more cases. | |
| This makes electrolyte physiology fun and knowledge thereof useful. | |

Authors' comments or interpretation are in brackets.

fidelity tools allow for the creation of prototypes with a richer user interface, more interactivity, and even basic conditional logic. They are usually more expensive, with a steeper learning curve, which may be daunting for nontechnical users. Software programs in this category include Flash and Flex (www.adobe.com), Axure (www.axure.com), and Irise (www.irise.com).

*Iterative development process.* An iterative development process usually involves the following: *1*) identifying the basic requirements; *2*) developing a first version of the application, typically a basic prototype; *3*) review(s) by the client and preferably end users; and *4*) revision by the developers based on the feedback received. There may be several iterations of *steps 3* and *4* until there is agreement on the user interface and functionality. The code is then written to transform the prototype into a dynamic, interactive application.

*Written agreement.* A written agreement concluded at the beginning of the project helps with managing the project and preventing disputes. Essential elements include software specifications, timelines and payments, ownership of the software, warranties, and dispute resolution (18). The developer should be required to fix software errors at no charge for a specified period of time. It is advisable to break down a sizable project into discrete parts and link payments to the completion of each part. This also makes it easier to monitor progress and avoids the danger of getting an unsatisfactory product at the end. The intellectual property rights to the software usually reside with the developer. However, many different options can be negotiated, ranging from sole ownership by the client to ownership by the developer with the client merely having a license to use the software. Finally, provision must be made for resolving disputes, preferably through mediation or arbitration.

### *User Satisfaction*

Overall, our Electrolyte Workshop was positively received by the participants in our initial evaluation, with the SUS score of 78.4 indicating a good level of usability. The additional questionnaire items confirmed the satisfaction of participants with the case-based approach and overall design of the application. They considered it useful, thought that it improved their understanding of the topic, and would recommend the resource to others. It was considered to be a suitable learning resource for residents and specialists, our target audience, and also for medical students. Of some concern was the data on the treatment simulation in the HandsOn case. Only 9 of 16 participants found it realistic, and only half felt that it increased their confidence for managing similar problems. Difficulties with the selection and application of treatments as well as inadequate guidance and feedback were highlighted as issues in this interactive part of the application.

### *Conclusions*

In our Electrolyte Workshop, we have the foundation of a multimedia resource that has the potential to offer a rich, immersive learning experience and assist students and colleagues to acquire expertise in the area of electrolyte and acid-base disorders. User testing with the aid of a standardized questionnaire indicated that we achieved a good level of usability. Further evaluation should include objective measures of usability and an assessment of gains in knowledge. The development of e-learning materials of high quality requires a multidisciplinary team that includes content experts, instructional designers, and developers. Implementing good project management, with clarification of roles and expectations, is important in ensuring a successful outcome. Finally, using an iterative development approach with the routine testing of usability is an essential aspect in realizing the full educational potential of the electronic medium.


## ACKNOWLEDGMENTS

The authors thank the team at PixelProject (www.pixelproject.com) for generously providing Flash expertise and Martin Schreiber for a critical review of the final drafts of the manuscript.

## GRANTS

This work was supported by grants from the South African Universities Health Sciences Information Technology Consortium and Stellenbosch University's Fund for Innovation and Research into Learning and Teaching.


## DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

## REFERENCES


1. **Bangor A, Kortum PT, Miller JT.** An empirical evaluation of the system usability scale. *Int J Hum Comp Interact* 24: 574–594, 2008.
2. **Bevan N.** Measuring usability as quality of use. *Software Qual J* 4: 115–130, 1995.
3. **Boling E, Sousa G.** Interface design issues in the future of business training. *Business Horizons* 36: 54, 1993.
4. **Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, Newhouse JP, Weiler PC, Hiatt HH.** Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *N Engl J Med* 324: 370–376, 1991.
5. **Brinck T, Gergle D, Wood SD.** *Usability for the Web: Designing Web Sites That Work*. San Francisco, CA: Morgan Kaufmann, 2002.
6. **Brooke J.** SUS: a "quick and dirty" usability scale. In: *Usability Evaluation in Industry*, edited by Jordan PW, Thomas B, Weerdmeester BA, McClelland IL. London: Taylor & Francis, 1996, p. 189–194.
7. **Chopra V, Gesink BJ, de Jong J, Bovill JG, Spierdijk J, Brand R.** Does training on an anaesthesia simulator lead to improvement in performance? *Br J Anaesth* 73: 293–297, 1994.
8. **Collins A, Brown J, Newman S.** *Cognitive Apprenticeship: Teaching the Craft of Reading, Writing, and Mathematics. Technical Report No. 403* (online). http://www.eric.ed.gov/PDFS/ED284181.pdf [25 May 2011].
9. **Dawson-Saunders B, Feltovich PJ, Coulson RL, Steward DE.** A survey of medical school teachers to identify basic biomedical concepts medical students should understand. *Acad Med* 65: 448–454, 1990.
10. **Elstein AS, Shulman LS, Sprafka SA.** *Medical Problem Solving: an Analysis of Clinical Reasoning*. Cambridge, MA: Harvard Univ. Press, 1978.
11. **Ericsson KA.** Deliberate practice and acquisition of expert performance: a general overview. *Acad Emerg Med* 15: 988–994, 2008.
12. **Ericsson KA.** Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med* 79: S70–S81, 2004.
13. **Ericsson KA, Nandagopal K, Roring RW.** Toward a science of exceptional achievement: attaining superior performance through deliberate practice. *Ann NY Acad Sci* 1172: 199–217, 2009.
14. **Eva KW, Neville AJ, Norman GR.** Exploring the etiology of content specificity: factors influencing analogic transfer and problem solving. *Acad Med* 73: S1–5, 1998.
15. **Eva KW, Norman GR, Neville AJ, Wood TJ, Brooks LR.** Expert-novice differences in memory: a reformulation. *Teach Learn Med* 14: 257–263, 2002.
16. **Faulkner L.** Beyond the five-user assumption: benefits of increased sample sizes in usability testing. *Behav Res Methods Instrum Comp* 35: 379, 2003.
17. **Finstad K.** The system usability scale and non-native English speakers. *J Usabil Studies* 1: 185–188, 2006.
18. **Fishman S.** *Legal Guide to Web & Software Development*. Berkeley, CA: Nolo, 2007, p. 528.

19. **Friedman CP.** Anatomy of the clinical simulation. *Acad Med* 70: 205–209, 1995.
20. **Gawande AA, Thomas EJ, Zinner MJ, Brennan TA.** The incidence and nature of surgical adverse events in Colorado and Utah in 1992. *Surgery* 126: 66, 1999.
21. **Gordon MS, Ewy GA, DeLeon AC Jr, Waugh RA, Felner JM, Forker AD, Gessner IH, Mayer JW, Patterson D.** "Harvey," the cardiology patient simulator: pilot studies on teaching effectiveness. *Am J Cardiol* 45: 791–796, 1980.
22. **Groen GJ, Patel VL.** Medical problem-solving: some questionable assumptions. *Med Educ* 19: 95–100, 1985.
23. **Hornbæk K.** Current practice in measuring usability: challenges to usability studies and research. *Int J Hum Comp Studies* 64: 79–102, 2006.
24. **Howarth J, Smith-Jackson T, Hartson R.** Supporting novice usability practitioners with usability engineering tools. *Int J Hum Comp Studies* 67: 533–549, 2009.
25. International Organization for Standardization. *ISO 9241-11. Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)–Part 11: Guidance on Usability* (online). http://www.it.uu.se/edu/course/homepage/acsd/vt09/ISO9241part11.pdf [25 May 2011].
26. **Jones MG, Farquhar JD.** User interface design for web-based instruction. In: *Web-Based Instruction*, edited by Khan BH. Englewood Cliffs, NJ: Educational Technology, 1997, p. 239–244.
27. **Kohn LT, Corrigan J, Donaldson MS.** *To Err is Human: Building a Safer Health System*. Washington, DC: National Academy, 2000.
28. **Leape LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA, Hebert L, Newhouse JP, Weiler PC, Hiatt H.** The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. *N Engl J Med* 324: 377–384, 1991.
29. **Mayer RE.** Principles of multimedia learning based on social cues: personalization, voice and image principles. In: *The Cambridge Handbook of Multimedia Learning*, edited by Mayer RE. New York: Cambridge Univ. Press, 2005.
30. Millward-Brown. *Adobe Plug-In Technology Study, September 2010* (online). http://www.adobe.com/products/player_census/flashplayer/ [28 November 2010].
31. **Nielsen J.** Guerrilla HCI: using discount usability engineering to penetrate the intimidation barrier. In: *Cost-Justifying Usability*, edited by Bias RG, Mayhew DJ. Boston, MA: Academic, 1994.
32. **Nielsen J.** Iterative user-interface design. *Computer* 26: 41, 1993.
33. **Nielsen J.** *Usability engineering*. Boston, MA: Academic, 1993.
34. **Nielsen J.** *Usability Laboratories: a 1994 Survey* (online). http://www.useit.com/papers/uselabs.html [25 May 2011].
35. **Nielsen J, Landauer TK.** A mathematical model of the finding of usability problems. In: *Proceedings of the INTERCHI 1993 Conference on Human Factors in Computing Systems*. Amsterdam: IOS, 1993, p. 206–213.
36. **Nielsen J, Mack RL:** *Usability Inspection Methods*. New York: Wiley, 1994, p. 413.
37. **Norman G.** The essential role of basic science in medical education: the perspective from psychology. *Clin Invest Med* 23: 47–51, 2000.
38. **Norman GR, Trott AL, Brooks LR, Smith EKM.** Cognitive differences in clinical reasoning related to postgraduate training. *Teach Learn Med* 6: 114–120, 1994.
39. **Patel VL, Cranton PA.** Transfer of student learning in medical education. *J Med Educ* 58: 126–135, 1983.
40. **Patel VL, Glaser R, Arocha JF.** Cognition and expertise: acquisition of medical competence. *Clin Invest Med* 23: 256–260, 2000.
41. **Patel VL, Groen GJ, Frederiksen CH.** Differences between medical students and doctors in memory for clinical cases. *Med Educ* 20: 3–9, 1986.
42. **Reimann P, Schult TJ.** Turning examples into cases: acquiring knowledge structures for analogical problem solving. *Educ Psychol* 31: 123–132, 1996.
43. **Renkl A.** The worked-out examples principle in multimedia learning. In: *The Cambridge Handbook of Multimedia Learning*, edited by Mayer RE. Cambridge: Cambridge Univ. Press, 2005.
44. **Sajid AW, Ewy GA, Felner JM, Gessner I, Gordon MS, Mayer JW, Shub C, Waugh RA.** Cardiology patient simulator and computer-assisted instruction technologies in bedside teaching. *Med Educ* 24: 512–517, 1990.
45. **Sandars J.** The importance of usability testing to allow e-learning to reach its potential for medical education. *Educ Prim Care* 21: 6–8, 2010.
46. **Schreiber M, Halperin ML.** The Paleolithic curriculum: figure it out (with the help of experts). *Adv Physiol Educ* 20: 185–194, 1998.
47. **Snyder C.** *Paper Prototyping: the Fast and Easy Way to Design and Refine User Interfaces*. San Diego, CA: Morgan Kaufmann, 2003.
48. Standish Group. *CHAOS Summary 2009: the 10 Laws of CHAOS* (online). http://www.statelibrary.state.pa.us/portal/server.pt/document/690719/chaos_summary_2009_pdf. [01 November 2010].
49. **Sweller J.** Cognitive load theory, learning difficulty, and instructional design. *Learn Instruct* 4: 295–312, 1994.
50. **Sweller J, Chandler P.** Why some material is difficult to learn. *Cogn Instruct* 12: 185–233, 1994.
51. **Sweller J, Cooper GA.** The use of worked examples as a substitute for problem solving in learning algebra. *Cogn Instruct* 2: 59–89, 1985.
52. **Tobias S, Duffy TM.** *Constructivist Instruction: Success or Failure?* New York: Routledge, 2009.
53. **Tullis TS, Stetson JN.** *A Comparison of Questionnaires for Assessing Website Usability* (online). http://home.comcast.net/~tomtullis/publications/UPA2004TullisStetson.pdf [25 May 2011].
54. **Virzi R.** Refining the test phase of usability evaluation: how many subjects is enough? *Hum Factors* 34: 457–468, 1992.
55. **Wexler S, Schlenker Bruce B, Clothier P, Miller D, Nguyen F.** *Authoring and Development Tools*. Santa Rosa, CA: The e-Learning Guild, 2008; http://www.elearningguild.com/research/.
56. Wikipedia. "ActionScript." http://en.wikipedia.org/wiki/ActionScript [25 May 2011].
57. Wikipedia. "Website wireframe." http://en.wikipedia.org/wiki/Website_wireframe [25 May 2011].
58. **Wong G, Greenhalgh T, Pawson R.** Internet-based medical education: a realist review of what works, for whom and in what circumstances. *BMC Med Educ* 10: 12, 2010.
59. **Woods NN, Brooks LR, Norman GR.** The value of basic science in clinical diagnosis: creating coherence among signs and symptoms. *Med Educ* 39: 107–112, 2005.
60. **Woods NN, Neville AJ, Levinson AJ, Howey EH, Oczkowski WJ, Norman GR.** The value of basic science in clinical diagnosis. *Acad Med* 81: S124–S127, 2006.
61. **Zalunardo N, Lemaire M, Davids MR, Halperin ML.** Acidosis in a patient with cholera: a need to redefine concepts. *QJM* 97: 681–696, 2004.
62. **Ziv A, Wolpe PR, Small SD, Glick S.** Simulation-based medical education: an ethical imperative. *Acad Med* 78: 783–788, 2003.

# CHAPTER 4

## USABILITY EVALUATION BY USER TESTING

### Testing involving postgraduate trainees and qualified medical specialists and

### exploring the number of participants needed for usability evaluation by conducting a Monte Carlo simulation

Published paper:

Davids MR, Chikte U, Grimmer-Somers K, Halperin ML. Usability testing of a multimedia e-learning resource for electrolyte and acid-base disorders. British Journal of Educational Technology 2014;45(2):367–381.

Manuscript under review:

Davids MR, Harvey J, Halperin ML, Chikte UME. Determining the number of participants needed for the usability evaluation of e-learning resources: a Monte Carlo simulation.

*British Journal of Educational Technology* *Vol 45 No 2 2014* 367–381
doi:10.1111/bjet.12042

# *Usability testing of a multimedia e-learning resource for electrolyte and acid-base disorders*

**Mogamat Razeen Davids, Usuf Chikte, Karen Grimmer-Somers and Mitchell L Halperin**

*Razeen Davids is Associate Professor in the Division of Nephrology and Usuf Chikte is Head of the Department of Interdisciplinary Health Sciences at Stellenbosch University, Cape Town, South Africa. Karen Grimmer-Somers is Professor of Allied Health and Director of the International Centre for Allied Health Evidence, University of South Australia, and a visiting professor at Stellenbosch University. Mitch Halperin is Emeritus Professor in the Keenan Research Building, Li Ka Shing Knowledge Institute of St. Michael's Hospital and Division of Nephrology at the University of Toronto. Address for correspondence: Prof Mogamat Razeen Davids, Division of Nephrology, Stellenbosch University & Tygerberg Hospital, Cape Town 7505, South Africa. Email: mrd@sun.ac.za*

**Abstract**

The usability of computer interfaces may have a major influence on learning. Design approaches that optimize usability are commonplace in the software development industry but are seldom used in the development of e-learning resources, especially in medical education. We conducted a usability evaluation of a multimedia resource for teaching electrolyte and acid-base disorders by studying the interaction of 15 medical doctors with the application. Most of the usability problems occurred in an interactive treatment simulation, which was completed successfully by only 20% of participants. A total of 27 distinct usability problems were detected, with 15 categorized as serious. No differences were observed with respect to usability problems detected by junior doctors as compared with more experienced colleagues. Problems were related to user information and feedback, the visual layout, match with the real world, error prevention and management, and consistency and standards. The resource was therefore unusable for many participants; this is in contrast to good scores previously reported for subjective user satisfaction. The findings suggest that the development of e-learning materials should follow an iterative design-and-test process that includes routine usability evaluation. User testing should include the study of objective measures and not rely only on self-reported measures of satisfaction.

## Introduction

e-Learning is considered to be as effective as educational interventions delivered by traditional media (Chumley-Jones, Dobbie & Alford, 2002; Cook *et al*, 2008) and has rapidly become part of the medical education mainstream (Ellaway & Masters, 2008). Creative educators are increasingly using animation, simulations and virtual 3-D learning environments (Hansen, 2008) to create engaging learning resources for students and health-care professionals. Virtual patients, for instance, hold particular promise for assisting in the development of clinical reasoning ability (Cook & Triola, 2009).

Developing innovative e-learning materials can be expensive and time-consuming. A survey of virtual patient development at US and Canadian medical schools revealed that the cases took an average of 16.6 months to complete and that 85% of them cost over $10 000 (Huang, Reynolds & Candler, 2007). It is therefore important to maximize the educational impact of these

**Practitioner Notes**

What is already known about this topic

- The usability of computer interfaces may have a major influence on learning.
- While design approaches that optimize usability are common in the software development industry, this is not the case with e-learning, especially in the area of medical education.

What this paper adds

- Neglecting the evaluation of usability may lead to the implementation of e-learning materials with poor usability, with failure to achieve desired educational outcomes.
- The results of objective user testing do not correlate well with evaluations based on self-reported user satisfaction.

Implications for practice and/or policy

- e-Learning development should include routine usability evaluation and follow an iterative design-test-redesign approach.
- Usability evaluation should include observing typical end-users interacting with the system and not be based only on subjective ratings of user satisfaction.

resources. One aspect that has not been sufficiently emphasized in the implementation of effective e-learning is the usability of the technology interface. This has a major impact on learning and should be considered when designing e-learning resources (Sandars, 2010; Zaharias, 2009). Usability is a concept from the field of human–computer interaction that describes the ease with which a technology interface can be used. The International Standard, ISO 9241-11, defines it as the "*Extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use*" (Abran, Khelifi, Suryn & Seffah, 2003). A user interface should be so intuitive and self-evident that even inexperienced users can accomplish tasks successfully (Krug, 2006).

High usability of learning resources is essential, though of course not sufficient, to achieving the desired educational impact (Sandars & Lafferty, 2010). This is likely to be especially relevant when the subject matter is complex and contains multiple interacting elements (Sweller, 2010). Such material presents a heavy intrinsic cognitive load in view of the limited capacity of working memory and is often perceived as difficult to learn. Poorly designed user interfaces can present an additional, extraneous cognitive load, as the user has to struggle with challenging content as well as with the technology interface. Reducing extraneous cognitive load has been shown to lead to large gains in learning efficiency (Mayer & Moreno, 2003; van Merrienboer & Sweller, 2010); optimizing the usability of e-learning resources therefore seems essential.

Iterative methodologies that include the routine evaluation of usability are common in the software development industry (Bygstad, Ghinea & Brevik, 2008; Holzinger, Errath, Searle, Thurnher & Slany, 2005; Mao, Vredenburg, Smith & Carey, 2005; Sohaib & Khan, 2010). As far back as the mid-80s, Gould and Lewis (1985) recommended the following design principles: an early focus on users and their tasks; empirical user testing starting early in the development process; and an iterative approach using cycles of design, testing and redesign until the application meets performance and usability goals. This approach is seldom used in the development and evaluation of e-learning resources, especially in medical education (Sandars, 2010). There are two main categories of usability evaluation techniques: *empirical user testing* involves studying typical end-users interacting with the application while *usability inspection methods* involves experts evaluating the application against a set of rules or design principles (Dumas & Salzman, 2006).

Selecting which methods and measures to use when evaluating an e-learning resource remains difficult. For example, we can evaluate usability, learner interactions, learner perceptions or learning outcomes, we can collect subjective or objective data, qualitative or quantitative data, and we can make use of experts or we can involve typical end-users (Dyson & Campello, 2003). Even if the focus is on usability as in this study, different approaches are available and each will have their own resource requirements, examine a particular aspect of usability and detect different usability problems. A common recommendation is to combine methods whenever resources allow and to alternate between inspection by experts and end-user testing.

User testing usually involves participants being asked to think aloud as they interact with the system being tested. Evaluations may be conducted in settings ranging from sophisticated usability laboratories to informal settings employing paper prototypes (Snyder, 2003). User testing has been rated by usability professionals as having a greater impact on product development than inspection methods, although the latter is also very commonly used (Mao *et al*, 2005; Rosenbaum, Rohn & Humburg, 2000). Developers are less likely to question the validity of the results when usability problems are identified by real users rather than by experts (Dumas & Salzman, 2006). However, real users may be expensive and difficult to recruit and the recording, coding and analysis of testing sessions may also be expensive and time-consuming. Nielsen has popularized simpler methods, pointing out that any testing is better than not testing at all, and demonstrating that four to five users are sufficient for each cycle of testing (Nielsen, 2012). This "discount usability" approach (Nielsen, 2009) may be an efficient option for improving the process of developing e-learning materials.

Inspection methods are often less expensive because they involve fewer people and can detect many problems in a limited amount of time. Evaluators may also suggest solutions to the problems they find. The most commonly used technique is heuristic evaluation, in which expert evaluators find usability problems by examining an interface and judging its compliance with well-established usability principles, called heuristics. The process is influenced by the skills of the evaluators, with the ideal evaluators being "double experts" at usability and the domain of the application being evaluated (Nielsen, 1992). However, such individuals may be difficult to find or very expensive to employ. Evaluators may also have their own biases regarding interface design or may have insufficient domain knowledge, causing domain-specific problems to be missed. They may miss problems that affect real users or identify many low priority problems that hardly affect real users.

