Evaluating the Human-Computer Interaction of 'ECGSim': A Virtual Simulator to Aid Learning in Electrocardiology

Raymond R Bond[1], Eelco Van Dam[2], Peter Van Dam[2], Dewar D Finlay[1], Daniel Guldenring[1]
[1]Ulster University, Jordanstown, Northern Ireland, United Kingdom
[2] Peacs, Arnhem, Netherlands

Abstract

Studying electrocardiography is intellectually challenging since it involves a myriad of theories and complex associations. For example, the electrocardiogram (ECG) is a non-invasive recording of the electrical activity of the heart and its interpretation requires an intricate understanding of how these electrical signals relate to cardiac mechanics (referred to as the electromechanical link). Furthermore, it is often difficult to revert signals recorded from the body surface to the internal health of an organ. To help alleviate these challenges, the ECGSim software application was developed to allow visual learners to modify transmembrane potentials on the myocardium and view how these changes affect the ECG at the body surface. ECGSim is based on a state of the art complex 'forward model' that can be used to aid understanding of how the electrical activity propagates from the myocardium to the epidermis. However, to maximize the uptake of ECGSim, this study quantifies and validates its usability using a series of metrics.

1. Introduction

The electrocardiogram (ECG) is a commonly used tool in medicine to detect abnormalities in a person's heart [1]. However, learning and fully understanding the ECG and its physiological origins is a difficult challenge for students [2,3]. However, the ECGSim software application has been developed to reduce the burden of understanding these difficult electrophysiological concepts [3,4]. ECGSim is a software package that simulates the ECG from a complex heart model. The software interface allows the user to modify parameters associated with the heart and view how these changes affect the resulting ECG signals. It has been recognized that the uptake of technology enhanced learning software is related to the usability and user acceptance of the software. In light of this we have conducted the study, reported here which evaluates the usability of the ECGSim software application and in turn maximise its adoption.

2. Methods

This work uses validated methods to determine the usability of ECGSim. Ethical approval was sought for the study and approval was granted by Ulster University's computing science research ethics filter committee. Subjects were recruited at the 2014 Computing in Cardiology conference in Cambridge, USA. All sessions involved the following:

1. Subject read an information sheet and provided informed consent.
2. A pre-test survey was completed to collect subject demographics.
3. Each subject attempted a series of typical representative tasks using ECGSim (Table 1).
4. To elicit cognitive insights, each subject was instructed to 'think-aloud' whilst attempting each of the tasks. This is known as the concurrent think-aloud protocol [5].
5. Before attempting a task each subject was asked to rate the expected/perceived difficulty of the task. The subject also stated the amount-of-time they expected to take in order to complete the task. After each task they rated the actual difficulty of the task.
6. Screen-recording software and a digital microphone was used to record user interactions in addition to the subject's verbalisation (think-aloud data).
7. After the study, the subject completed a post-test survey to capture subjective opinions of the user interface.

This protocol facilitated the acquisition of quantitative metrics such as task completion rates (% of subjects that completed a task), task completion times, mean difference between pre- and post task difficulty ratings and frequency of 'use errors'. We also calculated the mean difference between the subject's anticipated task completion time and their actual task completion time. These comparisons were tested using a paired *t*-test (where α=.05). The study also used the System Usability Scale (SUS) model, which is an objective measure of usability [6]. The SUS model is a post-experiment survey comprising of 10 Likert style questions (e.g. "I think that

I would like to use this system frequently" where 1= Strongly Disagree and 5 = Strongly Agree). These 10 answers are then used to calculate a composite score. The mean SUS score is derived using the following equation.

$$\overline{SUS} = \frac{1}{n}\sum_{i=1}^{n} norm \cdot \sum_{j=1}^{m} \begin{cases} q_{i,j} - 1, & q_{i,j} \bmod 2 > 0 \\ 5 - q_{i,j}, & otherwise. \end{cases}$$

where n is the number of subjects and m is the number of questions (m=10). Thus, $q_{i,j}$ is a rating from one question by one subject. All ratings from odd numbered questions are subtracted by 1 and all ratings from evenly numbered questions are subtracted from 5. This is due to the fact that the odd questions have a negative connotation and even numbered questions have a positive connotation. The *norm* coefficient is equal to 2.5 and is used to normalize the SUS score (out of 100). And $\overline{SUS}$ provides the mean SUS score and hence represents the subject cohort. Practically, each SUS score is calculated as such: SUS=(Q1-1)+(5-Q2)+(Q3-1)+(5-Q4)+(Q5-1)+(5-Q6)+(Q7-1)+(5-Q8)+(Q9-1)+(5-Q10)*2.5. This model is useful as a myriad of systems have been previously evaluated using this model and thus a Gaussian distribution exists where the accepted average SUS score is 68/100. This model facilitated a benchmark and thus a 'grading on the curve' approach was used to objectively classify the 'usability' of ECGSim.