The raw data generated by an evaluation need to be transformed before it can be used to improve the user interface (Howarth, Andre & Hartson, 2007). Each occurrence of a usability problem encountered by a user or evaluator is a problem instance. All related instances must be recognized and consolidated into distinct problems, and the problems may then be categorized according to the interface elements involved, the severity of the problems or the design principles violated. See Table 1 for a set of widely used principles for guiding good interface design. Categorizing the problems in this way makes it easier to identify solutions to address them and also to prioritize them for fixing during the subsequent redesign process.

We have developed a web-based learning resource to help students and practicing clinicians acquire expertise in the complex area of electrolyte, water and acid-base disorders, an area of medicine that students and clinicians find particularly difficult to master (Dawson-Saunders, Feltovich, Coulson & Steward, 1990). Patients with these disorders are usually encountered by doctors working in the fields of internal medicine or pediatrics, or in subdisciplines of these fields such as nephrology, endocrinology and intensive care medicine. Our Electrolyte

*Table 1: Principles of good interface design (heuristics). The first 10 are those proposed by Nielsen (2005), and the last is from Karat* et al. *(1992)*

| *Heuristic* | *Descriptors* |
|---|---|
| 1. Visibility of system status; feedback | Keep users informed through timely appropriate feedback. They always know where they are, which actions can be taken and how they can be performed. |
| 2. Match with the real world—language, conventions | Speak the users' language, use familiar terms and concepts; follow real-world conventions. |
| 3. Consistency and conformity to standards | Words, situations and actions mean the same thing; application uses commonly accepted conventions and conforms to user expectations. |
| 4. Minimize memory load; recognition rather than recall | Objects, actions and options accessed easily. The user should not have to remember information from one part of the application to another. |
| 5. Aesthetic and minimalist design | No irrelevant information as it competes with relevant information and diminishes their relative visibility. Animation and transitions should be used sparingly. |
| 6. Help and documentation | It is better if the system can be used without documentation. If required it should be concise, easy to search and task-centered. |
| 7. User control and freedom | The user can control the direction and pace of the application. Clearly marked exits if they take wrong options by mistake. Support undo and redo. |
| 8. Flexibility and efficiency of use | Users can modify the application to suit their individual capabilities and needs, for example, by using shortcuts. |
| 9. Error prevention and tolerance | Careful design to prevent errors occurring. Despite user errors, the intended result may still be achieved by error correction or good error management. |
| 10. Help users recognize, diagnose and recover from errors | Error messages should be in plain language (no codes or jargon) and suggest a solution. |
| 11. Intuitive visual layout | Position elements on screen to be easily perceived and understandable, and visually attractive. |

Workshop provides instruction and the opportunity to practice the treatment of electrolyte disorders through an interactive simulation. The application is freely accessible at http://www.learnphysiology.org/sim1/.

The underlying teaching approach and the initial development of the Electrolyte Workshop have been described previously (Davids, Chikte & Halperin, 2011). The application was built in Flash® and involved several iterations of development and review by the authors and the development team. This informal review process by content experts and experienced developers detected and corrected many usability problems with the application. Self-reported end-user satisfaction with the completed application was good as judged by positive comments and high ratings on the System Usability Scale (Brooke, 1996).

This paper reports on an evaluation that focuses on objective measures of usability obtained by observing, recording and analyzing the interaction of end-users with the application. The study did not address educational outcomes. Testing was conducted with doctors working in the field of internal medicine, our main target audience. The purpose was to determine how well our Electrolyte Workshop conforms to principles of good interface design and to inform further development. The study illustrates the importance of user testing in evaluating e-learning materials and, in particular, demonstrates the need to observe users and examine objective data rather than to rely solely on more easily obtained questionnaire data.

## Methods

Ethics approval for the study was granted by the Committee for Human Research at the Faculty of Health Sciences of Stellenbosch University (project no. N08/05/158).

*The e-learning resource*

The Electrolyte Workshop is built in Adobe® Flash® and consists of case-based tutorials. There are two sections: cases in the WalkThru section present a clinical problem, then demonstrate how an expert would analyze the data and make decisions about treatment. Animation is used to illustrate changes in body fluid compartment sizes, brain cell size and plasma sodium concentrations. The concept is "look and learn," analogous to the use of worked-out examples in other disciplines (Renkl, 2005), which allows students to appreciate the underlying principles rather than being focused on finding solutions to the problem presented.

Cases in the second section, called the HandsOn section, are interactive and include a treatment simulation where users can select from a menu of therapies and receive immediate feedback via animations and text messages. The HandsOn cases have introductory ("lead-in") slides that set the scene for the treatment simulation. These slides contain important clinical and laboratory data that are needed to complete the treatment simulation. After successful completion of the simulation a summary slide is displayed containing several "take-home messages."

Currently the application contains only two cases, one in each section. The WalkThru case is that of a young girl with acute hyponatremia related to Ecstasy use, and the HandsOn case is that of chronic hyponatremia in a patient with Addison's disease.

*Participants*

User testing was conducted with 15 doctors at an academic department of medicine. The group included 10 doctors who were undertaking postgraduate training in internal medicine ("registrars") and 5 qualified specialists in internal medicine, nephrology and endocrinology. This group is typical of our target population. We considered that the specialists and registrars were likely to be different in terms of subject knowledge and experience, and therefore recruited 15 participants to allow us to include sufficient participants from both groups and also to improve the overall usability problem detection rate (Faulkner, 2003).

*User testing equipment and procedures*

The application was loaded onto two 15-inch laptop computers, each equipped with a mouse and a webcam with an integrated microphone.

To facilitate the capture and analysis of information from each testing session we installed a usability software tool on each computer. We selected Morae® (http://www.techsmith.com) for this purpose because it is widely used and suited our requirements in terms of data collection and analysis options, cost and ease of use. Running unobtrusively in the background, it records all user interactions with a website or computer application. This includes the user's voice, webcam video of facial expressions and video of all on-screen activity. It also captures data like mouse clicks and keyboard activity. Recordings are marked up to log the start and end of tasks, instances of usability problems, user comments and occasions when help was needed. Metrics like time, task completion rates, usability problem counts and mouse activity are readily generated.

Participants received written instructions. They were required to work through the WalkThru and HandsOn cases and look carefully at the different panels on each slide. They were encouraged to try different options in the treatment simulation and were also asked to look at the glossary. No time limits were set.

*Measures of usability*

For the purposes of evaluating usability the WalkThru case, the introductory slides of the HandsOn case, the treatment simulation of the HandsOn case, and the glossary were each regarded as a separate task.

Binary task completion rates and the detection of usability problems were recorded for each task as measures of effectiveness. Time on task and input device activity (mouse clicks and mouse movement) was recorded for each task as measures of efficiency.

Successful task completion in the WalkThru case and the introductory slides of the HandsOn case simply required that participants navigate through that section from beginning to end, viewing all the information available. For completion of the interactive treatment simulation in the HandsOn case participants had to treat their patient effectively by applying appropriate therapy at the correct dosages, and then exit the simulation to end with a summary "take home messages" slide. In the case of the glossary, participants were simply required to open it by clicking a text hyperlink on a slide or by using its navigation tab at the top of the screen.

The usability problems detected by participants as they worked through the tasks were categorized by severity, the interface element involved and the design principle (heuristic) violated. Our definition of a serious usability problem is based on that of Nielsen (1997), which takes into account the impact, frequency and persistence of the problem; it refers to a problem that may cause unacceptable delays or even task failure for the user and which needs to be fixed before an application is released. Table 1 lists the heuristics we considered when analyzing the usability problems detected. They are based on those proposed by Nielsen (2005) and as used by Karat, Campbell and Fiegel (1992). Each problem identified was mapped to one or more heuristic.

*Statistical tests*

Binary task completion rates are reported as proportions, usability problems as counts, and time on task (in minutes) and mouse activity (clicks and movement in pixels) as means ± SD. For the comparisons between specialists and registrars, and between those participants who completed a task successfully and those who did not, Fisher's exact test was used to compare proportions, and the Wilcoxon rank sum test was used to compare usability problem detection, time on task and mouse activity. The significance level was set at .05.

## Results

User testing focused on measures of effectiveness and efficiency and yielded data that are described below and in Table 2. Although not the focus of this study, we also compared specialists

*Table 2: Measures of effectiveness: successful task completion rates and counts of usability problems detected by participants. Where the same problem was encountered by multiple participants these instances were merged to provide a count of unique or distinct problems*

| | *Task completion* | *All problems* | | *Serious problems* | |
|---|---|---|---|---|---|
| | *Rate (%)* | *Problem instances* | *Distinct problems* | *Problem instances* | *Distinct problems* |
| Task 1: WalkThru case | 15/15 (100) | 4 | 4 | 1 | 1 |
| Task 2: HandsOn lead-in slides | 8/15 (53) | 16 | 5 | 10 | 2 |
| Task 3: HandsOn treatment simulation | 3/15 (20) | 44 | 18 | 34 | 12 |
| Total | | 64 | 27 | 43 | 15 |

with registrars, and participants who completed a task successfully with those who did not, and summarize these results at the end of this section.

*Measures of effectiveness: task completion rates and usability problem detection (Table 2)*

Task completion rates

Participants all completed the WalkThru case with ease. The lead-in section of the HandsOn case was completed successfully by eight participants (53% task completion rate) while the treatment simulation was completed successfully by only three participants (20%). The glossary was viewed by nine participants, none of whom experienced any usability problems while accessing this feature of the application. All of them opened the glossary by clicking its main navigation tab at the top of the screen and not via a text hyperlink on one of the slides.

Usability problem detection

A total of 27 distinct usability problems were identified, 15 of which were categorized as serious. A median of 4 problems were detected per participant, and in the case of the serious problems the median detection rate was 3 per participant. Table 3 contains a sampling of the serious usability problems detected, and lists the interface elements involved, the heuristics violated, as well as proposed solutions for addressing these problems.

In the WalkThru case four distinct usability problems were detected: these related to user information and feedback (two problems), user control and freedom (unclear navigation, one problem) and match with the real world (a problem with case accuracy, one problem). The only error categorized as serious was the last mentioned, which violated the heuristic of matching with the real world. An animation showed fluid moving out of the intracellular fluid compartment then simply disappearing and not appearing in the extracellular fluid compartment (see the first line of Table 3 for details and Multimedia Appendix S1 for a video clip).

In the lead-in section of HandsOn case a total of five distinct usability problems were identified (16 separate instances were recorded). They related to user information (one problem), the visual layout (two) and match with the real world (two). Two problems were categorized as serious: one was related to inadequate user information and the other to the heuristic of providing an intuitive visual layout. A sliding panel displaying important laboratory data opens on clicking its tab on the side of the screen (Figure 1). This sliding panel was completely missed by seven participants (47%). One of these participants worked through the case twice, and two others worked through it three times without discovering the panel (see line 2 of Table 3 for details and Multimedia Appendix S2 for a video clip).

In the treatment simulation of the HandsOn case a total of 18 distinct usability problems were identified (44 separate instances were recorded). These were related to user information and feedback (five problems), visual layout (three), match with the real world (one), user control and freedom (one), consistency and conformity to standards (two), error prevention and tolerance (five) and error management (one). Twelve of these 18 problems were graded as serious, based on their impact and the frequency of their occurrence.

The first serious usability problem identified in the treatment simulation related to the fidelity of the case and lack of clarity regarding the correct treatment (Table 3 line 3). Two participants, both experienced specialists, were not convinced of the need to apply any fluid therapy in this case of Addison's disease.

The most frequently encountered problem related to the heuristic of designing for error prevention and tolerance. There were repeated unsuccessful attempts by 10 participants (67%) to apply multiple treatments simultaneously (Table 3 line 4 and Multimedia Appendix S3). The simulation was designed to allow treatments to be applied sequentially, not simultaneously, so that feedback could be given after each step. Groups of treatment options are displayed in separate panels.

*Table 3: Selected examples of serious usability problems detected with the interface element and heuristic involved, and proposed measures to address them. The first example is from the WalkThru case and the others from the HandsOn case. The number of participants encountering a particular problem is included in column 1. Quotes from participants are in italics*

| *Examples of usability errors* | *Interface element* | *Heuristic involved* | *Solution* |
|---|---|---|---|
| *The sums don't add up; 1.8L moved out [of the intracellular fluid] but I don't see it in the extracellular fluid!* (1/15) | WalkThru case: Case accuracy | Match with the real world: language, conventions, case accuracy | Revise the animation to show the extracellular fluid compartment increasing in volume as 1.8L of water moves into it from the intracellular fluid compartment. |
| Participants do not notice the sliding lab data panel; this panel contains important information on blood and urine chemistry (9/15) | Lead-in slides: lab data tab | Intuitive visual layout | Redesign the interface to avoid using the sliding panel—group all related data and display in plain view in the left panel. |
| *In this patient with Addison's, why can't you start with only the mineralocorticoid then wait?* (2/15) | Treatment simulation: case accuracy | Match with the real world: language, conventions, case accuracy | As the simulation is designed to provide practice at prescribing accurate fluid therapy, amend the case data so that the need for fluid treatment is clear. |
| Participants try unsuccessfully to select and apply multiple treatments simultaneously; the application is designed to have one treatment given at a time, with feedback supplied after each step. (10/15) | Treatment simulation: treatment selection | Error prevention and tolerance | Remove all panel covers from the treatment option groups so that users clearly see that only one option can be selected and applied at a time. Reinforce this in the information provided just before the simulation is attempted. |
| Clicking the slider rail does not indicate dose despite the "thumb" moving; have to drag the thumb or double-click on the rail. (2/15) | Treatment simulation: slider control | Consistency and conformity to standards | Reprogram the slider so that the thumb (also) moves with a single click on the rail. |
| Participants who do not understand the use of the slider control are applying treatments with a dose of zero. (5/15) | Treatment simulation: slider control | Error prevention and tolerance | Program an error message to pop up if the dose applied is zero; explain that the slider must be dragged to indicate the dose. |
| Error message "Please select a radio option" is not clear to all participants. (1/15) | Treatment simulation: error messages | Help users recognize, diagnose and recover from errors | Avoid jargon or provide links to the glossary. |
| Messages displayed are sometimes vague or unhelpful. (3/15) | Treatment simulation: error messages | Help users recognize, diagnose and recover from errors | Review the algorithms underlying the error messages; ensure that all messages are relevant and useful. |
| Some participants seem to struggle to end the HandsOn case—the ending is not clear especially if the treatment applied in the simulation was not successful. (2/15) | Treatment simulation: navigation | User control and freedom | Offer all users access to the summary "take home messages" slide, even when they have not applied treatment successfully. |

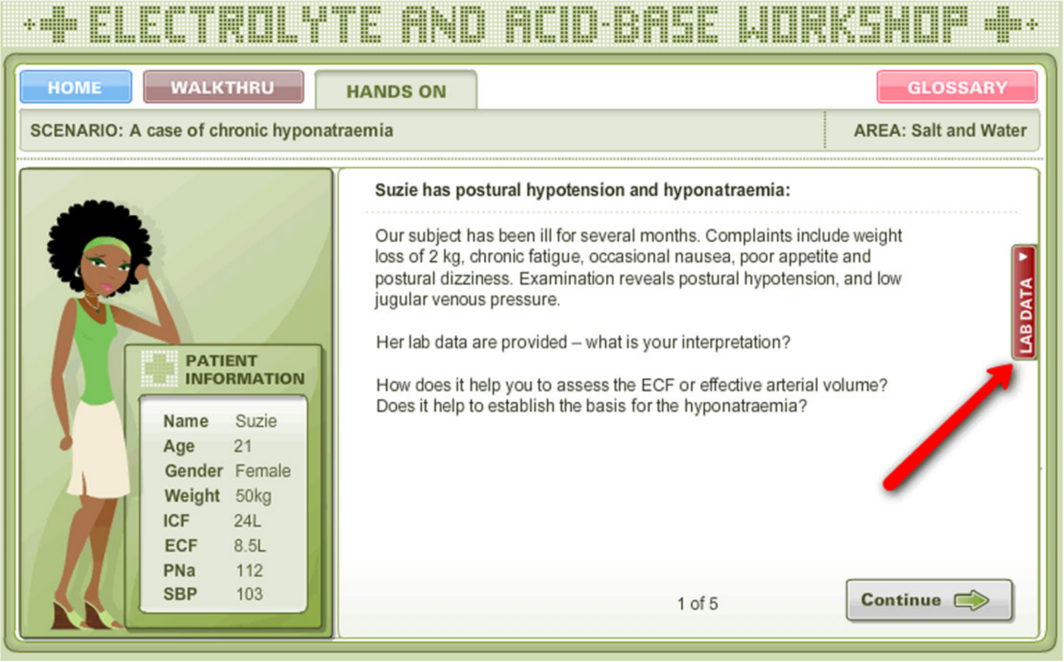



*Figure 1: The lab data panel slides open on clicking its tab at the side of the screen (arrow). This was missed by several participants*

Clicking on a panel cover causes it to slide open to reveal the options for that treatment group. When clicking on the panel for another treatment group, those options are revealed while the previous panel closes. This design led to much confusion and frustration. Most participants did not realize that a selected option was deselected once they clicked on another panel to try and add a second treatment.

A second serious problem, also related to error prevention and error management, was that some participants were unable to use the slider control (Table 3 lines 5 and 6). They would select a therapy but fail to indicate the dose by dragging the "thumb" along the rail of the slider control and would therefore apply a dose of zero (Figure 2 and Multimedia Appendix S4). As a result there was no change in plasma Na concentration or fluid compartment volumes. The impact of this usability problem was compounded by the display of poor feedback messages. For example, "*Your patient remains stable. Would you like to try something else?*" was vague and unhelpful, and contributed to participants' frustration. Additional usability problems related to the slider control are illustrated in Figure 3 and through video clips in Multimedia Appendices S5 and S6.

Problems with respect to user control and freedom were exposed when some participants appeared to have difficulty ending the simulation. The summary "take-home messages" slide was displayed only after successful completion of the simulation. After unsuccessful treatment attempts, participants were only offered the choice to try again, or to exit without any further feedback.

*Measures of efficiency: time on task and mouse activity*

Time on task

Participants spent a mean of 8.4 minutes on the WalkThru case, 6.8 minutes on the lead-in section of the HandsOn case and 9.9 minutes on the treatment simulation. The participants who accessed the glossary spent a mean of 1 minute on that part of the application.

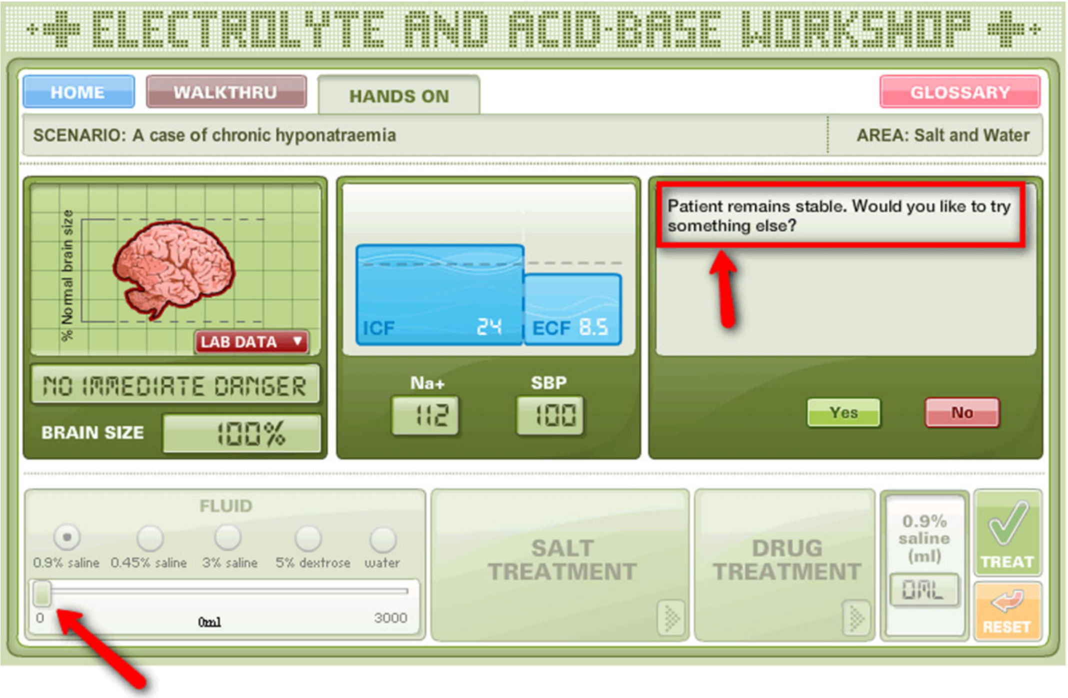



*Figure 2: The participant has clicked "Treat" without using the slider to indicate the dose of 0.9% saline, and there is therefore no change in any patient parameter. The feedback message is unhelpful*

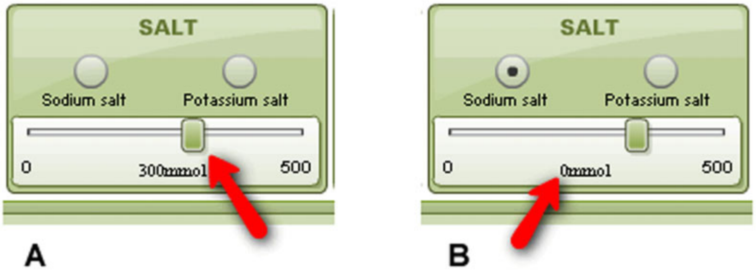



*Figure 3: A. The participant has indicated a dose of 300 mmol without first having to select a treatment option by clicking one of the radio buttons. B. Clicking on the slider rail causes the thumb to jump to the point clicked but the dose indicated is still 0 mmol*

Mouse clicks and mouse movement
As expected, the treatment simulation, being the most interactive section of the application, had the greatest mouse activity with a mean of 98.3 clicks and 38 884 pixels of mouse movement per participant.

*Specialists versus registrars*
There were no differences between the specialists and registrars with respect to task completion rates, or in the median number of total usability problems or serious usability problems detected.

The time spent on each task by specialists and registrars was similar. However, with regard to mouse activity for the treatment simulation, the specialists had a lower mean mouse click count

(47.4 vs. 123.8, $p = .010$) and less mouse movement (20 090 vs. 48 281 pixels, $p = .020$) than the registrars. Mouse activity for the other tasks was similar.

*Successful task completion versus failed task completion*
The two tasks that were not completed successfully by all participants were the lead-in section of the HandsOn case and the treatment simulation of the HandsOn case. Successful participants on the lead-in section had a higher mean mouse click count (20.4 vs. 11.7, $p = .042$) while those who completed the simulation successfully had a lower mouse click count (41.3 vs. 112.6, $p = .030$) and less mouse movement (14 859 vs. 44 890 pixels, $p = .030$). The successful participants on the lead-in section had a lower usability problem detection rate on this task ($p = .017$) while there was no difference on the treatment simulation.

## Discussion

Despite having followed an iterative design-and-review process involving the authors and developers, this evaluation with typical end-users detected several serious usability problems that had not been exposed during the initial development. Almost all were related to the interactive HandsOn tutorial and, in particular, the treatment simulation. Based on this evaluation, our e-learning application fell short with respect to principles of good interface design and would have been unusable for a large proportion of users, thus severely limiting the potential educational impact.

This finding is in contrast to the satisfactory self-reported user feedback previously obtained and confirms the observation that subjective measures of users' perceptions are often poorly correlated with objective measures (Bangor, Kortum & Miller, 2008). Our participants were aware who had developed the system and might well have been less critical in their responses because of this. When the aim is to improve the usability of a product, it is clear that it is not sufficient to employ only subjective measures of user satisfaction. The problems detected by employing user testing have allowed us to compile a detailed list of suggested revisions for the next iteration of the application.

Employing specialized usability software provided us with a rich source of data in the form of video recordings and usability metrics, giving us unique insights into participants' experiences. This allowed us to appreciate the full impact of the usability problems detected. For example, the levels of frustration—visible on participants' faces as they struggled with the slider control and repeatedly applied dosages of zero with no change in any patient parameters—may have been missed without the webcam video data stream. Another example was where the recordings provided accurate quantitative data that helped us to evaluate the utility of the glossary. Only 60% of participants accessed this section of the application—a mere 1 minute was spent there by those who did—and not one participant reached the glossary by clicking on a text hyperlink to access an explanation or definition as was intended. It would seem that participants only opened the glossary because this was required by the written instructions provided. It is probable that our participants were familiar with the terminology and concepts used and hence had little need to consult the glossary. This type of user support might be of more value to undergraduate students.

Registrars tended to spend more time on each task and had more mouse activity, although this was statistically significant only for mouse activity in the treatment simulation. This might reflect them finding the content more unfamiliar and challenging as opposed to their senior colleagues but may also reflect a greater inclination to explore the application. As expected, fewer mouse clicks were recorded by participants who missed the sliding data panel in the lead-in section of the HandsOn case. In the treatment simulation, participants who could not complete the task successfully had much more mouse activity as they made one failed attempt after another.

While there were few usability problems detected in the WalkThru case, the interactive HandsOn section was effectively unusable for the majority of our participants. This was true for both experienced clinicians and their junior colleagues. It was therefore not possible for these participants to achieve the intended objective of improving their skill and confidence in treating hyponatremia through practice in a simulated environment.

The design flaws causing the poor usability violated a number of heuristics. The principle of ensuring visibility of system status means that users should always know what was happening through clear information and appropriate feedback. Our feedback messages were often unhelpful or irrelevant. The heuristic of error prevention and management was not well implemented, as evidenced by the problems with the slider control and the repeated attempts at multiple treatment selection, which was compounded by the unhelpful error messages. The sliding lab data panel that was missed by many participants indicated that we did not succeed in providing an intuitive visual layout. While user control and freedom was reasonably well ensured by the clear navigation and the self-paced nature of the application, several users appeared to be unclear how to exit the HandsOn case, as it did not display the closing summary slide unless treatment had been successful.

The question of how many users are sufficient to evaluate a technology interface has long been debated in the usability literature. Five users will, on average, uncover 80% of usability problems (Turner, Lewis & Nielsen, 2006). This well-known "five users is enough" approach is appropriate when the probability of each user discovering a given problem is around 0.3, when applications are not too large and complex, when testing is done at an early stage of development and when several cycles of design-and-test are envisaged (Turner *et al*, 2006). When the application is larger and complex or when later versions are tested after the most obvious problems are already fixed then the probability of problem detection will fall and five users will not be enough (Spool & Schroeder, 2001). When the application is designed for more than one target group, then users from each subpopulation will need to be recruited and once again a greater number of users will be required.

Faulkner (2003) found that while five users detected a mean of 80% of problems present, wide confidence intervals implied that a particular set of five users detected as few as 55% of the problems. With 10 users, the lowest percentage of problems detected was 80%, and with 15 users it was 90%. Faulkner recommended testing the maximum number of users that resources allow to increase the confidence that the problems that need to be fixed will be found (Faulkner, 2003). We followed this recommendation. Our group of 15 participants enabled us to include both specialists and registrars, who differed in subject knowledge and clinical experience. We believed that this might impact on the detection of usability problems; however, we found that the usability problem detection rates were similar in the two groups and thus independent of differences in expertise.

Large increases in key metrics have been documented using an iterative approach to improve the usability of websites and software applications (Marcus, 2005). At least two cycles of usability testing should be undertaken, starting in the early stages of development and using simple prototypes or wireframes. Another cycle of testing should be undertaken with the fully functional "live" version of the product. Additional testing is advisable since new problems may be introduced when fixing the old ones. Ideally, this should continue until no new problems of significance are detected, but this iterative process will often be cut short by practical considerations. We believe that our Electrolyte Workshop requires at least one more cycle of revision and evaluation before we will have a robust and well-designed e-learning resource.

Several lessons were learned in the course of doing the study. End-users need to be involved much earlier, ideally before or at the stage where simple prototypes or wireframes are being built. Even experienced developers will not anticipate all the problems that a novice user may encounter, as was starkly demonstrated here. It is also clear that using only satisfaction ratings is insufficient, as these may correlate poorly with other, more objective, measures of usability.

Usability inspection methods, especially heuristic evaluation, may offer another efficient option in evaluating e-learning materials if an expert panel with the required experience and expertise can be assembled. We have learned that user testing can be resource intensive with suitable users difficult or expensive to recruit, and conducting, recording and analyzing testing sessions very time-consuming. It may therefore be most efficient to first use heuristic evaluation to find and fix the most obvious problems and then to undertake testing with a small number of end-users.

Our informal reviews during the initial development did not involve usability experts or the use of formal guidelines or checklists, and overlooked many serious problems. Inspection techniques as well as user testing can be used from the very early stages of the development process. If usability evaluation is only done at the end of the design cycle, changes to the interface are usually more costly and difficult to implement.

## Conclusions

Our usability evaluation, which was facilitated by specialized usability software, allowed us to identify many problems that were missed during the initial development process. These problems would otherwise have gone undetected and we would have released a resource with very limited potential educational impact. Our findings will inform a careful revision of the application and guide further content development. Future studies will examine the effect of optimizing usability on measures of learning as well as on users' motivation and engagement with the application.

The design of e-learning materials, modules and programs for medical education should include routine usability evaluation and follow an iterative design-and-test process. This is essential if we are to exploit the full potential of the electronic medium and maximize learning outcomes for all users in our target populations. User testing should be employed from the earliest phases of development and should include the study of objective measures obtained by observing the interaction of users with the system being tested, and not rely only on subjective measures of user satisfaction.

## Conflicts of interest

None.

## Acknowledgements

We thank Justin Harvey of the Stellenbosch University Centre for Statistical Consultation for assistance with data analysis and Martin Schreiber for a critical review of the manuscript.

## Funding

This work was supported by grants from the South African Universities Health Sciences IT Consortium and Stellenbosch University's Fund for Innovation and Research into Learning and Teaching.

## References

Abran, A., Khelifi, A., Suryn, W. & Seffah, A. (2003). Usability meanings and interpretations in ISO standards. *Software Quality Journal*, *11*, 325–338.

Bangor, A., Kortum, P. T. & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, *24*, 574–594.

Brooke, J. (1996). SUS: a "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester & I. L. McClelland (Eds), *Usability evaluation in industry* (pp. 189–194). London: Taylor & Francis.

Bygstad, B., Ghinea, G. & Brevik, E. (2008). Software development methods and usability: perspectives from a survey in the software industry in Norway. *Interacting with Computers*, *20*, 375–385.

Chumley-Jones, H. S., Dobbie, A. & Alford, C. L. (2002). Web-based learning: sound educational method or hype? A review of the evaluation literature. *Academic Medicine*, *77*, S86–S93.

Cook, D. A., Levinson, A. J., Garside, S., Dupras, D. M., Erwin, P. J. & Montori, V. M. (2008). Internet-based learning in the health professions: a meta-analysis. *Journal of the American Medical Association*, *300*, 1181–1196.

Cook, D. A. & Triola, M. M. (2009). Virtual patients: a critical literature review and proposed next steps. *Medical Education*, *43*, 303–311.

Davids, M. R., Chikte, U. M. E. & Halperin, M. L. (2011). Development and evaluation of a multimedia e-learning resource for electrolyte and acid-base disorders. *Advances in Physiology Education*, *35*, 295–306.

Dawson-Saunders, B., Feltovich, P. J., Coulson, R. L. & Steward, D. E. (1990). A survey of medical school teachers to identify basic biomedical concepts medical students should understand. *Academic Medicine*, *65*, 448–454.

Dumas, J. S. & Salzman, M. C. (2006). Usability assessment methods. *Reviews of Human Factors and Ergonomics*, *2*, 109–140.

Dyson, M. C. & Campello, S. B. (2003). Evaluating virtual learning environments: what are we measuring. *Electronic Journal of E-Learning*, *1*, 11–20.

Ellaway, R. & Masters, K. (2008). AMEE Guide 32: e-Learning in medical education Part 1: learning, teaching and assessment. *Medical Teacher*, *30*, 455–473.

Faulkner, L. (2003). Beyond the five-user assumption: benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers*, *35*, 379–383.

Gould, J. D. & Lewis, C. (1985). Designing for usability: key principles and what designers think. *Communications of the ACM*, *28*, 300–311.

Hansen, M. M. (2008). Versatile, immersive, creative and dynamic virtual 3-D healthcare learning environments: a review of the literature. *Journal of Medical Internet Research*, *10*, e26.

Holzinger, A., Errath, M., Searle, G., Thurnher, B. & Slany, W. (2005). From extreme programming and usability engineering to extreme usability in software engineering education (XP+ UE→ XU). Proceedings of the 29th Annual International Computer Software and Applications Conference (COMPSAC'05) (pp. 169–172). Washington, DC: IEEE Computer Society.

Howarth, J., Andre, T. S. & Hartson, R. (2007). A structured process for transforming usability data into usability information. *Journal of Usability Studies*, *3*, 7–23.

Huang, G., Reynolds, R. & Candler, C. (2007). Virtual patient simulation at U.S. and Canadian medical schools. *Academic Medicine*, *82*, 446–451.

Karat, C.-M., Campbell, R. & Fiegel, T. (1992). Comparison of empirical testing and walkthrough methods in user interface evaluation. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. (pp. 397–404). Monterey, California, USA, ACM.

Krug, S. (2006). *Don't make me think! A common sense approach to Web usability*. Berkeley, CA: New Riders.

Mao, J. Y., Vredenburg, K., Smith, P. W. & Carey, T. (2005). The state of user-centered design practice. *Communications of the ACM*, *48*, 105–109.

Marcus, A. (2005). User interface design's return on investment: examples and statistics. In R. G. Bias & D. J. Mayhew (Eds), *Cost-Justifying usability: an update for the internet age* (2nd ed.). (pp. 17–39). San Francisco, CA: Elsevier.

Mayer, R. E. & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*, 43–52.

Nielsen, J. (1992). Finding usability problems through heuristic evaluation. Proceedings of the 1992 SIGCHI conference on Human Factors in Computing Systems (pp. 373–380). Monterey, CA: ACM.

Nielsen, J. (1997). Severity ratings for usability problems. Retrieved 2011/09/07, from http://www.useit.com/papers/heuristic/severityrating.html

Nielsen, J. (2005). Ten usability heuristics. Retrieved 31/08/2012, from http://www.useit.com/papers/heuristic/heuristic_list.html

Nielsen, J. (2009). Discount usability: 20 years. Retrieved 28/08/2012, from http://www.useit.com/alertbox/discount-usability.html

Nielsen, J. (2012). How many test users in a usability study? Retrieved 28 August 2012, from http://www.useit.com/alertbox/number-of-test-users.html

Renkl, A. (2005). The worked-out examples principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 229–245). Cambridge, UK: Cambridge University Press.

Rosenbaum, S., Rohn, J. A. & Humburg, J. (2000). A toolkit for strategic usability: results from workshops, panels, and surveys. Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 337–344). The Hague, The Netherlands: ACM.

Sandars, J. (2010). The importance of usability testing to allow e-learning to reach its potential for medical education. *Education for Primary Care*, *21*, 6–8.
Sandars, J. & Lafferty, N. (2010). Twelve Tips on usability testing to develop effective e-learning in medical education. *Medical Teacher*, *32*, 956–960.
Snyder, C. (2003). *Paper prototyping: the fast and easy way to design and refine user interfaces*. San Diego, CA: Morgan Kaufmann.
Sohaib, O. & Khan, K. (2010). Integrating usability engineering and agile software development: a literature review. International Conference on Computer Design and Applications (pp. 32–38).
Spool, J. & Schroeder, W. (2001). Testing web sites: five users is nowhere near enough. Proceedings of the CHI 2001 conference on Human Factors in Computing Systems (pp. 285–286). Seattle, Washington: ACM.
Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, *22*, 123–138.
Turner, C. W., Lewis, J. R. & Nielsen, J. (2006). Determining usability test sample size. In W. Karwowski (Ed.), *International encyclopedia of ergonomics and human factors* (2nd ed.). (pp. 3084–3088). Boca Raton, FL: CRC/Taylor & Francis.
van Merrienboer, J. J. & Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Medical Education*, *44*, 85–93.
Zaharias, P. (2009). Usability in the context of e-learning. *International Journal of Technology and Human Interaction*, *5*, 37–59.

**Supporting information**

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Multimedia Appendix S1: This animation illustrates the movement of 1.8 L of fluid out of cells when hypertonic saline is used to treat acute hyponatremia. There is a problem with fidelity in that this fluid does not appear in the extracellular fluid compartment.

Multimedia Appendix S2: The lab data panel slides open on clicking its tab on the right side of the screen. This panel is easily missed and obscures on-screen text when open.

Multimedia Appendix S3: Participants tried unsuccessfully to select and apply multiple treatments simultaneously. Most participants did not realize that their first option was deselected once they clicked on another panel to try and add a second treatment.

Multimedia Appendix S4: Some participants failed to indicate the dose by dragging the "thumb" along the rail of the slider and therefore applied dosages of zero. The impact of this usability problem was compounded by the display of inappropriate feedback messages.

Multimedia Appendix S5: After a single click on the rail of the slider, the slider thumb jumps to the point clicked but the dose indicated is still 0 mmol. The dose is only registered when the thumb is dragged or the rail is double-clicked.

Multimedia Appendix S6: The slider is visible, and the participant is able to indicate a dose without first selecting the treatment to be applied.

# Determining the number of participants needed for the usability evaluation of e-learning resources: a Monte Carlo simulation

**Mogamat Razeen Davids, Justin Harvey, Mitchell L Halperin and Usuf ME Chikte**

***Running title:*** *Monte Carlo simulation*

*Razeen Davids is Associate Professor in the Division of Nephrology and Usuf Chikte is Professor and Head of the Department of Interdisciplinary Health Sciences at Stellenbosch University, Cape Town, South Africa.*

*Justin Harvey is a consultant at the Centre for Statistical Consultation at Stellenbosch University, South Africa.*

*Mitch Halperin is Emeritus Professor in the Division of Nephrology, Li Ka Shing Knowledge Institute, St Michael's Hospital and University of Toronto, Canada.*

*Address for correspondence: Prof MR Davids, Division of Nephrology, Department of Medicine, Stellenbosch University and Tygerberg Hospital, Cape Town 7505, South Africa. Email: mrd@sun.ac.za*

## Abstract

The usability of computer interfaces may have a major influence on learning and optimizing the usability of e-learning resources is therefore essential. However, this may be neglected because of time and monetary constraints. User testing is a common approach to usability evaluation and involves studying typical end-users interacting with the application being tested. Determining the minimum number of users that are required for such an evaluation is important as it has a direct bearing on the costs and time requirements. This issue has long been a subject of debate and the widely cited recommendation of five users being enough has been questioned. We conducted a usability evaluation of an e-learning resource for electrolyte and acid-base disorders by studying the interaction of medical doctors with the application. A total of 15 serious usability problems were detected, most of these related to an interactive treatment simulation. With this report we are making available the data on the detection of serious usability problems by each of our participants. We have used these data to run a Monte Carlo simulation and examine how many users would be sufficient to test our application. The MATLAB® code is supplied, as are our calculations of problem discovery rates. The e-learning application which was evaluated is freely available, together with a revised version with all identified usability problems addressed.

***Keywords:***
user testing, discount usability, usability problem detection rates

## The dataset

The first worksheet of this Excel file shows the detection of serious usability problems by each of our 15 study participants. A total of 15 distinct problems were identified. For each one we identify the relevant part of the application, link it to a particular heuristic, and provide a brief description. The participants included 5 medical specialists and 10 postgraduate trainees.

The second worksheet demonstrates our calculations of problem discovery rates and includes the calculation of a rate which is adjusted for small samples.

The application on which the dataset is based may be accessed at http://www.learnphysiology.org/sim1/ and the revised version, with the usability problems addressed, at http://www.learnphysiology.org/sim2/.

## Introduction

Optimizing the usability of e-learning resources is essential to achieving the desired educational impact (Sandars and Lafferty, 2010). Poorly designed user interfaces can present an extraneous cognitive load causing the user to struggle with the challenges of the technology interface as well as with the content to be learnt.

The two main approaches to evaluating usability are *empirical user testing* and *usability inspection.* User testing involves studying typical end-users interacting with the application being tested while usability inspection methods involve experts evaluating the application against a set of rules or design principles (Dumas and Salzman, 2006). User testing is often viewed as having greater validity than usability inspection and appears to have a greater impact on product development (Rosenbaum, Rohn and Humburg, 2000, Mao, Vredenburg, Smith and Carey, 2005, Dumas and Salzman, 2006). However, real users may be expensive and difficult to recruit and the recording and analysis of testing sessions may also be expensive and time-consuming.

The question of how many users are required is an important one which has long been debated in the usability literature. Increasing the number of users adds substantially to the cost and the time required for an evaluation. Nielsen (Nielsen, 2009) has popularized simpler, “discount usability” methods and suggested that four to five users are sufficient as they will, on average, uncover 80% of usability problems (Nielsen and Landauer, 1993, Turner, Lewis and Nielsen, 2006, Nielsen, 2012). Thereafter, the law of diminishing returns seems to apply, as fewer and fewer new problems are identified by involving additional users (Virzi, 1992, Turner, Lewis and Nielsen, 2006).

This “five users are enough” approach is widely cited, but needs to be put into context before it is taken as an efficient approach which can be applied to the improvement of our medical e-learning resources. It assumes that a formative evaluation is being conducted where several iterations of testing and redesign are envisaged, and that the probability of each user discovering a given problem is around .3. It may therefore be appropriate when testing at an early stage of the development process and when applications are not too large or complex (Turner, Lewis and Nielsen, 2006).

When the application is complex or when testing is done after the most obvious problems are already fixed, the probability of problem detection falls and five users may not be enough (Spool and Schroeder, 2001). Faulkner (Faulkner, 2003) found that five users detected a mean of 80% of the usability problems present, but wide confidence intervals implied that a particular set of five users might detect as few as 55% of the problems present.

We have reported on a usability evaluation of our multimedia e-learning resource for electrolyte and acid-base disorders (Davids, Chikte, Grimmer-Somers and Halperin, 2014). User testing was conducted by studying the interaction of postgraduate trainees and specialists in medicine with the application. In this report we provide data from that usability study, which contributes to the debate on how many users is enough.

The low problem detection rates we observed may reflect the complexity of the application or the fact that the most obvious problems had already been fixed in the initial development process. Had we involved only five participants, we would have detected a mean of just over half – but possibly as few as one-quarter – of the serious problems, depending on which five individuals had been recruited.

## Methods

Ethics approval was granted by the Committee for Human Research at the Faculty of Health Sciences of Stellenbosch University (project no. N08/05/158).

***The e-learning resource***

The Electrolyte Workshop is built in Adobe Flash® and consists of case-based tutorials. There are two sections: the WalkThru and the HandsOn section. Cases in the WalkThru section present a clinical problem and then demonstrate how an expert would analyse the data and make decisions about treatment. Animation is used to illustrate changes in body fluid compartment sizes, brain cell size and plasma sodium concentrations. The HandsOn section is interactive and includes a treatment simulation where users can select from a menu of therapies and receive immediate feedback via animations and text messages. The application is accessible at http://www.learnphysiology.org/sim1/.