Table 1. The tasks performed using ECGSim.

#	Task (a larger description was made available)
T1	Can you find a way to show the chest electrodes in the heart panel?
T2	Modify a transmembrane potential located anywhere on the heart. Set the depolarization time to 100ms and view the changes this makes on the 12-lead ECG.
T3	Show the electrogram associated with the transmembrane potential you had modified?
T4	Reset the transmembrane potential to its original settings.
T5	Simulate ischemia in the LAD region and discuss the effects this has on the 12-lead ECG. Ischemia can be simulated by: (1) Shortening the duration of the transmembrane potential, (2) Reducing the amplitude of the of the transmembrane potential and (3) Delaying the depolarization of the transmembrane potential.
T6	Zoom into the specific changes this made to the T waves?
T7	Simulate ischemia on the posterior region of the heart and view how this affects the ECG.
T8	Visualize the heart vector and the vectorcardiogram.

3. Results

Fourteen subjects were recruited at the conference (10 males, 4 females, mean age=35±10). Ten subjects had previously used software simulators for medical purposes. The mean level of expertise in electrocardiography was 6 / 10 (this aided inclusion as all subjects were required to have basic knowledge of the electrocardiography and electrophysiology). There was also a high level of computer literacy amongst the subjects (mean computer literacy rating = 9 / 10), which helped alleviate any bias or confounding factors.

The task completion rates are shown in Table 2. This highlights that the usability of the ECGSim interface is sub-optimal for performing tasks 2, 6 and 7. This was due to the fact that these tasks required the use of 'hidden unaided features', e.g. using right-click button to rotate the heart and using the mouse scroll to amplify the signal. Nevertheless, 5/8 tasks achieved a 100% completion rate.

Table 2. Task completion rates/ratios for all eight tasks.

#	Task Completion Rate
T1	100% (14/14)
T2	**93% (13/14)**
T3	100% (14/14)
T4	100% (14/14)
T5	100% (14/14)
T6	**64% (9/14)**
T7	**79% (11/14)**
T8	100% (14/14)

Figure 1 illustrates the differences in the subject's anticipated task completion times and their actual task completion times. This figure is complemented by Table 3, which indicates that there is no statistical significance between user anticipated task completion times and actual task completion times. This fact may go some way to validate the usability of the interface. However, the comparison for task 7 is almost statistically significant (p-value=0.07) where subjects needed on average a further 53 seconds to complete the task in addition to the task completion time they had initially estimated. Figure 2 illustrates the differences in anticipated task difficulty ratings and post-task difficulty ratings. This figure is also complemented by Table 3, which indicates that tasks 1, 6 and 7 were statistically significant in being more difficult than expected by the user. On average subjects provided a higher difficultly rating for task 6 after having attempted it (mean difference between pre- and post difficulty rating for task 6 = 2.57). Conversely, with statistical significance task 5 exceeded the user's expectation.

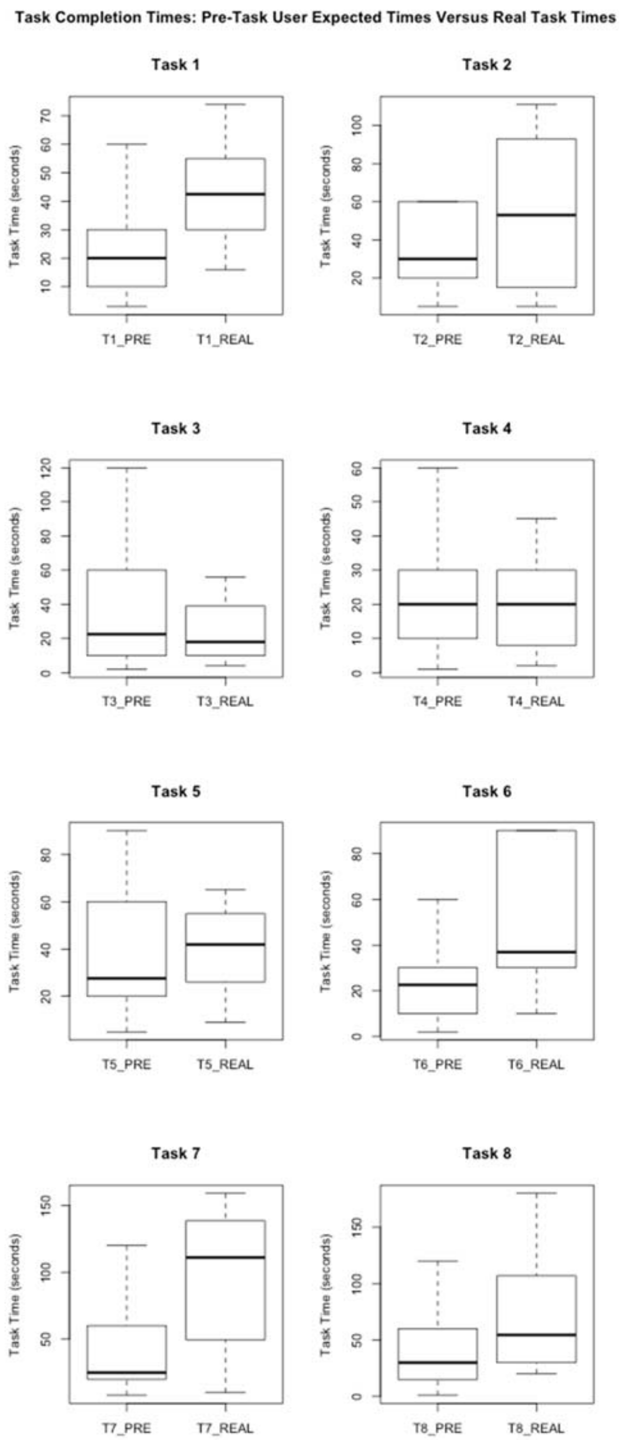

Figure 1. Box plots comparing the anticipated task completion times (Pre) and actual task completion times (Real).

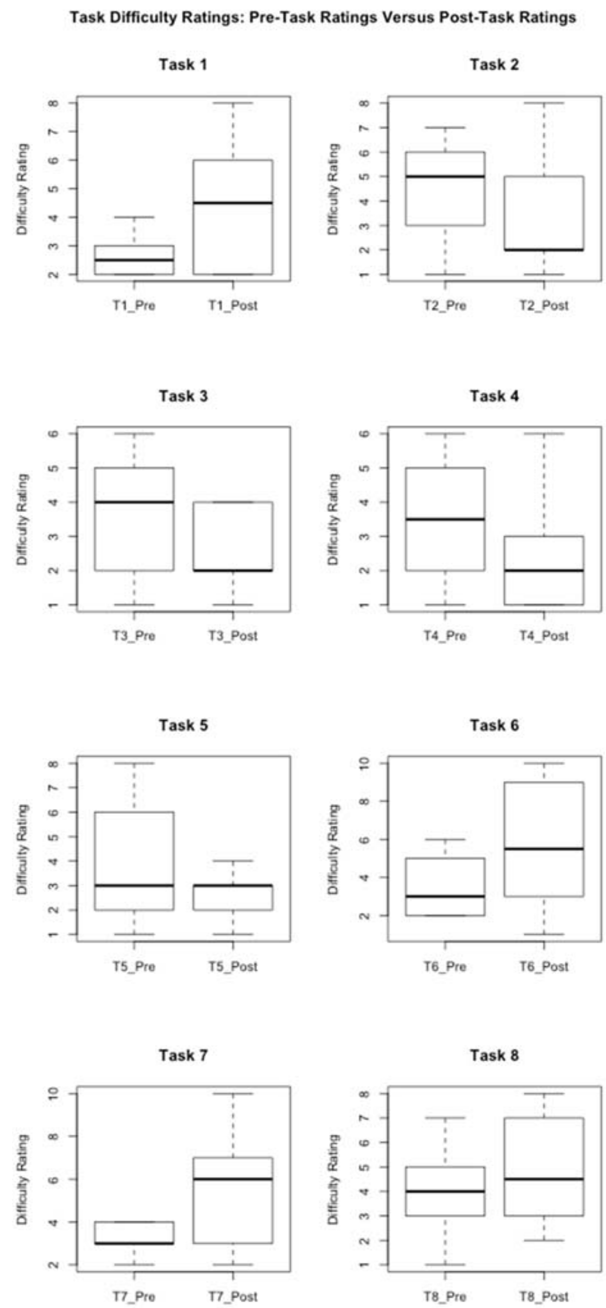

Figure 2. Box plots comparing the anticipated task difficulty ratings (Pre) and the post-task difficulty rating (Post), where 10 = 'Difficult' and 1 = 'Easy'.