***Participants***

User testing was conducted with 15 doctors at one academic department of medicine. They included 10 registrars in Internal Medicine and 5 specialists in Internal Medicine, Nephrology and Endocrinology.

***User testing equipment and procedures***

The application was loaded onto two 15-inch laptop computers, each equipped with a mouse and a webcam with an integrated microphone. We installed Morae® usability software (www.techsmith.com) to facilitate the capture and analysis of information from each testing session. Participants were required to navigate through the WalkThru case, viewing all the information available. For the HandsOn case, participants had to view the patient information available on the introductory slides, and then go on to the simulation where they would attempt to treat their patient effectively, finally exiting with a summary slide which provided key ‘take-home messages’.

***Detection of usability problems***

Usability problems detected were categorized by severity, the design principle (heuristic) violated and the interface element involved. The average problem detection rate by our participants was calculated by dividing the total number of problem occurrences by the number of participants, times the number of unique problems. An adjusted rate was also calculated as this is recommended to reduce the bias toward over-estimation which occurs with small sample sizes (Lewis, 2001, Lewis, 2006). This adjustment involves averaging a method based on Good-Turing discounting and a normalization method proposed by Hertzum and Jacobsen (Hertzum and Jacobsen, 2001).

Given our problem detection rates, the number of users required to detect 80% of the usability problems was estimated using the well-known equation below which is based on the binomial probability formula:

$$\text{Proportion of unique problems detected (80\%)} = 1\text{-}(1\text{-}P)^n,$$

where $P$ is the mean problem detection rate and $n$ is the number of participants (Turner, Lewis and Nielsen, 2006).

***Monte Carlo simulation***
A Monte Carlo simulation can be used to predict the performance of a process or system by simulating sampling from an actual population. It requires a mathematical model of the process being studied, an estimate of the variation of each input variable and an idea of what output performance would be acceptable. Random simulated data are then generated without the need to conduct large numbers of experiments or build large numbers of samples.

We conducted a Monte Carlo simulation using MATLAB v 7.1 to examine the effect of using fewer participants on problem detection rates. Sample sizes from 5 to 12 were studied. For each group size, 1 000 trials were run to select groups of that size randomly from the 15 participants and then derive the mean percentage of unique problems detected (as a percentage of the total found by all participants) and the minimum percentage of problems detected.

## Results in brief

A total of 27 usability problems were identified, 15 of which were categorized as serious. There were no significant differences between the specialists and the trainees with respect to the number of problems detected. The average problem detection rate as a proportion of total problems detected was .191 for the serious problems. After applying the adjustment recommended by Lewis the rate was .123. These problem detection rates predicted that we would need 8 participants to detect 80% of the serious problems based on the initial rates, and 13 participants based on the adjusted rates.

The Monte Carlo simulation (Table 1) revealed that 10 of our participants would have been sufficient to detect an average of 80% of the serious usability problems. Had we recruited only 5 of the 15 participants that were involved in the study, we would have detected an average of 58% of the problems, but potentially as few as 27%. With 12 participants we would have detected an average of 90% of the problems, but potentially as few as 67%.

## Acknowledgements and funding

This work was supported by grants from the South African Universities Health Sciences IT Consortium and Stellenbosch University's Fund for Innovation and Research into Learning and Teaching.

***Table 1.*** *Percentage of serious problems found in a model when sampling varying group sizes in a Monte Carlo simulation**

| No. of users | Min % Found | Mean % Found | 95% Conf. Interval | |
|---|---|---|---|---|
| 5 | 26.7% | 57.5% | 33.3% | 80.0% |
| 6 | 26.7% | 63.4% | 40.0% | 86.7% |
| 7 | 33.3% | 68.9% | 46.7% | 86.7% |
| 8 | 40.0% | 73.3% | 53.3% | 93.3% |
| 9 | 46.7% | 78.2% | 60.0% | 93.3% |
| 10 | 53.3% | 82.2% | 66.7% | 100.0% |
| 11 | 60.0% | 86.6% | 66.7% | 100.0% |
| 12 | 66.7% | 90.1% | 73.3% | 100.0% |

* 1000 random samples were taken for each group size

## References


Davids, M. R., Chikte, U., Grimmer-Somers, K. & Halperin, M. L. (2014). Usability testing of a multimedia e-learning resource for electrolyte and acid-base disorders. *British Journal of Educational Technology*, 45**,** 367-381.

Dumas, J. S. & Salzman, M. C. (2006). Usability assessment methods. *Rev Hum Factors Ergon*, 2**,** 109-140.

Faulkner, L. (2003). Beyond the five-user assumption: benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers*, **35,** 379-383.

Hertzum, M. & Jacobsen, N. E. (2001). The evaluator effect: a chilling fact about usability evaluation methods. *International Journal of Human-Computer Interaction*, 13**,** 421-443.

Lewis, J. R. (2001). Evaluation of procedures for adjusting problem-discovery rates estimated from small samples. *International Journal of Human-Computer Interaction*, 13**,** 445-479.

Lewis, J. R. (2006). Sample sizes for usability tests: mostly math, not magic. *Interactions*, 13**,** 29-33.

Mao, J. Y., Vredenburg, K., Smith, P. W. & Carey, T. (2005). The state of user-centered design practice. *Communications of the ACM*, 48**,** 105-109.

Nielsen, J. (2009). *Discount usability: 20 years* [WWW document]. Retrieved online 28/08/2012 at: http://www.useit.com/alertbox/discount-usability.html.

Nielsen, J. (2012). *How many test users in a usability study?* [WWW document]. Retrieved online 28 August 2012 at: http://www.useit.com/alertbox/number-of-test-users.html.

Nielsen, J. & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. *Proceedings of the INTERACT '93 and CHI '93 conference on Human Factors in Computing Systems* (pp. 206-213). Amsterdam, The Netherlands: ACM.

Rosenbaum, S., Rohn, J. A. & Humburg, J. (2000). A toolkit for strategic usability: results from workshops, panels, and surveys. *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 337-344). The Hague, The Netherlands: ACM.

Sandars, J. & Lafferty, N. (2010). Twelve Tips on usability testing to develop effective e-learning in medical education. *Medical Teacher*, 32**,** 956-960.

Spool, J. & Schroeder, W. (2001). Testing web sites: five users is nowhere near enough. *Proceedings of the CHI 2001 conference on Human Factors in Computing Systems* (pp. 285-286). Seattle, Washington: ACM.

Turner, C. W., Lewis, J. R. & Nielsen, J. (2006). Determining usability test sample size. In Karwowski, W. (Ed.), *International encyclopedia of ergonomics and human factors* (pp. 3084-3088). 2nd edn. Boca Raton, FL: CRC/Taylor & Francis.

Virzi, R. (1992). Refining the test phase of usability evaluation: how many subjects is enough? *Human Factors*, **34,** 457-468.

| | Sen. | UP1 | UP2 | UP3 | UP4 | UP5 | UP6 | UP7 | UP8 | UP9 | UP10 | UP11 | UP12 | UP13 | UP14 | UP15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | WT | IS | IS | TS | TS | TS | TS | TS | TS | TS | TS | TS | TS | TS | TS |
| **P1** | **1** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| **P2** | **1** | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **P3** | **1** | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **P4** | **1** | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| **P5** | **1** | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| **P6** | **2** | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| **P7** | **2** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **P8** | **2** | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **P9** | **2** | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **P10** | **2** | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 |
| **P11** | **2** | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| **P12** | **2** | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **P13** | **2** | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **P14** | **2** | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| **P15** | **2** | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**P: Participant no. 1-15**
**Sen.: seniority - 1=specialist 2=postgraduate trainee**
**UP: Usability problems 1-15.**
**WT: WalkThru case**
**IS: Introductory slides of the HandsOn case**
**TS: Treatment simulation of the HandsOn case**

**0 = problem not detected**
**1 = problem detected**

| | Interface element | Heuristic violated | Usability problem |
|---|---|---|---|
| UP1 | WT - slide 10 | Match with the real world | Fluid moves out of the intracellular fluid compartment but simply disappers and does not appear in the extracellular fluid compartment. |
| UP2 | IS - slides 1-4 | Intuitive visual layout | Sliding Lab Data panel missed; the tab at side of screen needs to be clicked. |
| UP3 | IS - instructions | User information and feedback | "*Do I answer somewhere?*" Users need information about whether input is required. |
| UP4 | TS - treatment panels | Intuitive visual layout | Users attempt to apply multiple treatments unsuccessfully. |
| UP5 | TS - lab data panel | Intuitive visual layout | Sliding Lab Data panel also missed in treatment simulation (after missing it in the introductory slides). |
| UP6 | TS - text messages | Intuitive visual layout | Long feedback message cut off. |
| UP7 | TS - slider | Conformity to standards | The slider does not work with a single click - it has to be dragged or double-clicked. |
| UP8 | TS - case accuracy | Match with the real world | Unclear to users (2 experienced clinicians) why fluid therapy is required in this case. |
| UP9 | TS - instructions | User information and feedback | "*How do I answer?*" Users need more information about how to use the treatment simulation. |
| UP10 | TS - text messages | User information and feedback | Unhelpful feedback messages. Problem with underlying algorithm. |
| UP11 | TS - instructions | User information and feedback | The time over which treatment is being administered is not clear. |
| UP12 | TS - treatment selection | User information and feedback | Users do not realize that the application "remembers" previous treatments applied. |
| UP13 | TS - feedback | User information and feedback | The correct treatment of the patient is not clearly communicated at the end. |
| UP14 | TS - slider | Error prevention and management | The slider allows doses of zero to be given. |
| UP15 | TS - navigation | User control and freedom | Ending the application is difficult, especially if the correct treatment is not given. Summary messages only shown after successful treatment. |

| | Sen. | UP1 | UP2 | UP3 | UP4 | UP5 | UP6 | UP7 | UP8 | UP9 | UP10 | UP11 | UP12 | UP13 | UP14 | UP15 | Count | P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **P1** | **1** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 3 | 0.200 |
| **P2** | **1** | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 0.133 |
| **P3** | **1** | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 3 | 0.200 |
| **P4** | **1** | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 4 | 0.267 |
| **P5** | **1** | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 3 | 0.200 |
| **P6** | **2** | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 5 | 0.333 |
| **P7** | **2** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.000 |
| **P8** | **2** | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0.067 |
| **P9** | **2** | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0.133 |
| **P10** | **2** | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 7 | 0.467 |
| **P11** | **2** | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 3 | 0.200 |
| **P12** | **2** | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0.067 |
| **P13** | **2** | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0.067 |
| **P14** | **2** | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 5 | 0.333 |
| **P15** | **2** | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0.200 |
| **Count** | | 1 | 7 | 1 | 10 | 5 | 1 | 2 | 1 | 1 | 3 | 1 | 1 | 2 | 5 | 2 | 43 | |
| **P** | | 0.067 | 0.467 | 0.067 | 0.667 | 0.333 | 0.067 | 0.133 | 0.067 | 0.067 | 0.2 | 0.067 | 0.067 | 0.133 | 0.333 | 0.133 | | **0.191** |
| | | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | | |

**P: Participant no. 1-15**
**Sen.: seniority - 1=specialist 2=postgraduate trainee**
**UP: Usability problems 1-15.**
**0 = problem not detected**
**1 = problem detected**

### Adjustment according to Lewis

| | |
|---|---|
| n= | 15 |
| **p_est =** | **0.191** |
| E(N1) = | 7 |
| N = | 15 |
| **p_adj =** | **0.123** |

### Users required if p=.123

| Goal discovery rate | User no. |
|---|---|
| 80% | 13 |
| 85% | 15 |
| 90% | 18 |
| 95% | 23 |

### Users required if p=.191

| Goal discovery rate | User no. |
|---|---|
| 80% | 8 |
| 85% | 9 |
| 90% | 11 |
| 95% | 15 |

```
% MATLAB CODE

% Load data for serious usability problems (errser)

ser = ...

[1      0       0       0       0       0       0       1       0       0       0       0       0       1       0; ...
0       1       0       0       0       0       0       0       0       0       0       0       0       0       1; ...
0       1       0       1       0       0       0       0       0       0       0       0       0       0       1; ...
0       1       0       1       1       0       0       0       0       0       0       0       0       1       0; ...
0       0       0       1       0       0       0       0       0       0       1       0       1       0       0; ...
0       1       0       1       1       0       0       0       0       0       0       1       0       1       0; ...
0       0       0       0       0       0       0       0       0       0       0       0       0       0       0; ...
0       0       0       0       0       1       0       0       0       0       0       0       0       0       0; ...
0       0       0       1       0       0       1       0       0       0       0       0       0       0       0; ...
0       1       1       1       1       0       0       0       1       1       0       0       0       1       0; ...
0       1       0       0       1       0       0       0       0       1       0       0       0       0       0; ...
0       0       0       1       0       0       0       0       0       0       0       0       0       0       0; ...
0       0       0       1       0       0       0       0       0       0       0       0       0       0       0; ...
0       0       0       1       0       0       1       0       0       1       0       0       1       1       0; ...
0       1       0       1       1       0       0       0       0       0       0       0       0       0       0];


% Full set parameters

[nser errser] = size(ser);

% set number of iterations to 1000, and set group sizes for sampling from 5 through to 12

num = 1000;

DATA = zeros(8,13);
```

```
for size = 5:12

    ERROR = zeros(num,1);
    for sample = 1:num

        %Select users at random for each data set
        x = randperm(15);
        sertmp = sortrows([x' ser]);
        sertmp1 = sertmp(1:size,2:errser+1);

            %Calculate the required stats
            serr = sum(sertmp1);
            nserr = 0;
            end
            for i = 1:length(serr)
                if serr(i) > 0
                    nserr = nserr + 1;
                end
            end

            ERROR(sample,1) = nserr/15;
    end
```

```
minerr = min(ERROR);
meanerr = mean(ERROR);
stderr = std(ERROR);
steerr = (1/sqrt(num))*stderr;
conf = sort(ERROR);
confl = conf(num*0.025,:);
confu = conf(num*0.975,:);
DATA(size-4,:) = [size minerr meanerr stderr steerr confl confu];
end
        %Calculate the stats with adjustment according to Lewis
        pe_ser = sum(serr)/(n*Nser);
        padj_ser = 0.5*((pe_ser - 1/n)*(1 - 1/n) + (pe_ser/(1 + ENser/Nser)));
        pdr_ser = 1 - (1 - padj_ser)^size;
        Goal = 0.9;
        reqn_ser = (log(1 - Goal))/(log(1-padj_ser));
        PROP(sample,:) = [pe_ser padj_ser pdr_ser reqn_ser];
end

PROPs = sort(PROP);
DATAmeans(size-4,:) = [size mean(PROP)];
DATAminimum(size-4,:) = [size min(PROP)];
DATAconfl(size-4,:) = [size PROPs(0.025*num,:)];
DATAconfu(size-4,:) = [size PROPs(0.975*num,:)];
end
```

# CHAPTER 5

# USABILITY EVALUATION BY HEURISTIC EVALUATION

Published paper:

Davids MR, Chikte UME, Halperin ML. An efficient approach to improve the usability of e-learning resources: the role of heuristic evaluation. Advances in Physiology Education 2013;37:242-8.

*Adv Physiol Educ* 37: 242–248, 2013;
doi:10.1152/advan.00043.2013.

# An efficient approach to improve the usability of e-learning resources: the role of heuristic evaluation

**Mogamat Razeen Davids,[1] Usuf M. E. Chikte,[2] and Mitchell L. Halperin[3]**

[1]*Division of Nephrology and Department of Medicine, Stellenbosch University and Tygerberg Hospital, Cape Town, South Africa;* [2]*Department of Interdisciplinary Health Sciences, Stellenbosch University, Cape Town, South Africa; and* [3]*Keenan Research Building, Li Ka Shing Knowledge Institute of St Michael's Hospital and Division of Nephrology, University of Toronto, Toronto, Ontario, Canada*

Submitted 6 May 2013; accepted in final form 8 July 2013

**Davids MR, Chikte UME, Halperin ML.** An efficient approach to improve the usability of e-learning resources: the role of heuristic evaluation. *Adv Physiol Educ* 37: 242–248, 2013; doi:10.1152/advan.00043.2013.—Optimizing the usability of e-learning materials is necessary to maximize their potential educational impact, but this is often neglected when time and other resources are limited, leading to the release of materials that cannot deliver the desired learning outcomes. As clinician-teachers in a resource-constrained environment, we investigated whether heuristic evaluation of our multimedia e-learning resource by a panel of experts would be an effective and efficient alternative to testing with end users. We engaged six inspectors, whose expertise included usability, e-learning, instructional design, medical informatics, and the content area of nephrology. They applied a set of commonly used heuristics to identify usability problems, assigning severity scores to each problem. The identification of serious problems was compared with problems previously found by user testing. The panel completed their evaluations within 1 wk and identified a total of 22 distinct usability problems, 11 of which were considered serious. The problems violated the heuristics of visibility of system status, user control and freedom, match with the real world, intuitive visual layout, consistency and conformity to standards, aesthetic and minimalist design, error prevention and tolerance, and help and documentation. Compared with user testing, heuristic evaluation found most, but not all, of the serious problems. Combining heuristic evaluation and user testing, with each involving a small number of participants, may be an effective and efficient way of improving the usability of e-learning materials. Heuristic evaluation should ideally be used first to identify the most obvious problems and, once these are fixed, should be followed by testing with typical end users.

simulation; iterative design; user-centered design; interface design

THE DEVELOPMENT of engaging e-learning materials for students and professionals in the health sciences is often resource intensive. It therefore becomes critical to evaluate and optimize these materials to maximize their educational impact. The usability of user interfaces is an important element that needs to be considered when designing e-learning resources. This is an underappreciated factor that, if ignored, may have a major impact on learning (30, 34). Usability describes the ease with which a technology interface can be used and has been defined as the "Extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (1). A poorly designed user interface imposes an additional, extraneous, cognitive load and impedes learning as users struggle with the interface as well as with the challenges of the content presented.

More recently, this traditional view of usability is being extended and affective dimensions such as aesthetics, fun, and flow are receiving increased attention as designers seek to enhance user motivation and ensure pleasurable user experiences (11, 16, 34). For example, a study by Miller (19) reported that students working in an online environment with enhanced aesthetic design had reduced cognitive load, increased motivation, and increased performance compared with those working with a low-aesthetic interface. It also seems that users' perception of the aesthetics of an interface may be negatively affected by poor usability (32).

Design approaches that routinely include usability evaluation are well established in the software development industry (3, 10, 13, 18, 21, 24, 31), but this is seldom the case in the development of e-learning resources, especially in the area of medical education (30). The aim of usability evaluation is to improve a system or application by identifying usability problems and then prioritizing fixing them based on their impact. A usability problem can be defined as any aspect of a design that, if changed, could result in an improvement in usability. There may be several iterations of design, testing, and redesign before an application is released.

Usability can be evaluated by empirical user testing, where typical end users are observed using an application in laboratory or field settings. Think-aloud protocols, largely based on the work of Ericsson and Simon (9), are often used. Users are encouraged to speak their thoughts aloud while working with the application or immediately afterward (8). This increases the number of problems identified compared with simply observing users. Formal modeling is an approach that can be applied early in the development cycle and aims to determine and model the knowledge a user needs and/or the actions a user should perform to accomplish specified tasks. By considering users' mental models, designers attempt to predict and therefore prevent potential usability problems (6).

Another approach is to use usability inspection by experts. This approach, which is the focus of this report, relies on the considered judgment of expert inspectors and includes methods such as heuristic evaluation, cognitive walkthroughs, guideline review, and consistency inspection (26). Heuristic evaluation and cognitive walkthroughs are two commonly used methods. Heuristic evaluation involves experts evaluating an interface against a set of generally accepted principles for good design (the heuristics) (25), whereas cognitive walkthrough is based

Address for reprint requests and other correspondence: M. R. Davids, Div. of Nephrology and Dept. of Medicine, Stellenbosch Univ. and Tygerberg Hospital, PO Box 19063, Tygerberg 7505, Cape Town, South Africa (e-mail: mrd@sun.ac.za).

on a theory of learning by exploration and involves a group of inspectors walking through the interface and doing a step-by-step analysis of a hypothetical user's potential actions and mental processes while performing particular tasks (17, 33).

Each approach has its own cost and time requirements and examines a particular aspect of usability. With user testing, end users may be expensive or difficult to recruit, and the recording and analysis of testing sessions may be expensive and time consuming. Cost and time pressures are common in many environments and may lead to the evaluation of new resources being neglected. Inspection methods offer appealing options in such resource-constrained situations, since skilled experts could evaluate the application quickly, without the need to involve end users.

It should be noted that the average problem detection rate of individual inspectors is generally low (22), and, therefore, using small groups of inspectors is recommended. A review of 11 usability studies (12) found that inspectors evaluating an interface detected different sets of problems, with the average agreement between any 2 inspectors ranging from 5% to 65%. This "evaluator effect" appears to exist for both novice and experienced inspectors and for both the detection of usability problems as well as the assessment of problem severity. The authors of this review also recommend that this unavoidable effect be dealt with by involving multiple inspectors.

Heuristic evaluation is the most commonly used of the inspection methods. Each inspector evaluates the application independently, usually working through it at least twice. On the first pass, the overall flow of the application is evaluated, and on the second pass, each interface element is examined in detail. Inspectors may be asked to categorize the problems found with respect to their severity and the heuristic(s) violated, and they may also suggest solutions to the problems identified. Compared with other inspection methods, heuristic evaluation appears to be a better predictor of problems that are encountered by end users and also identifies more severe usability problems (14, 22). The ideal inspectors would be "double experts" at usability and the domain of the application being evaluated (22), but such individuals are likely to be hard to find and may be expensive to employ.

While heuristic evaluation is often the most common approach used by practitioners in the field of human-computer interaction, its impact on influencing software design is often rated by these usability professionals as being well below that of tests conducted with real users (18, 29). Software developers and project managers appear less willing to make design changes based on expert reviews, which they believe may include many "false alarms" that may not necessarily affect real users, than when end users have been observed first hand encountering problems with the interface (8). The comparative usability evaluation study of Molich and Dumas (20), however, found no significant differences between the results of usability testing and expert reviews. They consider reviews by expert practitioners comparable to usability testing and point out that usability testing should not be seen as a "gold standard" as it overlooks usability problems like any other method.

We (5) have developed a Web-based multimedia application to help medical students and practicing colleagues acquire expertise in the diagnosis and treatment of electrolyte and acid-base disorders. This e-learning resource is available at http://www.learnphysiology.org/sim1/. It provides instruction and hands-on practice via an interactive treatment simulation. We (4, 5) previously described the development of our "Electrolyte Workshop" and the results of user testing with 15 residents and fellows in internal medicine and its subdisciplines. Briefly, the usability software tool Morae was used to facilitate the recording and analysis of the interaction of participants with the application. Measures of effectiveness (task completion rates and usability problem counts) and measures of efficiency (time on task and mouse activity) were studied. This evaluation revealed several serious problems that rendered the application unusable for a large proportion of study

Table 1. *Principles of good interface design*

| Heuristic | Descriptor |
|---|---|
| *1.* Visibility of system status; feedback | Keep users informed through timely, appropriate feedback. They should always know where they are, which actions can be taken, and how they can be performed. |
| *2.* Match with the real world: language and conventions | Speak the users' language, with familiar words, phrases, and concepts. Follow real-world conventions, making information appear in a natural and logical order. |
| *3.* Consistency and conformity to standards | Words, situations, and actions mean the same thing; application uses commonly accepted platform conventions and conforms to user expectations. |
| *4.* Minimize memory load; recognition rather than recall | Make objects, actions, and options visible. The user should not have to remember information from one part of the application to another. Instructions should be visible or easily retrievable. |
| *5.* Aesthetic and minimalist design | No irrelevant information as it competes with relevant information and diminishes their relative visibility. Animation and transitions should be used sparingly. |
| *6.* Help and documentation | It is better if the system can be used without documentation. If required it should be concise, easy to search, and task centered. |
| *7.* User control and freedom | The user can control the direction and pace of the application. There should be clearly marked exits available if they take wrong options by mistake. Support undo and redo. |
| *8.* Flexibility and efficiency of use | Users can modify the application to suit their individual capabilities and needs, for example, by using shortcuts. |
| *9.* Error prevention and tolerance | Careful design to prevent errors occurring. Despite user errors, the intended result may still be achieved by error correction or good error management. |
| *10.* Help users recognize, diagnose, and recover from errors | Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution. |
| *11.* Intuitive visual layout | Position elements on screen to be easily perceived, understandable, and visually attractive. |

The heuristics used by our expert panel to evaluate the application are those of Nielsen (25), with the last item from Karat et al. (15).

participants. An interactive treatment simulation, for example, was successfully completed by only 20% of participants.

While the evaluation with typical end users was extremely valuable, it was, however, very resource intensive, especially in regard to recruiting suitable participants and in terms of the time required to log and analyze the recordings of the testing sessions. The study was eventually completed over the course of several months. We therefore followed up this study by exploring whether usability inspection by experts might provide an equally effective but more efficient alternative.

This report details the heuristic evaluation of our Electrolyte Workshop conducted by a panel of experts. The findings were also compared with those previously obtained by user testing to try and identify the most efficient method for improving our e-learning resources.

## METHODS

Ethics approval for the project was granted by the Committee for Human Research of the Faculty of Medicine and Health Sciences of Stellenbosch University (project no. N08/05/158).

*The e-learning resource: our Electrolyte Workshop.* This Adobe Flash application (http://www.learnphysiology.org/sim1/) consists of case-based tutorials and can be accessed over the internet by any Web browser. Each case consists of a series of slides, with the navigation and therefore the pace of the tutorial controlled by the user. There are two main sections to the Electrolyte Workshop: the first uses a "look and learn" approach and is called the WalkThru section. A clinical problem is presented, followed by a demonstration of how an expert would analyze the data and embark on treatment. Animation is used to illustrate changes in body fluid compartment sizes, brain cell size, and plasma $Na^{+}$ concentrations. The "look and learn" concept is analogous to the use of worked-out examples in disciplines such as

Table 2. *Examples of usability problems detected by heuristic evaluation*

| Heuristic(s) Involved | Interface Element | Usability Problem | Solution |
|---|---|---|---|
| Visibility of system status | Loading of application | *Loads in 18 s with ADSL connection. [suggests adding a progress indicator] | Add indicator to indicate progress with loading application. |
| Visibility of system status | Treatment simulation: treatment selection | *Wasn't sure where to click first to start treatment. No indication of rate of [fluid] administration. How is salt/K administered? | Clear instructions needed on the slides preceding the simulation. |
| Visibility of system status. Help users recognize, diagnose and recover from errors | Treatment simulation: text messages | *Messages not very helpful, e.g., "This is not a useful option. . ."–need to explain why. "Please select radio option" is not easily understandable. | Review the algorithms underlying the text messages; ensure that all messages are relevant and useful. |
| Visibility of system status. Minimize memory load; recognition rather than recall | Treatment simulation: feedback | Show the user his/her treatment attempts, with feedback. Tell me if I am on wrong track and nudge [me] where to go. | Display treatments previously applied, with feedback. |
| User control and freedom | Treatment simulation: navigation | *Couldn't go back to "lead-in" slides from the treatment simulation. | Add "back" button. |
| User control and freedom | Treatment simulation: navigation | *All users need [to see summary] "take-home messages" when they finish. [not only those completing the simulation successfully] | Display "take-home messages" slide for all users. |
| User control and freedom | WalkThru case: animations | *Add "Replay animation" button on relevant slides. | Add function to replay animation without navigating away from the slide. |
| Match with the real world | Character used as "the patient" | *Using Suzie again as the patient for the HandsOn case is confusing. [different illness, same character] | Use a different character for each case. |
| Match with the real world | WalkThru case: the patient | Suzie still upright after having a seizure! [patient looks too well] | Modify illustration appropriately. |
| Match with the real world | WalkThru case: patient data panel | Update patient data on panel after successful treatment. [clinical and lab data should change] | Update details on panel as treatment is applied. |
| Intuitive visual layout | HandsOn case: lab data sliding panel | *Sliding panel easily missed. *Obscures the text on the slide when open. | Redesign to avoid using the sliding panel; display data in plain view in left panel. |
| Consistency and conformity to standards | Text on slides | *Difficult to read small text. [not familiar with increasing the font size in a browser] | Use bigger font size and/or inform users how to zoom in. |
| Aesthetic and minimalist design | Text and animation on slides | Trim words on slides. Why all the animated lines around Suzie? | Reduce word count and extraneous animation where possible. |
| Error prevention and tolerance | Treatment simulation: treatment selection | *Can't select [and apply] multiple treatments [simultaneously]. Prevent users from trying to select multiple treatments. | Remove covers from the treatment option panels so users see that only one option can be selected at a time. |
| Help and documentation | Glossary | No hyperlinks to glossary for terms in the HandsOn section. | Add hyperlinks for terms that may be unfamiliar. |

Quotes from the inspectors are shown in the "Usability Problem" column, with the authors' comments or interpretations in brackets. *Serious problem.

mathematics and physics and allows students to appreciate underlying principles rather than being focused on finding solutions to the problem presented (28).

The second section, called the HandsOn section, is more interactive. Each case includes a simulation that provides the opportunity for deliberate practice of the treatment of patients with electrolyte disorders and, in particular, the accurate prescription of intravenous fluid therapy. HandsOn cases begin with a series of "lead-in" slides containing the clinical and laboratory data, which set the scene for the treatment simulation. Within the simulation, users select from a menu of therapies and receive immediate feedback on the treatment applied via on-screen text messages and animations. Upon completion of the simulation, a final summary slide displays several "take-home messages."

At present, there is one case in each section. The WalkThru case is that of a young woman with acute hyponatremia related to the use of the drug Ecstasy, and the HandsOn case is that of chronic hyponatremia due to Addison's disease. There is also a glossary that can be accessed via text hyperlinks on the slides or via a tab in the main navigation menu.

*Heuristic evaluation procedures.* In this study, a panel of six experts conducted a heuristic evaluation. The panel consisted of a usability expert, two e-learning experts with expertise in instructional design, an internist with an additional qualification in medical informatics, and two experienced nephrologists as the subject matter experts. Inspectors were supplied with a website link to the application and worked independently. Written instructions included information about the purpose of the application and stated that their participation was aimed at improving the application and formed part of a research project. They were required to work through the different sections of the application and evaluate it according to a set of commonly used heuristics (Table 1) based on those of Nielsen (25) and as used by Karat et al. (15). A template for recording and grading the usability problems detected was provided. Inspectors were asked to indicate the heuristic(s) relevant to each problem and to assign severity scores based on its frequency, persistence, and impact. The severity rating scale of Nielsen (23) was used as follows: *1* = cosmetic problem only, need not necessarily be fixed; *2* = minor usability problem, fixing this should be a low priority; *3* = major usability problem, fixing this should be a high priority; and *4* =

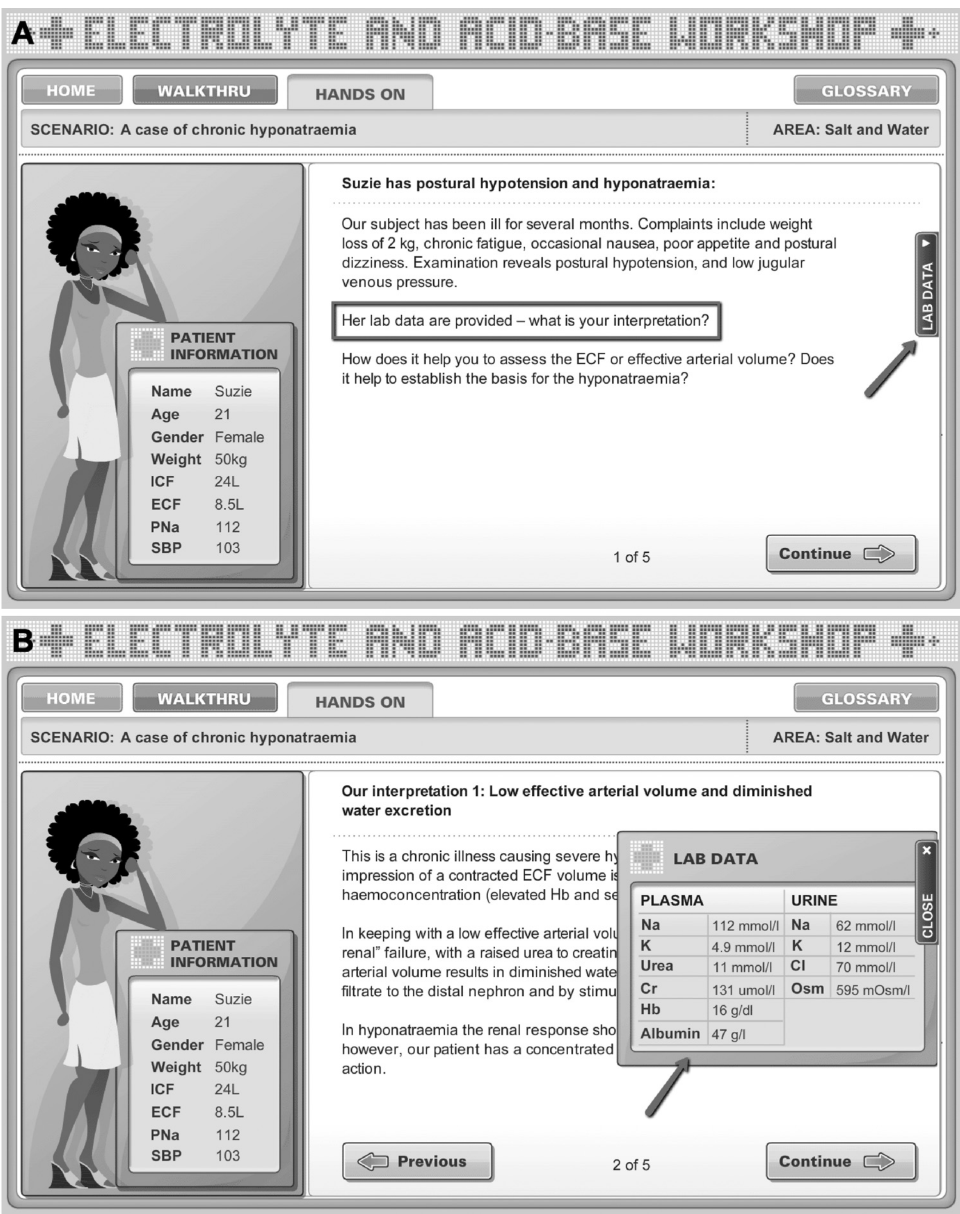



Fig. 1. The "hidden" laboratory data panel. *A*: the panel displays important laboratory data needed for assessing the case and deciding on appropriate treatment. It slides open on clicking the tab (arrow) at the side of the screen. Inspectors felt that this would be missed by some users despite the cue (surrounded by the rectangle) provided on *slide 1*. *B*: the open panel (arrow) also obscures other on-screen information and remains open when a user navigates to the next or to previous slides. The tab has to be clicked again to cause the panel to close.

usability catastrophe, may cause task failure and must be fixed before releasing the application. Each inspector submitted a written report based on the template provided.

All problems found were cataloged and categorized according to severity, the interface element involved, and the heuristic(s) involved. Problems with severity scores of *3* and *4* were grouped together as serious problems and were then compared with the serious problems previously found by user testing.

**RESULTS**

The evaluation was completed within 1 wk of supplying the inspectors with their documentation and the link to the application. Their overall impression of the application was uniformly positive, with comments such as “easy to use,” “good visuals,” and “an excellent application.”

A total of 22 distinct usability problems were identified. Examples of these, with the interface element involved, the heuristics violated, and potential solutions, are shown in Table 2. There were 11 problems categorized as serious; each of these was detected by a median of 2 inspectors (range: 1–4). Each inspector detected a median of 4 of the 11 serious problems (range: 3–7).

Several usability problems were identified that related to the heuristic of ensuring the visibility of system status and providing appropriate user feedback. Two inspectors were concerned about the long loading time of the application over slower internet connections; one inspector suggested adding a progress indicator to keep users informed during the loading process. Inappropriate or unhelpful feedback and error messages in the interactive simulation were highlighted by four inspectors. It was also suggested that treatments previously applied by users be displayed to them, accompanied by useful feedback.

Problems related to the heuristic of user control and freedom included the inability to navigate back to the lead-in slides after entering the treatment simulation. This was identified by four inspectors. It was suggested by three inspectors that the take-home message summary slide after the completion of the simulation be displayed to all users and not only to those who had completed the simulation successfully. In the WalkThru case, inspectors recommended adding the functionality to allow users to replay animations on the slides rather than requiring them to navigate away from the slide and then back again to have the animation replayed.

The heuristic of ensuring a match with the real world was violated by the use of the same character, Suzie, in both the WalkThru and HandsOn cases (and with different diagnoses). This was highlighted as potentially confusing. In the WalkThru case, clinical and laboratory parameters on the patient data panel were not updated appropriately after the successful treatment of the patient, also violating this heuristic.

Two problems were identified that resulted from violations of the heuristic of providing an intuitive visual layout. With regards to the lead-in slides of the HandsOn case, two inspectors pointed out that a laboratory data panel displaying the patient’s blood and urine chemistry results could easily be missed by users and suggested that their attention be drawn to it in some way. This panel slides open on clicking its tab at the side of the screen (Fig. 1). The problem of the open panel obscuring other on-screen information was also identified.

The heuristic of consistency and conformity to standards was violated by the use of too-small font sizes for the text on the slides. This was highlighted by two inspectors.

Inspectors also recommended reducing the word count and eliminating unnecessary animation on certain slides to conform to the heuristic of aesthetic and minimalist design.

The heuristic of error prevention and tolerance was violated in the design of the selection and application of treatments in the simulation. This was identified by four inspectors as a serious usability problem. The simulation was designed to permit treatments to be applied sequentially, and not simultaneously, so that appropriate feedback could be given after each step. Treatment options are grouped and displayed in separate panels (Fluid, Salt Treatment, and Drug Treatment) with only one panel open at a time (Fig. 2). Moving from one panel to the next causes the previous panel to be closed and any selected option in that panel to be deselected. Because the first panel

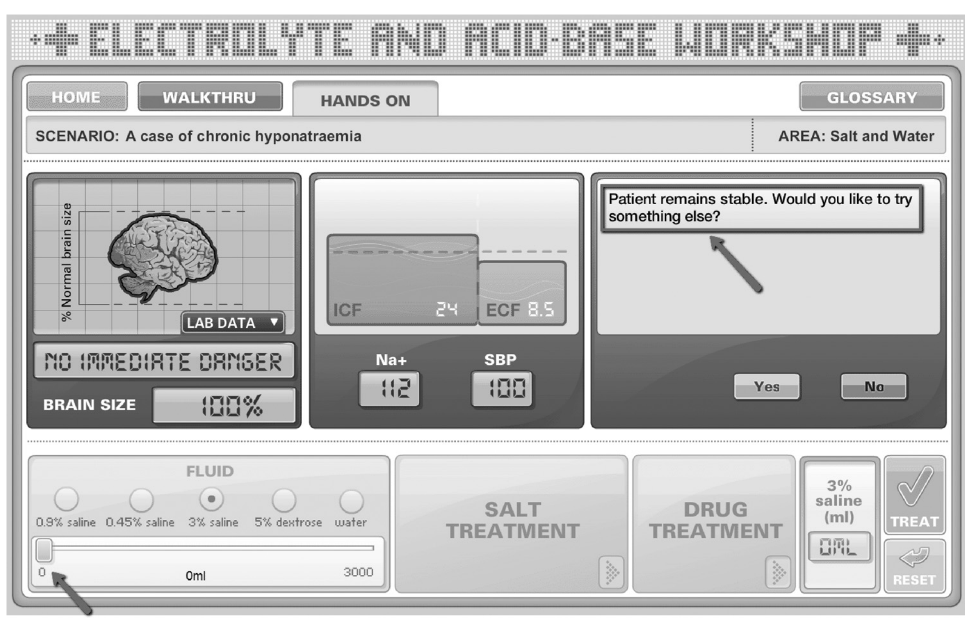



Fig. 2. Usability problems with treatment selection in the simulation. Treatment options are grouped and displayed in separate panels (Fluid, Salt Treatment, and Drug Treatment), with only one panel of options open and active at a time. Navigating from one panel to another causes the first panel to close and a selected option in that panel to be deselected. Here, the user has selected 3% saline from the Fluid panel (*bottom left* arrow) and then clicked the “Treat” button without first using the slider to indicate the dose. As a dose of zero has been administered, there is no change in any patient parameter. The feedback message (*top right* arrow) does not bring this problem to the user’s attention but is inappropriate and unhelpful.

Table 3. *Comparison of the serious usability problems detected by heuristic evaluation and user testing*

| | Problems Detected by Heuristic Evaluation | Problems Detected by User Testing |
|---|---|---|
| Slow loading of application; no progress indicator | ++ | ND |
| Same character as “patient” in both cases is confusing | ++ | ND |
| Text on slides uses too-small font size | ++ | ND |
| Inability to replay animations on slides | ++ | ND |
| Sliding laboratory data panel easily missed | ++ | ++ |
| Open laboratory data panel obscures other information | + | ++ |
| Insufficient information on treatment options, e.g., rate of administration | ++ | ND |
| Failed attempts at multiple treatment selection | ++ | ++ |
| Applying zero dosages with slider control | ND | ++ |
| Additional slider control problems* | ND | ++ |
| Unhelpful or inappropriate feedback messages | ++ | ++ |
| Cannot navigate backward once in the simulation | ++ | ND |
| No feedback/summary if simulation not successfully completed | ++ | ++ |

+, problem detected; ++, serious problem (detected with high frequency or impact); ND, problem not detected. *Figure 3 shows additional slider control problems revealed by user testing.

closes, users might not realize that the first treatment option was no longer selected and unsuccessfully attempt to select and apply multiple treatments simultaneously.

Inspectors also made suggestions relating to cosmetic changes and relatively minor usability problems. Examples of these included suggestions for font changes, adding a period after each glossary entry, and using the singular “Select your character” and not “characters” to indicate that only a single case scenario was presently available in each section of the application. There were also new feature requests that did not address an identified usability problem. An example of this was the suggestion that users have the ability to print summary notes of the cases upon completion.

A comparison of the detection of the most important usability problems by heuristic evaluation versus user testing is shown in Table 3. Among the problems identified by heuristic evaluation but not user testing were the need for a progress indicator while loading the application, text with too-small font sizes, unnecessary words and animation, the need to be able to replay the animations, and the problem with navigation. The most important problems identified by user testing but missed by the heuristic evaluation were the difficulties with using the slider control to select dosages in the treatment simulation (Figs. 2 and 3). User testing also highlighted the underutilization of the glossary: no participants accessed it from text hyperlinks as they worked through the slides. Those who viewed the glossary did so via the main navigation tab and at the end of the session, most likely only because this was required by the written instructions.