Table 3. Mean differences in (1) pre- and post-task difficulty ratings and (2) between anticipated and actual task completion times.

#	Δ Difficulty Ratings	*p*-value	Δ Task Times	*p*-value
T1	1.43	***0.05**	26.57	0.21
T2	-1.07	0.16	23.46	0.38
T3	-0.50	0.56	-14.14	0.59
T4	-0.86	0.12	-8.86	0.36
T5	-1.21	***0.02**	1.14	0.93
T6	2.57	***0.02**	63.89	0.12
T7	2.07	***0.01**	53.27	0.07
T8	0.79	0.20	36.50	0.10

Figure 3 presents a boxplot of the individual SUS scores. The mean SUS score is 72 (SD=17), the mode is 78 and the median is also 78. The median SUS score of 78 achieved a percentile rank of 82 /100. This indicates that ECGSim is quantifiably more useable than 82% of all other software systems that have been previously evaluated using this model. In addition, when grading on the curve, the ECGSim software application achieved a usability classification of B+ (from possible classes A-F, where A = best).

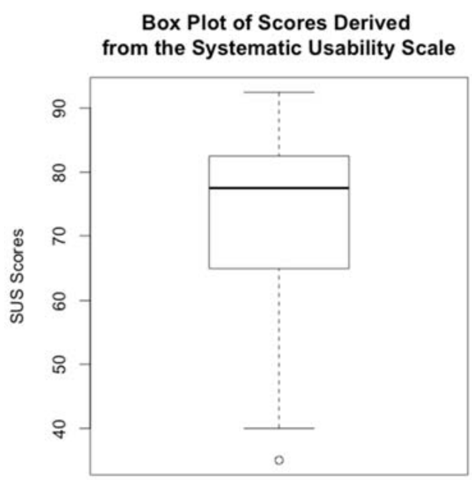

Figure 3. Box plot of the individual subject SUS scores (median = 78).

From the post-test survey, subjects rated the ‘usefulness’ of ECGSim (mean=9, where excellent=10) and its ‘look and feel’ (mean=8). All subjects stated that ECGSim would enhance their knowledge of electrophysiology and 12 out of 14 subjects stated that ECGSim could be understood without any formal training on how to use the software.

4. Discussion

Interestingly, the mean difference between the pre-task and post-task difficulty ratings proved to be a potent metric for discriminating the intuitive nature of the different tasks whereas the mean difference of anticipated and actual task times did not provide any statistical significance. This is a potential contribution to the human-computer interaction or user experience research discipline given the idea of ‘quantifying usability’ is a relatively new research domain that looks to identify usability metrics that have discriminant power [7].

5. Conclusion

In conclusion, ECGSim is an intuitive software platform and scores highly when using quantitative and qualitative metrics for measuring usability. In addition, the shortcomings identified in this study has allowed the user experience of ECGSim to be optimised further. Hence, the potential adoption of this software platform as a technology enhanced learning tool has in turn been maximised.

Acknowledgements

This work was funded by StitPro.

References

[1] Wagner GS, Marriott HJL. Marriott's practical electrocardiography. 11th ed. Philadelphia: Wolters Kluwer Lippincott Williams & Wilkins; 2008.

[2] Little B, Mainie I, Ho KJ, Scott L. Electrocardiogram and rhythm strip interpretation by final year medical students. Ulster Med J 2001 Nov;70(2):108-110.

[3] ECGSIM: Interactive Simulation of the ECG for Teaching and Research Purposes. Proceedings of the International Conference on Computing in Cardiology; 2010.

[4] Van Oosterom A, Oostendorp T. ECGSIM: an interactive tool for studying the genesis of QRST waveforms. Br Med J 2004;90(2):165-168.

[5] Lewis C. Using the "thinking-aloud" method in cognitive interface design. IBM TJ Watson Research Center; 1982.

[6] Brooke J. SUS - A quick and dirty usability scale. Usability evaluation in industry 1996;189(194):4-7.

[7] Sauro J, Lewis JR. Quantifying the user experience: Practical statistics for user research. Elsevier; 2012.

Address for correspondence.

Dr Raymond R. Bond, University of Ulster (UUJ), Shore Road Newtownabbey, Co. Antrim, BT370QB

rb.bond@ulster.ac.uk