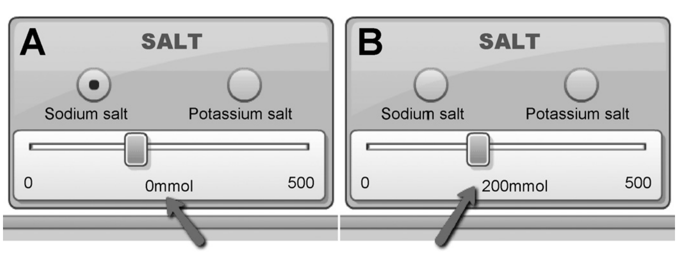



Fig. 3. Additional usability problems related to the slider control. *A*: here, the user has selected “Sodium salt” by clicking the appropriate radio button and then tried to indicate the dose by a single click on the rail of the slider. The slider thumb has jumped to the point clicked, but the dose still indicates 0 mmol (arrow). The correct dose was registered only when the slider thumb was dragged or the rail double-clicked. *B*: the slider thumb has been dragged to indicate the dose without the user first being required to select a treatment option by clicking one of the radio buttons.

## DISCUSSION

Heuristic evaluation of our Electrolyte Workshop by a panel of experts proved to be an efficient approach to improving usability. The evaluation was completed in a short space of time and detected most of the serious usability problems found by previous user testing as well as serious additional problems not identified by user testing. Heuristic evaluation thus presents an appealing option when time and financial resources are limited, as is often the case when developing e-learning materials. An additional advantage of using heuristic evaluation is that expert inspectors may often suggest solutions to problems found and may also highlight the strengths of a design.

A team of four to five experts can be expected to identify ~70% of usability problems (27). However, more problems will be missed when inspectors are inexperienced or lack domain expertise. Nielsen (22) found that novice inspectors uncovered 22% of problems, general usability professionals discovered 41%, and “double experts” who were specialists in usability as well as in the particular domain of the interface being tested were best found 60% of the problems. It is therefore important to have a good mix of experience and expertise when assembling a panel of inspectors.

Observing typical end users interacting with the application remains important however, as they may expose problems that experts, with their advanced computer skills, would not encounter (Table 3). The problem with the slider control is a case in point, where none of our expert panel had any difficulty dragging the slider to indicate the treatment dose in the simulation, whereas several participants in our earlier user testing study (4) could not work out how to use this at all, rendering the simulation unusable for these individuals. Another potential drawback of only using heuristic evaluation is that problems identified by inspection methods do not seem to have the same credibility with software developers and managers as those identified through testing “real” users (7).

User testing with the collection of subjective data by validated questionnaires is another attractive option when resources are limited. The System Usability Scale (2), for example, is freely available and easy to administer and yields a score of overall usability, which is useful for comparison with other applications and with different iterations of the same application. However, it does not generate a list of usability problems to fix and, on its own, would be of limited use when the aim is improving the application.

Heuristic evaluation and user testing each appear to identify important usability problems overlooked by the other method. It has therefore been suggested that both methods be used to supplement each other, with heuristic evaluation being used first to identify and correct the more obvious problems and, after the subsequent redesign, user testing be used to try and uncover the remaining problems (8, 23).

*Conclusions.* Heuristic evaluation is an efficient way of improving the design of e-learning materials in resource-constrained environments, considerably reducing the cost and time of evaluating usability. In terms of effectiveness, it compares well with user testing where typical end users are directly observed while using the application. In our study, heuristic evaluation detected several serious usability problems with our Electrolyte Workshop, each of which could have resulted in a substantial loss of educational impact. However, at least one serious problem was missed by heuristic evaluation, and we therefore support the recommendation that a combination of methods be used whenever possible, to increase the likelihood that most of the serious usability problems are detected and addressed. Ideally, heuristic evaluation should be used first and at an early stage in the development cycle. Combining heuristic evaluation with user testing, and involving a small number of participants with each cycle of testing, should provide valuable and rapid feedback to guide the development of usable e-learning materials for our health sciences programs.

## GRANTS

This work was supported by grants from the South African Universities Health Sciences IT Consortium and Stellenbosch University's Fund for Innovation and Research into Learning and Teaching.

## DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

## AUTHOR CONTRIBUTIONS

Author contributions: M.R.D., U.M.E.C., and M.L.H. conception and design of research; M.R.D. analyzed data; M.R.D., U.M.E.C., and M.L.H. interpreted results of experiments; M.R.D. prepared figures; M.R.D. drafted manuscript; M.R.D. and U.M.E.C. edited and revised manuscript; M.R.D., U.M.E.C., and M.L.H. approved final version of manuscript.

## REFERENCES

1. **Abran A, Khelifi A, Suryn W, Seffah A.** Usability meanings and interpretations in ISO standards. *Software Qual J* 11: 325–338, 2003.
2. **Brooke J.** SUS: a "quick and dirty" usability scale. In: *Usability Evaluation in Industry*, edited by Jordan PW, Thomas B, Weerdmeester BA, McClelland IL. London: Taylor & Francis, 1996, p. 189–194.
3. **Bygstad B, Ghinea G, Brevik E.** Software development methods and usability: perspectives from a survey in the software industry in Norway. *Interact Comput* 20: 375–385, 2008.
4. **Davids MR, Chikte U, Grimmer-Somers K, Halperin ML.** Usability testing of a multimedia e-learning resource for electrolyte and acid-base disorders. *Br J Educ Technol*; doi:10.1111/bjet.12042.
5. **Davids MR, Chikte UME, Halperin ML.** Development and evaluation of a multimedia e-learning resource for electrolyte and acid-base disorders. *Adv Physiol Educ* 35: 295–306, 2011.
6. **de Haan G, van der Veer GC, van Vliet JC.** Formal modelling techniques in human-computer interaction. *Acta Psychol* 78: 27–67, 1991.
7. **Desurvire HW.** Faster, cheaper!! Are usability inspection methods as effective as empirical testing? In: *Usability Inspection Methods*, edited by Nielsen J, Mack RL. New York: Wiley, 1994, p. 173–202.
8. **Dumas JS, Salzman MC.** Usability assessment methods. *Rev Hum Factors Ergonomics* 2: 109–140, 2006.
9. **Ericsson KA, Simon HA.** *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: MIT Press, 1984.
10. **Gould JD, Lewis C.** Designing for usability: key principles and what designers think. *Commun ACM* 28: 300–311, 1985.
11. **Hancock PA, Pepe AA, Murphy LL.** Hedonomics: the power of positive and pleasurable ergonomics. *Ergonomics Design* 13(1): 8–14, 2005.
12. **Hertzum M, Jacobsen NE.** The evaluator effect: a chilling fact about usability evaluation methods. *Int J Hum Comput Interact* 13: 421–443, 2001.
13. **Holzinger A, Errath M, Searle G, Thurnher B, Slany W.** From extreme programming and usability engineering to extreme usability in software engineering education (XP + UE→ XU). In: *Proceedings of the 29th Annual International Computer Software and Applications Conference*. Washington, DC: IEEE Computer Society, 2005, p. 169–172.
14. **Jeffries R, Miller JR, Wharton C, Uyeda K.** User interface evaluation in the real world: a comparison of four techniques. In: *Proceedings of the 1991 CHI conference on Human Factors in Computing Systems*. New Orleans, LA: ACM, 1991, p. 119–124.
15. **Karat CM, Campbell R, Fiegel T.** Comparison of empirical testing and walkthrough methods in user interface evaluation. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Monterey, CA: ACM, 1992, p. 397–404.
16. **Locher P, Overbeeke K, Wensveen S.** Aesthetic interaction: a framework. *Design Issues* 26: 70–79, 2010.
17. **Mahatody T, Sagar M, Kolski C.** State of the art on the cognitive walkthrough method, its variants and evolutions. *Int J Hum Comput Interact* 26: 741–785, 2010.
18. **Mao JY, Vredenburg K, Smith PW, Carey T.** The state of user-centered design practice. *Commun ACM* 48: 105–109, 2005.
19. **Miller C.** Aesthetics and e-assessment: the interplay of emotional design and learner performance. *Dist Educ* 32: 307–337, 2011.
20. **Molich R, Dumas JS.** Comparative usability evaluation (CUE-4). *Behav Inform Technol* 27: 263–281, 2008.
21. **Myers BA, Rosson MB.** Survey on user interface programming. In: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. Monterey, CA: ACM, 1992, p. 195–202.
22. **Nielsen J.** Finding usability problems through heuristic evaluation. In: *Proceedings of the 1992 SIGCHI conference on Human Factors in Computing Systems*. Monterey, CA: ACM, 1992, p. 373–380.
23. **Nielsen J.** Heuristic evaluation. In: *Usability Inspection Methods*, edited by Nielsen J, Mack RL. New York: Wiley, 1994, p. 25–62.
24. **Nielsen J.** Iterative user-interface design. *Computer* 26: 32–41, 1993.
25. **Nielsen J.** *Nielsen Norman Group*. *Ten Usability Heuristics* (online). http://www.nngroup.com/articles/ten-usability-heuristics/ [10 July 2013].
26. **Nielsen J, Mack RL.** *Usability Inspection Methods*. New York: Wiley, 1994.
27. **Nielsen J, Molich R.** Heuristic evaluation of user interfaces. In: *Proceedings of the CHI 1990 conference on Human Factors in Computing Systems*. Seattle, WA: ACM, 1990, p. 249–256.
28. **Renkl A.** The worked-out examples principle in multimedia learning. In: *The Cambridge Handbook of Multimedia Learning*, edited by Mayer RE. Cambridge, UK: Cambridge Univ. Press, 2005, p. 229–245.
29. **Rosenbaum S, Rohn JA, Humburg J.** A toolkit for strategic usability: results from workshops, panels, and surveys. In: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. The Hague, The Netherlands: ACM, 2000, p. 337–344.
30. **Sandars J.** The importance of usability testing to allow e-learning to reach its potential for medical education. *Educ Prim Care* 21: 6–8, 2010.
31. **Sohaib O, Khan K.** Integrating usability engineering and agile software development: a literature review. In: *International Conference on Computer Design and Applications*. Qinhuangdao, China: IEEE, 2010, p. 32–38.
32. **Tuch AN, Roth SP, HornbækKK, Opwis K, Bargas-Avila JA.** Is beautiful really usable? Toward understanding the relation between usability, aesthetics, and affect in HCI. *Comput Hum Behav* 28: 1596–1607, 2012.
33. **Wharton C, Rieman J, Lewis C, Polson P.** The cognitive walkthrough method: a practitioner's guide. In: *Usability Inspection Methods*, edited by Nielsen J, Mack RL. New York: Wiley, 1994, p. 105–140.
34. **Zaharias P.** Usability in the context of e-learning. *Int J Technol Hum Interact* 5: 37–59, 2009.

# CHAPTER 6

# EFFECT OF IMPROVING THE USABILITY OF AN E-LEARNING RESOURCE

## A randomized trial comparing the original version of the Electrolyte Workshop with a revised version optimised for usability

Published paper:

Davids MR, Chikte UME, Halperin ML. Effect of improving the usability of an e-learning resource: a randomized trial. Advances in Physiology Education 2014;38(2):155-60.

*Adv Physiol Educ* 38: 155–160, 2014;
doi:10.1152/advan.00119.2013.



# Effect of improving the usability of an e-learning resource: a randomized trial

**Mogamat Razeen Davids,[1] Usuf M. E. Chikte,[2] and Mitchell L. Halperin[3]**

[1]*Division of Nephrology, Department of Medicine, Stellenbosch University and Tygerberg Hospital, Cape Town, South Africa;* [2]*Department of Interdisciplinary Health Sciences, Stellenbosch University, Cape Town, South Africa; and* [3]*Li Ka Shing Knowledge Institute of St Michael's Hospital and Division of Nephrology, University of Toronto, Toronto, Ontario, Canada*

Submitted 10 October 2013; accepted in final form 13 January 2014

**Davids MR, Chikte UME, Halperin ML.** Effect of improving the usability of an e-learning resource: a randomized trial. *Adv Physiol Educ* 38: 155–160, 2014; doi:10.1152/advan.00119.2013.—Optimizing the usability of e-learning materials is necessary to reduce extraneous cognitive load and maximize their potential educational impact. However, this is often neglected, especially when time and other resources are limited. We conducted a randomized trial to investigate whether a usability evaluation of our multimedia e-learning resource, followed by fixing of all problems identified, would translate into improvements in usability parameters and learning by medical residents. Two iterations of our e-learning resource [*version 1* (V1) and *version 2* (V2)] were compared. V1 was the first fully functional version and V2 was the revised version after all identified usability problems were addressed. Residents in internal medicine and anesthesiology were randomly assigned to one of the versions. Usability was evaluated by having participants complete a user satisfaction questionnaire and by recording and analyzing their interactions with the application. The effect on learning was assessed by questions designed to test the retention and transfer of knowledge. Participants reported high levels of satisfaction with both versions, with good ratings on the System Usability Scale and adjective rating scale. In contrast, analysis of video recordings revealed significant differences in the occurrence of serious usability problems between the two versions, in particular in the interactive HandsOn case with its treatment simulation, where there was a median of five serious problem instances (range: 0–50) recorded per participant for V1 and zero instances (range: 0–1) for V2 ($P < 0.001$). There were no differences in tests of retention or transfer of knowledge between the two versions. In conclusion, usability evaluation followed by a redesign of our e-learning resource resulted in significant improvements in usability. This is likely to translate into improved motivation and willingness to engage with the learning material. In this population of relatively high-knowledge participants, learning scores were similar across the two versions.

usability; e-learning; multimedia; simulation

THE USABILITY OF TECHNOLOGY INTERFACES may have a major impact on learning, thus limiting the potential benefit obtained from using e-learning resources (2, 25, 31, 33, 37). We conducted a randomized trial to determine whether evaluating and optimizing the usability of a medical e-learning resource would result in improved measures of usability or learning.

The concept of usability derives from the field of human-computer interaction (HCI) and has been defined as the "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (1). Usability evaluation is well established in the software development industry (5, 12, 16, 22, 26, 28, 34), and there are often several cycles of testing and redesign before an application is released. This, however, is not common practice in medical education (33), where the importance of usability testing of e-learning resources is not yet widely recognized. Cost and time pressures are additional factors that may cause the evaluation of new resources to be neglected, with failure to achieve desired learning outcomes.

In the field of education, researchers have proposed guidelines for the design of e-learning resources based on cognitive load theory (CLT) (36) and the cognitive theory of multimedia learning (23, 24). These are based on a model of human cognitive architecture that views learning as involving active processing of information by working memory via separate visual and auditory channels. This system has a limited capacity. Any load that does not contribute to learning is considered extraneous and is likely to impede learning when the material is difficult and has a high intrinsic cognitive load (35). Mayer (23) has recommended several evidence-based principles to reduce extraneous cognitive load when designing multimedia learning resources. For example, according to the coherence principle, all irrelevant material should be eliminated, the signaling principle involves highlighting essential material, and the contiguity principle involves placing printed words near corresponding graphics.

There have been limited interactions to date between the fields of HCI and CLT (15). A recent review (15) reported that CLT concepts were mentioned in only 65 of >1.2 million citations in the Guide to the Computing Literature database. The two fields clearly share important concepts in that both strive to reduce extraneous cognitive load. In the case of HCI, this takes the form of usability guidelines such as "do not require the user to remember information from one screen to the next," designing for "recognition, not recall," encouraging "aesthetic and minimalist design," and "offer functionality only when needed" (27). In the case of CLT, there are instructional design principles such as the coherence, signaling, and contiguity principles. Hollender et al. (15) have proposed that the cognitive load induced by poor usability of e-learning interfaces be viewed as a specific component of extraneous cognitive load. This adds to the load resulting from poor instructional design.

Our interest is in developing learning resources to assist medical students and qualified practitioners in acquiring expertise in the diagnosis and treatment of electrolyte and acid-base disorders. This is a particularly challenging area of medicine (10). One of the resources we have developed is a web-based multimedia application called the "Electrolyte Workshop" (8). It is built in Adobe Flash and provides instruction and the opportunity for deliberate practice via an interac-

Address for reprint requests and other correspondence: M. R. Davids, Division of Nephrology, Dept. of Medicine, Stellenbosch Univ. and Tygerberg Hospital, PO Box 19063, Tygerberg 7505, Cape Town, South Africa (e-mail: mrd@sun.ac.za).

tive treatment simulation. The content has a high intrinsic cognitive load, and we therefore attempted to minimize any extraneous load by optimizing the usability of the application. We conducted a usability evaluation of the application by testing it with typical end users (7) and followed this by conducting a heuristic evaluation with a panel of experts (9). The information gained from these evaluations informed a comprehensive revision of our application.

This article reports on the effects of addressing the usability problems identified in our Electrolyte Workshop. Using a randomized trial, we investigated whether this had resulted in measurable improvements in usability and in improvements in learning. The reader is invited to examine the original and revised versions of the application at http://www.learnphysiology.org/sim1/ and http://www.learnphysiology.org/sim2/.

## METHODS

Ethics approval for the project was granted by the Committee for Human Research of the Faculty of Medicine and Health Sciences of Stellenbosch University (project no. N08/05/158).

### *The e-Learning Resource*

The application consists of case-based tutorials, each consisting of a series of slides, with the navigation controlled by the user. There are two main sections to the application. The first, called the WalkThru section, has cases with a "look-and-learn" approach similar to the use of worked examples in other disciplines (32). A clinical problem is presented followed by a demonstration of how an expert would analyze the data and embark on treatment. Animations illustrate changes in body fluid compartment sizes, brain cell size, and plasma $Na^+$ concentrations. The second section of the application, the HandsOn section, is interactive, with each case including a treatment simulation that provides the opportunity for deliberate practice. Users receive immediate feedback via on-screen text messages and animations.

### *Study Participants and Procedures*

Residents and subspecialty trainees (fellows) were recruited from the Departments of Medicine and Anesthesiology at Stellenbosch University in Cape Town, South Africa. Participants were randomly assigned to the different versions of the two cases using a computer-generated random number sequence, blocked randomization, and stratification by discipline (internal medicine vs. anesthesiology) and seniority (residents vs. specialists who were training in subdisciplines of internal medicine or anesthesiology). Allocation concealment was ensured using sequentially numbered, opaque envelopes.

The application was loaded onto two 15-in. laptop computers, which were each equipped with a mouse and a webcam with an integrated microphone. Morae usability software was installed on each computer to facilitate the recording and analysis of testing sessions. Participants were each required to work through the allocated versions of the WalkThru and HandsOn cases. No time limits were set. After each case, participants completed a user satisfaction questionnaire and answered a set of questions designed to test learning.

Technical problems resulted in the loss of certain of the Morae recordings and, hence, the objective data on some participants. Of the 18 participants allocated to each version of the WalkThru case, we had objective data for 17 participants in each group; of the 27 participants allocated to each version of the HandsOn case, we had objective data for 25 participants in the version 1 (V1) group and for 23 participants in the version 2 (V2) group.

### *Measures of Usability*

There is no single best measure of usability as each measure has its pros and cons and examines a particular aspect of usability. We followed the commonly recommended approach of using multiple usability measures and collected both subjective, self-reported data as well as objective data obtained by recording and analyzing the interactions of our participants with the application.

*Subjective measures.* A user satisfaction questionnaire that included the System Usability Scale (SUS) (4) was used to provide an overall measure of usability. The SUS can be used to compare different versions of a system and yields a single number (range: 0–100) with a score of 70 or greater regarded as acceptable. It is widely used, reliable, freely distributed, easy to administer, and easy to score (3, 4). We added a seven-point adjective rating scale as recommended by Bangor et al. (3). This item asked participants to rate the overall user friendliness of the application from being the worst imaginable (score of 1) through to the best imaginable (score of 7). Additional Likert-type questions asked participants to indicate whether the application increased their understanding and their confidence, whether navigation was difficult, whether they would recommend the application to others, and (for the HandsOn case) whether the simulation was realistic and engaging. The questionnaire also included two open-ended questions asking participants to comment on what they liked about the application and what they did not like or thought could be improved.

*Objective measures.* Successful task completion rates and the detection of usability problems were recorded for each task as measures of effectiveness, whereas time on task and input device activity (mouse clicks and mouse movement) were recorded as measures of efficiency. The WalkThru case, the introductory slides of the HandsOn case, and the treatment simulation of the HandsOn case were each regarded as a separate task. Task completion in the WalkThru case and the introductory slides of the HandsOn case simply required that participants view all the information available. For successful completion of the simulation, participants had to treat the patient effectively and end with the summary "take home messages" slide. The severity of each usability problem detected was determined by considering the frequency, persistence, and impact of the problem (29). A serious problem is one that may cause delays or task failure for the user and that needs to be fixed before an application is released.

### *Measures of Learning*

Eight questions related to the content of each tutorial were prepared. The first four questions tested recall, and the second four questions tested transfer. Participants were allowed 3 min/question, with each question printed on a separate sheet of paper and provided to them one at a time. Examples of the questions are shown in Table 1. The scores of the students were calculated by allocating one point for each correct answer; no penalties were given for incorrect answers. All answers were scored independently by a specialist physician and a nephrologist and were moderated by one of the authors (M. R. Davids).

### *Statistical Tests*

To compare scores across the two versions of the cases, the Wilcoxon rank-sum test was used for the SUS, adjective rating scale, and Likert-type questions. The *t*-test was used to compare SUS scores from the HandsOn case, as these were normally distributed. Fisher's exact test was used to compare the proportion of participants in each group with either positive or negative comments. It was also used to compare binary task completion rates and the proportion of participants encountering serious usability problems. Usability problem counts, time on task, mouse activity, and learning scores were compared using the Wilcoxon rank-sum test except where the data were

Table 1. *Measures of learning*

| Questions |
|---|
| Tests of retention |
| How did Suzie develop severe acute hyponatremia? Write down all the factors mentioned in the case that could have contributed. |
| Describe the major body fluid compartments in healthy individuals with respect to their volumes. |
| How would we know that antidiuretic hormone is acting on the kidney? |
| Write down all the case data you can remember. If you don't know the number you can simply indicate whether a parameter was normal (N), increased (↑), or decreased (↓). |
| Tests of transfer |
| "Runners hyponatremia" related to water overload may occur with long-distance races. You are advising the medical support team of next year's Two Oceans Ultramarathon. List all possible "risk factors" that could identify runners with a greater likelihood of developing acute hyponatremia during the race. |
| An athlete has a seizure at the end of a long-distance race. His plasma $Na^+$ concentration is 125 mmol/l. He is given 200 ml of 3% saline over 30 min. However, the followup plasma $Na^+$ concentration is 124 mmol/l and there is no clinical improvement. List the possible reasons why the plasma $Na^+$ concentration did not rise in response to treatment. |
| How much water would a 72-kg woman have to take in (and retain) to drop her plasma $Na^+$ concentration from 140 to 126 mmol/l? Show your calculations. |
| A 90-kg male patient developed acute hyponatremia from psychogenic polydipsia. You want to raise his plasma $Na^+$ concentration rapidly from 121 to 126 mmol/l. How many millimoles of $Na^+$ need to be administered? Show your calculations. |

Examples of questions designed to test the recall of information and questions to test the transfer of problem solving ability are shown. These are related to the WalkThru case.

normally distributed, in which case the *t*-test was used. The significance level was set at 0.05.

## RESULTS

### *Subjective Usability Data*

*SUS and adjective rating scale.* The results from the SUS and adjective rating scale are shown in Table 2. Mean scores were higher for the revised version of each case, but this difference was not significant. For the WalkThru case, mean SUS scores were 84.7 for the V1 group and 87.9 for the V2 group ($P = 0.27$). Scores on the adjective rating scale were 5.8 versus 5.9 for the V1 and V2 groups, respectively ($P = 0.36$). For the HandsOn case, SUS scores were 76.6 and 81.5 ($P = 0.13$) and adjective rating scale scores were 5.4 and 5.6 ($P = 0.20$) for the V1 and V2 groups, respectively. When the WalkThru and HandsOn cases were combined, SUS scores were significantly higher for revised versions ($P = 0.03$). There was a moderate to good correlation ($r = 0.68$) between the SUS scores and those of the adjective rating scale.

*Additional Likert-type questions.* The results from the additional Likert-type questions are shown in Table 3. Participants experienced navigation as more difficult in the first version of the HandsOn case compared with the revised version ($P = 0.02$). There were no other significant differences observed between the two versions of either case from this set of questions.

*Open-ended questions.* There were no clear differences in the number of positive or negative comments from participants in the different groups. A selection of quotes is shown in Table 4.

### *Objective Usability Data*

*Measures of effectiveness.* TASK COMPLETION RATES. The WalkThru case was successfully completed by all participants ($n = 17$ participants/group). With the more interactive HandsOn case, 18 of 25 participants successfully completed the first version, whereas 21 of 23 participants completed the second version ($P = 0.09$).

USABILITY PROBLEM COUNTS. As expected, participants encountered very few usability problems with the two versions of the WalkThru case. In total, five serious problem instances were recorded. These were encountered by five different participants: four participants from the V1 group and one participants from the V2 group ($P = 0.17$). With the interactive HandsOn case, serious usability problems were encountered by 22 of 25 participants in the V1 group as opposed to 2 of 23 participants in the V2 group ($P < 0.001$). The median number of serious problem instances recorded per participant was five (range: 0–50) for the V1 group and zero (range: 0–1) for the V2 group ($P < 0.001$). When these separate problem instances were consolidated into distinct usability problems for each participant, the median problem count was two (range: 0–4) for the V1 group and zero (range: 0–1) for the V2 group ($P < 0.001$). Of the 25 participants in the V1 group of the HandsOn case, 2 participants expressed frustration and 3 participants asked for help while using the application. There were no such events recorded in the V2 group.

*Measures of efficiency.* TIME ON TASK. Participants spent similar amounts of time on the two versions of each case. Mean times for the V1 and V2 groups were 11.8 ± 4.9 versus 12.7 ±

Table 2. *Scores for the two versions of each case and for both cases combined*

| | System Usability Scale | | | Adjective Rating Scale | | |
|---|---|---|---|---|---|---|
| | WalkThru case | HandsOn case | Both cases* | WalkThru case | HandsOn case | Both cases |
| V1 group | 84.7 ± 12.0 | 76.6 ± 18.2 | 79.8 ± 16.4 | 5.8 ± 0.5 | 5.4 ± 0.8 | 5.5 ± 0.7 |
| V2 group | 88.0 ± 14.0 | 81.5 ± 12.9 | 84.1 ± 13.6 | 5.9 ± 0.5 | 5.6 ± 0.6 | 5.7 ± 0.6 |

Values are means ± SD. V1 and V2, *versions 1* and *2*, respectively. The only significant difference observed was for System Usability Scale scores for both cases combined (*$P = 0.03$).

Table 3. *Answers to additional Likert-type questions*

| | Total Number of Participants | Item Description | | | | | |
|---|---|---|---|---|---|---|---|
| | | Increased my understanding | Increased confidence | Navigation difficult* | Would recommend | Simulation realistic | Simulation engaging |
| WalkThru case | | | | | | | |
| V1 group | 18 | 14 | 14 | 0 | 16 | | |
| V2 group | 18 | 13 | 14 | 2 | 15 | | |
| HandsOn case | | | | | | | |
| V1 group | 27 | 21 | 19 | 8 | 23 | 20 | 19 |
| V2 group | 27 | 21 | 20 | 1 | 23 | 22 | 23 |

Positive responses to the question items (i.e., agree and strongly agree) were combined. The only significant difference observed was for the item on navigation for the HandsOn case (*$P = 0.02$).

4.6 min for the WalkThru case ($P = 0.57$) and 19.2 ± 18.4 versus 18.4 ± 19.0 min for the HandsOn case ($P = 0.68$).

MOUSE ACTIVITY. Mouse activity was similar for the two versions except for a higher click count in the V1 group versus the V2 group of the interactive HandsOn case. For the WalkThru case, click counts for the V1 group versus the V2 group were 29.2 ± 14.4 versus 25.5 ± 15.0 clicks ($P = 0.89$) and mouse movement was 20,193 ± 29,978 versus 30,251 ± 27,082 pixels ($P = 0.05$). For the HandsOn case, click counts for the V1 group versus the V2 group were 142.6 ± 79.6 versus 89.0 ± 36.4 clicks ($P = 0.008$) and mouse movement was 73,259 ± 37,681 versus 66,724 ± 49,444 pixels ($P = 0.29$).

### *Measures of Learning*

*Tests of recall and transfer.* For the WalkThru case, recall test scores were 17.2 ± 2.6 and 16.6 ± 2.6 for the V1 and V2 groups ($P = 0.16$); for the HandsOn case, scores were 20.4 ± 5.0 and 21.0 ± 4.0 for the V1 and V2 groups ($P = 0.58$). For the WalkThru case, transfer test scores were 7.0 ± 3.1 and 6.9 ± 2.5 for the V1 and V2 groups ($P = 0.91$); for the HandsOn case, scores were 7.4 ± 3.2 and 6.6 ± 2.5 for the V1 and V2 groups ($P = 0.31$).

## DISCUSSION

A thorough evaluation followed by an extensive revision of our application resulted in measurable improvements in usability, in particular with regard to the HandsOn case with its interactive treatment simulation. The most striking finding was the large number of serious usability problems participants encountered in the original version of the HandsOn case compared with very few in the revised version. Nearly all the participants in the V1 group were affected but only two

Table 4. *Selection of responses to open-ended questions*

| Participants Liked the Following | Participants Did Not Like or Thought the Following Could Be Improved |
|---|---|
| *WalkThru case* | |
| "Bright, very good visuals. Clear smooth integration. Visuals and words coupled well together. Makes a sometimes daunting subject approachable/fun." (V1 group) | "Font very small!" (V1 group) |
| "Giving you a case, explain the treatment in a stepwise, easy to understand fashion. Also not too much detail." (V1 group) | "Perhaps a bit 'wordy' in places. Less paragraphs and more bullets/points perhaps." (V1 group) |
| "Contemporary example. Flows like a story–easier to remember the facts." (V2 group) | "Suzie's 'blinking eyes' distracted from the text on the last slide." (V1 group) |
| "Visually pleasing. Simple yet clear message. Useful animations that demonstrate the concept well." (V2 group) | "Too many different things to look @ at one time → gets distracting–I tended to ignore the graphics and just read the text." (V2 group) |
| *HandsOn case* | |
| "Excellent the way it responds and gives feedback. Take home messages are good too!" (V1 group) | "I did not clearly follow the last management steps. Do electrolyte and fluid administration and IV steroid use all impact on the outcome of the case simulation? Are all these maneuvers considered by the computer?" (V1 group) |
| "Being able to play around and see the effect of treatments administered." (V1 group) | "I think the way to use treatment options needs to be a bit explained before use." (V1 group) |
| "Animation again was excellent → seeing the consequences immediately of certain therapies was excellent." (V2 group) | "Lab data hidden (only found it after 5 minutes)." (V1 group) |
| "Real life case and can see what actually will happen if you give certain amount of fluids and sodium." (V2 group) | "Took me a while to figure out SBP was systolic blood pressure." (V1 group) |
| "Initially I was not thinking of the actual solution, rather fooling around with slides to see what would happen to brain if I give inappropriate therapy." (V2 group) | "Should add the appropriate management at the end of the case, as a teaching tool." (V1 group) |
| | "Have to drag slider; doesn't work if click at a certain point." (V2 group) |
| | "I would have liked a model answer with explanation." (V2 group) |
| | "When answering doesn't indicate which part of the answer was wrong → a bit frustrating." (V2 group) |

Shown are verbatim quotes from participants followed by the version group.

participants in the V2 group, suggesting that we had succeeded in eliminating most of the serious usability problems. Task completion rates and user satisfaction scores were also higher for the V2 group, although these were not statistically significant.

Expressions of frustration and requests for help were documented for participants in the V1 group but not in the V2 group.

We observed an interesting disconnect between subjective and objective measures of usability. Participants awarded high SUS and adjective rating scale scores to both versions of each case, even to the original version of the HandsOn case where many serious usability problems were encountered. This phenomenon has been noted previously (3, 7) and underlines the importance of not relying only on subjective measures of usability when evaluating e-learning resources or programs.

The improvements in usability were not accompanied by differences in learning, with scores on tests of retention and transfer being similar between the groups. A possible reason for the lack of impact on learning measures might be that our participants, all practicing clinicians, were not novices with regard to the subject area. All had received instruction on electrolyte and acid-base disorders as undergraduate students, some had received additional instruction in the course of their postgraduate training, and all of them had at least some experience in managing patients with these conditions. High-knowledge learners obtain less benefit when learning materials are designed to reduce cognitive load and, in some cases, may even suffer a decrease in performance, a phenomenon called the expertise-reversal effect (17). Another reason for the absence of a learning effect might be that our application implemented the segmenting principle (23) by allowing participants to control the navigation. This breaks the lesson into user-paced segments and is likely to minimize the negative impact of any extraneous load caused by poor usability.

While some researchers have reported significant learning effects from optimizing usability (2, 25), others have not observed differences but have found improvements in efficiency, satisfaction, or motivation. These effects are important in the light of the alarmingly high dropout rate from e-learning courses (38). Highly motivated and self-regulated learners are more likely to persist and succeed in e-learning environments, and optimizing usability can make an important contribution to their satisfaction and motivation. A study (13) among medical undergraduate students found that perceived quality of the e-learning program was an important determinant of their attitudes toward computer-based learning. In other studies, better usability resulted in improved task completion rates and less time on task (18) and in increased self-regulation by learners (21). Levy (20) found satisfaction to be a key indicator in the completion of online courses, whereas Zaharias and Poylymenakou (38) reported a strong relationship between learners' perceptions of system usability and their motivation to learn.

Traditional usability goals usually involve designing for effective and rapid task completion; however, systems that are less efficient to use or more difficult to learn may sometimes have a positive influence on motivation or learning. In a recent study (11), students who used disfluent learning materials with harder-to-read fonts had improved retention of the content compared with control students. This inclusion of "desirable difficulties" in their learning materials appeared to promote deeper processing and thereby improved learning.

Our study is a real-world example of the benefit of optimizing the usability of e-learning resources for medical education. The study participants were representative of our primary target audience, and the lack of a learning effect with these relatively high-knowledge learners is not surprising. As we also intend to use our Electrolyte Workshop for teaching undergraduate students, followup studies could investigate whether improved usability may translate into better learning for these novice learners. Compared with their senior colleagues, they are more likely to experience the content matter as having a high intrinsic cognitive load and should therefore be more sensitive to the addition of an extraneous load imposed by poor usability.

e-Learning has now become part of the medical education mainstream, with increasing investments in developing e-learning materials, modules, and programs. We would recommend that the usability of these resources be evaluated and optimized as a matter of routine. An iterative development process should be followed, with usability evaluation beginning early and involving both subjective and objective methods. Educators need to be aware that any existing digital divide will be widened by educational software that is poorly designed and that improving usability will lead to accessibility for a wider range of learners (6). Optimizing usability may therefore contribute to improved rates of persistence and success in e-learning environments.

Future research should examine the effect of optimizing usability and cognitive load on learning in learners who are novices regarding the subject matter, especially when the material to be learnt is complex, and in learners from the wrong side of the digital divide. A wider range of measures to evaluate the user experience will increasingly be used, including measures such as engagement, motivation, aesthetics, fun, and pleasure (14, 19, 30, 38).

In conclusion, the adoption of a design-test-redesign approach led to significant improvements in the usability of our multimedia e-learning application. This is likely to result in improved motivation and engagement with the learning resource and increases the chances of achieving desired educational outcomes. We support the recommendation that the development of e-learning materials should integrate user-centered technology design with learner-centered instructional design (15). The process should be iterative and focused on optimizing usability as well as on implementing principles of good instructional design based on CLT.

## ACKNOWLEDGMENTS

The authors thank Althea Goosen for the valuable assistance with data collection.

## GRANTS

This work was supported by a Doctoral Fellowship Award from Stellenbosch University's Faculty of Medicine and Health Sciences.

## DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

## AUTHOR CONTRIBUTIONS

Author contributions: M.R.D., U.M.E.C., and M.L.H. conception and design of research; M.R.D. performed experiments; M.R.D. analyzed data;

M.R.D. and U.M.E.C. interpreted results of experiments; M.R.D. drafted manuscript; M.R.D., U.M.E.C., and M.L.H. edited and revised manuscript; M.R.D., U.M.E.C., and M.L.H. approved final version of manuscript.

## REFERENCES

1. **Abran A, Khelifi A, Suryn W, Seffah A.** Usability meanings and interpretations in ISO standards. *Software Qual J* 11: 325–338, 2003.
2. **Avouris NM, Dimitracopoulou A, Daskalaki S, Tselios NK.** Evaluation of distance-learning environments: Impact of usability on student performance. *Int J Educ Telecommum* 7: 355–378, 2001.
3. **Bangor A, Kortum PT, Miller JT.** An empirical evaluation of the system usability scale. *Int J Hum Comput Interact* 24: 574–594, 2008.
4. **Brooke J.** SUS: a "quick and dirty" usability scale. In: *Usability Evaluation in Industry*, edited by Jordan PW, Thomas B, Weerdmeester BA, McClelland IL. London: Taylor & Francis, 1996, p. 189–194.
5. **Bygstad B, Ghinea G, Brevik E.** Software development methods and usability: perspectives from a survey in the software industry in Norway. *Interact Comput* 20: 375–385, 2008.
6. **Chalmers PA.** The role of cognitive theory in human-computer interface. *Comput Hum Behav* 19: 593–607, 2003.
7. **Davids MR, Chikte U, Grimmer-Somers K, Halperin ML.** Usability testing of a multimedia e-learning resource for electrolyte and acid-base disorders. *Br J Educ Technol*; doi:10.1111/bjet.12042.
8. **Davids MR, Chikte UME, Halperin ML.** Development and evaluation of a multimedia e-learning resource for electrolyte and acid-base disorders. *Adv Physiol Educ* 35: 295–306, 2011.
9. **Davids MR, Chikte UME, Halperin ML.** An efficient approach to improve the usability of e-learning resources: the role of heuristic evaluation. *Adv Physiol Educ* 37: 242–248, 2013.
10. **Dawson-Saunders B, Feltovich PJ, Coulson RL, Steward DE.** A survey of medical school teachers to identify basic biomedical concepts medical students should understand. *Acad Med* 65: 448–454, 1990.
11. **Diemand-Yauman C, Oppenheimer DM, Vaughan EB.** Fortune favors the bold (and the italicized): effects of disfluency on educational outcomes. *Cognition* 118: 111–115, 2011.
12. **Gould JD, Lewis C.** Designing for usability: key principles and what designers think. *Commun ACM* 28: 300–311, 1985.
13. **Hahne AK, Benndorf R, Frey P, Herzig S.** Attitude towards computer-based learning: determinants as revealed by a controlled interventional study. *Med Educ* 39: 935–943, 2005.
14. **Hancock PA, Pepe AA, Murphy LL.** Hedonomics: The power of positive and pleasurable ergonomics. *Ergonom Design* 13: 8–14, 2005.
15. **Hollender N, Hofmann C, Deneke M, Schmitz B.** Integrating cognitive load theory and concepts of human-computer interaction. *Comput Hum Behav* 26: 1278–1288, 2010.
16. **Holzinger A, Errath M, Searle G, Thurnher B, Slany W.** From extreme programming and usability engineering to extreme usability in software engineering education (XP+ UE→ XU). In: *Proceedings of the 29th Annual International Computer Software and Applications Conference*. Washington, DC: IEEE Computer Society, 2005, p. 169–172.
17. **Kalyuga S, Ayres P, Chandler P, Sweller J.** The expertise reversal effect. *Educ Psychol* 38: 23–31, 2003.
18. **Kanuka H, Szabo M.** Conducting research on visual design and learning: pitfalls and promises. *Can J Learn Technol* 27: 105–123, 1999.
19. **Khalid HM.** Embracing diversity in user needs for affective design. *Appl Ergonom* 37: 409–418, 2006.
20. **Levy Y.** Comparing dropouts and persistence in e-learning courses. *Comput Educ* 48: 185–204, 2007.
21. **Liaw SS, Huang HM.** Perceived satisfaction, perceived usefulness and interactive learning environments as predictors to self-regulation in e-learning environments. *Comput Educ* 60: 14–24, 2013.
22. **Mao JY, Vredenburg K, Smith PW, Carey T.** The state of user-centered design practice. *Commun ACM* 48: 105–109, 2005.
23. **Mayer RE.** Applying the science of learning to medical education. *Med Educ* 44: 543–549, 2010.
24. **Mayer RE.** Cognitive theory of multimedia learning. In: *The Cambridge Handbook of Multimedia Learning*, edited by Mayer RE. New York: Cambridge Univ. Press, 2005, p. 31–48.
25. **Meiselwitz G, Sadera W.** Investigating the connection between usability and learning outcomes in online learning environments. *J Online Learn Teach* 4: 9, 2008.
26. **Myers BA, Rosson MB.** Survey on user interface programming. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Monterey, CA: ACM, 1992, p. 195–202.
27. **Nielsen J.** *Nielsen Norman Group. 10 Usability Heuristics for User Interface Design* (online). http://www.nngroup.com/articles/ten-usability-heuristics/ [27 January 2014].
28. **Nielsen J.** Iterative user-interface design. *Computer* 26: 32–41, 1993.
29. **Nielsen J.** *Nielson Norman Group. Severity Ratings for Usability Problems* (online). http://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/ [27 January 2014].
30. **O'Brien HL, Toms EG.** The development and evaluation of a survey to measure user engagement. *J Am Soc Inform Sci Technol* 61: 50–69, 2010.
31. **Oviatt S.** Human-centered design meets cognitive load theory: designing interfaces that help people think. In: *Proceedings of the 14th ACM International Conference on Multimedia*. Santa Barbara, CA: ACM, 2006, p. 871–880.
32. **Renkl A.** The worked-out examples principle in multimedia learning. In: *The Cambridge Handbook of Multimedia Learning*, edited by Mayer RE. Cambridge, UK: Cambridge Univ. Press, 2005, p. 229–245.
33. **Sandars J.** The importance of usability testing to allow e-learning to reach its potential for medical education. *Educ Prim Care* 21: 6–8, 2010.
34. **Sohaib O, Khan K.** Integrating usability engineering and agile software development: a literature review. In: *International Conference on Computer Design and Applications*. Qinhuangdao, China: IEEE, 2010, p. 32–38.
35. **Sweller J, Chandler P.** *Why Some Material Is Difficult to Learn* (online). http://ammonwiemers.com/IdetPortfolio/articles/Instructional%20Design%20Theory/Why%20Some%20Material%20is%20Difficult%20to%20Learn.pdf [27 January 2014].
36. **Sweller J, van Merrienboer JJG, Paas F.** Cognitive architecture and instructional design. *Educ Psychol Rev* 10: 251–296, 1998.
37. **Zaharias P.** Usability in the context of e-learning. *Int J Technol Hum Interact* 5: 37–59, 2009.
38. **Zaharias P, Poylymenakou A.** Developing a usability evaluation method for e-learning applications: beyond functional usability. *Int J Hum Comput Interact* 25: 75–98, 2009.

# CHAPTER 7

# DISCUSSION AND CONCLUSIONS

# CHAPTER 7 – DISCUSSION AND CONCLUSIONS

The development and usability evaluation of an innovative multimedia e-learning resource, our Electrolyte Workshop, was successfully completed. This led to a comprehensive revision of the application and a comparison of the original and revised versions in a randomized trial.

This chapter summarises our results, then discusses implications and recommendations related to our study, other factors required for implementing successful e-learning, immersive learning environments, and ends with a section on future directions.

## Synopsis of results

### *Aim 1: Building the artefact (Chapter 2)*

Using Adobe Flash®, we developed a Web-based, multimedia application to provide students and practicing clinicians with instruction and the opportunity for deliberate practice in managing electrolyte and acid-base disorders. Our Electrolyte Workshop consists of case-based tutorials organized into two main sections. The WalkThru section uses a look and learn approach. A case of acute hyponatraemia is presented, followed by a demonstration of how an expert would interpret the data and embark on therapy. Animations illustrate changes in body compartment sizes, brain size, blood pressure, and plasma sodium concentrations. The navigation, and therefore the pace of the lesson, is controlled by the user. The HandsOn section is interactive and includes a treatment simulation. A case of chronic hyponatraemia is presented and users are able to select from a menu of therapies to 'treat' their patient. Immediate feedback is provided via animations and text messages. Chapter 2 provides hyperlinks to the Electrolyte Workshop and describes its submission to an international, peer-reviewed repository of health education resources.

*Aim 2: Description of our teaching approach, the development of the Electrolyte Workshop, and an initial usability evaluation (Chapter 3)*

In Chapter 3 we described the teaching approach that we have applied to the development of our learning resources. This approach is built around a good understanding of the relevant physiology and makes use of real cases and storytelling to engage the learner. There is good evidence that learning around the basic sciences promotes learning in disciplines like nephrology; that using authentic cases and a narrative approach improves learner engagement and motivation; and that active learning and the transfer of expertise is then more likely to occur.

The development of our Electrolyte Workshop involved a team of Flash® developers and began with the construction of a PowerPoint wireframe. We went through several cycles of testing and redesign, until we had a fully functional version which included an interactive treatment simulation. We have documented the challenges encountered during this process and make recommendations for the managing of similar projects. These include the use of wireframes and prototypes, following an iterative development process, and having a written agreement with the developers which includes software specifications, timelines and payments, ownership of the software, warranties, and provisions for dispute resolution.

An initial usability evaluation using the System Usability Scale, a widely used and validated user satisfaction questionnaire, is also reported in this chapter. On the 0 to 100 scale, a mean score of 78 was obtained, indicating a good level of usability.

*Aim 3: Usability evaluation by user testing (Chapter 4)*

Our user testing study involved postgraduate trainees and practicing clinicians in internal medicine, nephrology and endocrinology. Morae® usability software was used to facilitate the capture and analysis of information from testing sessions. A large number of usability problems were identified, many of them considered serious. Most were related to the interactive HandsOn tutorial and its treatment simulation, which was completed successfully by only 20% of participants. Our application was therefore unusable for most participants, thus severely limiting its potential educational impact. There was a striking disconnect between the objective measures of usability examined in this study and the subjective data reported in Chapter 3. The poor results obtained on objective measures such as task completion and problem counts stand in stark contrast to the good ratings given by participants on the System Usability Scale.

Our data on problem detection rates and the subsequent Monte Carlo simulation make a useful contribution to the debate on how many users are required to find 80% of the usability problems. We had a mean problem detection rate per participant of .123, substantially lower than the .3 which is the basis of the oft-quoted "five users are enough" recommendation. This means that we would need 13 users to find a mean of 80% of the total problems in our application.

*Aim 4: Usability evaluation by heuristic evaluation (Chapter 5)*

This study investigated whether using one of the inspection methods might be an efficient alternative to testing with end-users. This would be a very attractive solution for teachers who operate in resource-constrained environments.  We assembled a panel of inspectors with expertise in e-learning, instructional design, nephrology, medical informatics and usability. Our panel conducted an inspection by heuristic evaluation, and we compared the identification of serious usability problems by this method with the problems identified by user testing. Heuristic evaluation proved to be a very efficient method as compared to user testing, being completed in a short space of time and uncovering most – but not all – of the serious usability problems. Both user testing and heuristic evaluation detected serious problems which were missed with the other method.

*Aim 5: Revision of the Electrolyte Workshop*

The usability evaluations described in Chapters 3, 4 and 5 informed a comprehensive revision of our application. We managed this by first consolidating all the instances of problems into unique usability problems, mapping the problems to specific interface elements, and categorizing the problems by severity to prioritize them for fixing. All identified usability problems were addressed in the subsequent redesign, which again involved several iterations. We were then able to compare the original with an optimized version in a randomized trial. Chapter 2 provides hyperlinks to the two versions of the Electrolyte Workshop**.**

*Aim 6: Effect of improving the usability of an e-learning resource: a randomized trial (Chapter 6)*

Using a randomized trial, we investigated whether addressing the usability problems identified in our Electrolyte Workshop had resulted in measurable improvements in usability and in improvements in learning. Postgraduate trainees in internal medicine and anaesthesiology were randomly assigned to the original or optimized versions. We found large improvements in objective measures of usability but similar scores on measures of learning. As our participants were relatively high-knowledge learners and not novices, the absence of a clear learning effect was not altogether surprising. The marked improvement in usability is a very important finding, as optimizing usability contributes to the satisfaction and motivation of learners, improving the chances that they will persist and succeed in online environments.

## Implications and recommendations

Our studies clearly indicate that the usability evaluation of e-learning resources is critical. They provide a striking example of how an e-learning resource which was costly and time-consuming to produce was unusable for a substantial proportion of users. Had we not subjected our Electrolyte Workshop to usability evaluation we would have released a resource with very limited educational impact. This is a real-world example from the medical e-learning area where there is little published to date and it makes a strong argument for the benefit of evaluating and optimizing the usability of e-learning resources for medical education. We have illustrated how clinician-teachers who are not usability experts can set about improving the usability of the resources they develop and we believe that our experiences will be of practical value to teachers in medical education and other areas. We have demonstrated that a combination of methods should be employed and have also highlighted the utility of heuristic evaluation, which can be conducted in less time, and usually at much less expense, than user testing. Our problem discovery rates and Monte Carlo simulation support the notion that five users are not always enough and that more users should be involved, if resources allow this.

We strongly recommend that usability be evaluated as a routine part of the development and implementation of e-learning materials, modules or programmes. This should start with the earliest versions of the resource, ideally at the wireframe or prototype stage, when making changes is easier and much less costly. An iterative approach should be followed, with several cycles of testing and redesign, and involving a small number of participants each time. More participants will be required if the resource is complex, or if more mature versions are being tested. Heuristic evaluation by experts should be used first and, once the obvious problems have been identified and fixed, followed by testing with real users. User testing should always include the study of objective usability measures and not rely only on self-reported measures of user satisfaction.

### Other factors required for successful e-learning

While we have highlighted usability and cognitive load as two important elements, there are many other factors that must be addressed to ensure successful e-learning [1]. These include ensuring institutional buy-in and change management, ensuring that appropriate hardware and software are available, providing relevant skills training as well as technical and administrative support, proper integration of e-learning into the curriculum, including ensuring that assessments also include this material, using a blended learning approach, and continuously evaluating costs and benefits.

### Immersive learning environments

The focus must always remain on sound pedagogy, especially when highly interactive e-learning resources are being developed. Game-based e-learning and virtual learning environments are exciting innovations which allow the creation of immersive experiences which have the potential to increase learner engagement and motivation. Imperial College London, for example, has used the *Second Life*® virtual world platform to create a virtual hospital where medical students, through their avatars, can interact with patients in a virtual ward, order tests and review the results, and embark on treatment [2]. Such immersive technology is, however, very resource intensive and there is currently limited evidence available to support its educational impact [3, 4]. Issues which have been identified include a lack of realism, the linear design of patient cases providing an insufficient challenge, the need for reliable broadband Internet access and computer hardware with high specifications, and high overall costs [5, 6]. In low-resource settings such as South Africa these issues – and the paucity of evidence on educational impact – are especially relevant and should be carefully considered by educators who are contemplating the use of immersive learning environments. A report from the UK Joint Information Systems Committee [6] on virtual worlds and virtual learning environments concluded that they currently remain a fringe technology and that many challenges need to be addressed before they could be part of everyday educational practice.

## Future directions

### *Further content development*

In our Electrolyte Workshop we have the beginnings of a resource that can offer a rich learning experience and assist students and colleagues to acquire expertise in the challenging area of electrolyte and acid-base disorders. We now need to continue the process of content development, so that we present multiple examples of each type of disorder, allowing our users to encounter key concepts and the same physiology-based approach in a variety of contexts.

### *Technology considerations: from Flash to HTML5*

Adobe Flash® was selected as our development platform as it was the leading platform for animation on the Web when we started this project. That is no longer the case. Apple’s iOS operating system and many Android devices do not support Flash, putting Flash content out of the reach of many users who use mobile devices to access the Web. Adobe has since stopped its development of Flash for mobile devices and has joined the move to HTML5 as the future of rich application development for mobile devices. HTML (HyperText Markup Language) is the standard language used to create web pages. The advantages of HTML5, the latest HTML standard, include an open source structure, native support for rendering multimedia content without the need to install browser plug-ins, better power efficiency as regards device battery life, and better accessibility of content to search engines.

For the future development of our multimedia e-learning resources we are exploring authoring tools which can publish to multiple formats, including HTML5. The eLearning Guild (www.elearningguild.com) publishes regular reports on authoring and development tools [7] that could serve as a starting point for educators interested in developing their own materials.

*Future studies*

Our future research on the effect of optimizing e-learning resources for usability and cognitive load should be conducted with participants who are novices regarding the subject area. Undergraduate students, for example, are more likely to experience the content of the Electrolyte Workshop as having a high intrinsic cognitive load and should be more sensitive to any extraneous load imposed by poor usability. This will increase the chances of finding differences in learning as we test the effect of applying usability guidelines and principles of instructional design based on cognitive load theory.

*Reframing the user experience*

In planning new studies, we will need to take cognizance of the recent trend to use a broader range of measures to evaluate the user experience. This includes measures such as engagement, motivation, aesthetics, and fun [8-11]. The affective features of instructional messages can influence the level of learner motivation and engagement in deep processing. We should, for example, consider incorporating instructional design features aimed at priming motivation, while being careful not to overload the learner’s working memory [12].

*A need for integration between the fields of usability and cognitive load theory*

The fields of human-computer interaction and cognitive load theory share important concepts in that both strive to reduce extraneous cognitive load. The usability guidelines and instructional design principles that they have provided should be seen as complementary and should be considered as constituting best practice when it comes to the design of e-learning. However, to date there has been limited interaction between these two disciplines. We concur with Hollender et al. [13] that the cognitive load induced by poor usability of e-learning materials can be viewed as a specific component of extraneous cognitive load, adding to the load resulting from poor instructional design. We support their recommendation [13] that the development of e-learning materials should integrate user-centred technology design with learner-centred instructional design, resulting in the optimization of usability as well as the implementation of design principles based on cognitive load theory.

## REFERENCES


1. Childs S, Blenkinsopp E, Hall A and Walton G. *Effective e-learning for health professionals and students—barriers and their solutions. A systematic review of the literature—findings from the HeXL project.* Health Info Libr J 2005; **22**:20-32.
2. Toro-Troconis M, Kamat A and Partridge MR. *Design and development of a component-based system for virtual patients in the virtual world of Second Life®.* Journal of Emerging Technologies in Web Intelligence 2011; **3**(4):308-16.
3. Akl EA, Pretorius RW, Sackett K, Erdley WS, Bhoopathi PS, Alfarah Z, et al. *The effect of educational games on medical students' learning outcomes: a systematic review: BEME Guide No 14.* Med Teach 2010; **32**(1):16-27.
4. Sandars J. *Technology and the delivery of the curriculum of the future: opportunities and challenges.* Med Teach 2012; **34**(7):534-8.
5. Lee A and Berge ZL. *Second life in healthcare education: virtual environment's potential to improve patient safety.* Knowledge Management & E-Learning 2011; **3**(1):17-23.
6. Livingstone D, Peachey A, Callaghan M and Hemani A. *Supporting learning in Virtual Worlds with Virtual Learning Environments*. 2011, UK Joint Information Systems Committee.
7. Shank P and Ganci J. *eLearning authoring tools 2013: what we're using, what we want*. 2013, The eLearning Guild: Santa Rosa, CA.
8. Hancock PA, Pepe AA and Murphy LL. *Hedonomics: the power of positive and pleasurable ergonomics.* Ergon Des 2005; **13**(1):8-14.
9. Khalid HM. *Embracing diversity in user needs for affective design.* Appl Ergon 2006; **37**(4):409-18.
10. O'Brien HL and Toms EG. *The development and evaluation of a survey to measure user engagement.* J Assoc Inf Sci Technol 2010; **61**(1):50-69.
11. Zaharias P and Poylymenakou A. *Developing a usability evaluation method for e-learning applications: beyond functional usability.* Int J Hum-Comput Interact 2009; **25**(1):75-98.
12. Mayer RE. *Incorporating motivation into multimedia learning.* Learn Instr 2014; **29**:171-3.
13. Hollender N, Hofmann C, Deneke M and Schmitz B. *Integrating cognitive load theory and concepts of human-computer interaction.* Comput Hum Behav 2010; **26**(6):1278-88.

# APPENDIX 1

# LIST OF PRESENTATIONS

## LIST OF PRESENTATIONS

1. Davids MR, Halperin ML. A web-based system for teaching integrative physiology [scientific exhibit]. Stellenbosch University Faculty of Health Sciences Annual Academic Day, August 2001.

   [Won the award for Best Scientific Exhibit for our teaching website at www.learnphysiology.org]

2. Davids MR, Halperin ML. Development of an interactive simulation for the teaching of electrolyte and acid-base disorders. Second Annual Conference on the Scholarship of Teaching and Learning, Stellenbosch University, May 2008.

   [Won joint first prize for Best Presentation]

3. Davids MR, Halperin ML. Development of an interactive simulation for the teaching of electrolyte and acid-base disorders. SAAHE 1st National Health Sciences Education Conference. Stellenbosch University Faculty of Health Sciences, June 2008.

4. Davids MR, Halperin ML. Development of an interactive simulation for the teaching of electrolyte and acid-base disorders. ICERI 2008 - International Conference of Education, Research and Innovation. Madrid, Spain, November 2008.

5. Davids MR, Chikte UME, Halperin ML. An interactive simulation for the teaching of electrolyte and acid-base disorders. Second African Regional Cooperative Agreement (AFRA) Conference on ICT, Stellenbosch University, November 2009.

6. Davids MR, Chikte UME, Halperin ML. Usability evaluation of a multimedia resource for teaching electrolyte and acid-base disorders. Stellenbosch University Faculty of Health Sciences Annual Academic Day, August 2010.

   [Won prize for Best Presentation in the session on Health Sciences Education]

7. Davids MR, Chikte UME, Halperin ML. Development and usability evaluation of a web-based simulation for teaching electrolyte and acid-base disorders. American Society of Nephrology Congress, Denver, Colorado, November 2010.

8. Davids MR, Chikte UME, Halperin ML. Development and usability evaluation of a multimedia learning resource for electrolyte and acid-base disorders. Stellenbosch University Faculty of Health Sciences Annual Academic Day, August 2011.

9. Davids MR, Chikte UME, Halperin ML. Development and usability evaluation of a multimedia learning resource for electrolyte and acid-base disorders. South African Congress of Nephrology, Pretoria, September 2012.

10. Davids MR, Chikte UME, Halperin ML. Effect of improving the usability of an e-learning resource: a randomized trial. Combined 16th Ottawa Conference on Assessment and 12th Canadian Conference on Medical Education, Ottawa, Canada, April 2014.

11. Davids MR, Chikte UME, Halperin ML. Improving e-learning materials by optimizing usability and cognitive load. Invited lecture at the Division of Nephrology, St Michaels Hospital and University of Toronto, Toronto, Canada, April 2014.

12. Davids MR, Chikte UME, Halperin ML. Effect of improving the usability of an e-learning resource: a randomized trial. SAAHE National Health Sciences Education Conference, Cape Town, June 2014.

13. Davids MR, Chikte UME, Halperin ML. Improving e-learning materials by optimizing usability and cognitive load. Invited lecture at the Centre for Teaching and Learning, Stellenbosch University, August 2014.

# APPENDIX 2

## SUPPLEMENTARY PUBLICATIONS ILLUSTRATING OUR TEACHING APPROACH

# SUPPLEMENTARY PUBLICATIONS

The complete series of Masterclasses in Medicine articles is listed below, together with three other articles and a book chapter which illustrate our teaching approach.

## Masterclasses in Medicine articles


1. Cherney DZ, Davids MR, Halperin ML. Acute hyponatraemia and 'ecstasy': insights from a quantitative and integrative analysis. QJM 2002;95:475-83.

2. Davids MR, Edoute Y, Stock S, Halperin ML. Severe degree of hyperglycaemia: insights from integrative physiology. QJM 2002;95:113-24.

3. Davids MR, Lin SH, Edoute Y, Cheema-Dhadli S, Halperin ML. Hyponatraemia and hyperglycaemia during laproscopic surgery. QJM 2002;95:321-30.

4. Lin YF, Lin SH, Tsai WS, Davids MR, Halperin ML. Severe hypokalaemia in a Chinese male. QJM 2002;95:695-704.

5. Lin SH, Davids MR, Halperin ML. Hypokalaemia and paralysis. QJM 2003;96:161-9.

6. Edoute Y, Davids MR, Johnston C, Halperin ML. An integrative physiological approach to polyuria and hyponatraemia: a 'double-take' on the diagnosis and therapy in a patient with schizophrenia. QJM 2003;96:531-40.

7. Lin SH, Hsu YJ, Chiu JS, Chu SJ, Davids MR, Halperin ML. Osmotic demyelination syndrome: a potentially avoidable disaster. QJM 2003;96:935-47.

8. Luthra M, Davids MR, Shafiee MA, Halperin ML. Anorexia nervosa and chronic renal insufficiency: a prescription for disaster. QJM 2004;97:167-78.

9. Davids MR, Segal AS, Brunengraber H, Halperin ML. An unusual cause for ketoacidosis. QJM 2004;97:365-76.

10. Zalunardo N, Lemaire M, Davids MR, Halperin ML. Acidosis in a patient with cholera: a need to redefine concepts. QJM 2004;97:681-96.

11. Kamel KS, Cheema-Dhadli S, Shafiee MA, Davids MR, Halperin ML. Recurrent uric acid stones. QJM 2005;98:57-68.

12. Groeneveld JH, Sijpkens YW, Lin SH, Davids MR, Halperin ML. An approach to the patient with severe hypokalaemia: the potassium quiz. QJM 2005;98:305-16.

13. Bohn D, Davids MR, Friedman O, Halperin ML. Acute and fatal hyponatraemia after resection of a craniopharyngioma: a preventable tragedy. QJM 2005;98:691-703.

14. Alazami M, Lin SH, Cheng CJ, Davids MR, Halperin ML. Unusual causes of hypokalaemia and paralysis. QJM 2006;99:181-92.

15. Maccari C, Kamel KS, Davids MR, Halperin ML. The patient with a severe degree of metabolic acidosis: a deductive analysis. QJM 2006;99:475-85.

16. Carlotti APCP, Bohn D, Jankiewicz N, Kamel KS, Davids MR, Halperin ML. A hyperglycaemic hyperosmolar state in a young child: diagnostic insights from a quantitative analysis. QJM 2007;100:125-37.

17. Gowrishankar M, Carlotti APCP, St George-Hyslop C, Bohn D, Kamel KC, Davids MR, et al. Uncovering the basis of a severe degree of acidemia in a patient with diabetic ketoacidosis. QJM 2007;100:721-35.

18. Bockenhauer D, Cruwys M, Kleta R, Halperin LF, Wildgoose P, Souma T, et al. Antenatal Bartter's syndrome: why is this not a lethal condition? QJM 2008;101:927-42.

### Additional articles

19. Davids MR, Edoute Y, Halperin ML. The approach to a patient with acute polyuria and hypernatremia: a need for the physiology of McCance at the bedside. Neth J Med 2001;58:103-10.

20. Davids MR, Edoute Y, Mallie JP, Bichet DG, Halperin ML. Body compartment volumes and composition after giving a vasopressin antagonist: changes are revealed by a tonicity balance. Nephrol Dial Transpl 2002;17:300-3.

21. Lin SH, Lin YF, Cheema-Dhadli S, Davids MR, Halperin ML. Hypercalcaemia and metabolic alkalosis with betel nut chewing: emphasis on its integrative pathophysiology. Nephrol Dial Transpl 2002;17:708-14.

22. Davids MR, Edoute Y, Jungas RL, Cheema-Dhadli S, Halperin ML. Facilitating an understanding of integrative physiology: emphasis on the composition of body fluid compartments. Can J Physiol Pharmacol 2002;80:835-50.

23. Napolova O, Urbach S, Davids MR, Halperin ML. Assessing the degree of extracellular fluid volume contraction in a patient with a severe degree of hyperglycaemia. Nephrol Dial Transpl 2003;18:2674-7.

## Book chapter

24. Kamel KS, Davids MR, Lin SH, Halperin ML. Interpretation of electrolyte and acid-base parameters in blood and urine. In: Taal MW, Chertow GM, Marsden PA, Skorecki K, Yu ASL, Brenner BM, eds. Brenner & Rector's The Kidney 9th edition. Philadelphia, PA: Saunders; 2011:897-929.