Original Paper

Value of Eye-Tracking Data for Classification of Information Processing–Intensive Handling Tasks: Quasi-Experimental Study on Cognition and User Interface Design

Stephan Wegner[1], MSc; Quentin Lohmeyer[1], Dr-Ing; Dimitri Wahlen[1], MSc; Sandra Neumann[2], PhD; Jean-Claude Groebli[2], PhD; Mirko Meboldt[1], Dr-Ing

[1]Product Development Group Zurich, Institute of Design, Materials and Fabrication, Department of Mechanical and Process Engineering, Swiss Federal Institute of Technology in Zurich, Zürich, Switzerland

[2]Peripal AG, Zürich, Switzerland

Corresponding Author:
Stephan Wegner, MSc
Product Development Group Zurich, Institute of Design, Materials and Fabrication
Department of Mechanical and Process Engineering
Swiss Federal Institute of Technology in Zurich
Leonhardstrasse 21
Zürich, 8092
Switzerland
Phone: 41 446324862
Email: stehess@ethz.ch

Abstract

Background: In order to give a wide range of people the opportunity to ensure and support home care, one approach is to develop medical devices that are as user-friendly as possible. This allows nonexperts to use medical devices that were originally too complicated to use. For a user-centric development of such medical devices, it is essential to understand which user interface design best supports patients, caregivers, and health care professionals.

Objective: Using the benefits of mobile eye tracking, this work aims to gain a deeper understanding of the challenges of user cognition. As a consequence, its goal is to identify the obstacles to the usability of the features of two different designs of a single medical device user interface. The medical device is a patient assistance device for home use in peritoneal dialysis therapy.

Methods: A total of 16 participants, with a subset of seniors (8/16, mean age 73.7 years) and young adults (8/16, mean age 25.0 years), were recruited and participated in this study. The handling cycle consisted of seven main tasks. Data analysis started with the analysis of task effectiveness for searching for error-related tasks. Subsequently, the in-depth gaze data analysis focused on these identified critical tasks. In order to understand the challenges of user cognition in critical tasks, gaze data were analyzed with respect to individual user interface features of the medical device system. Therefore, it focused on the two dimensions of dwell time and fixation duration of the gaze.

Results: In total, 97% of the handling steps for design 1 and 96% for design 2 were performed correctly, with the main challenges being task 1 insert, task 2 connect, and task 6 disconnect for both designs. In order to understand the two analyzed dimensions of the physiological measurements simultaneously, the authors propose a new graphical representation. It distinguishes four different patterns to compare the eye movements associated with the two designs. The patterns identified for the critical tasks are consistent with the results of the task performance.

Conclusions: This study showed that mobile eye tracking provides insights into information processing in intensive handling tasks related to individual user interface features. The evaluation of each feature of the user interface promises an optimal design by combining the best found features. In this way, manufacturers are able to develop products that can be used by untrained people without prior knowledge. This would allow home care to be provided not only by highly qualified nurses and caregivers, but also by patients themselves, partners, children, or neighbors.

XSL·FO
RenderX

Introduction

Chronically ill patients cared for at home experience a higher health-related quality of life and a normalization of everyday life that is less dominated by the disease [1-3]. Therefore, 82% of end-stage renal disease patients and their families, if fully informed about their treatment options, would choose a home modality [4]. However, only 14% of dialysis patients in Europe are treated at home [5]. The main obstacle to home care is the availability of caregivers such as community nurses, neighbors, or relatives [6,7]. In order to allow a broad range of people the opportunity to ensure and support home care, one approach is to design medical devices with greater ease of use. This allows nonexperts to use medical devices that were originally too complicated to use. For user-centric development of such medical devices, it is essential to understand which user interface (UI) designs best support patients, caregivers, and health care professionals [8,9].

Human factors engineering drives user-oriented design and must test customized product UIs with intended users to determine the ideal level of mental workload. According to Kantowitz [10], mental workload is a subset of attention and the link between the demands of the environment and the capacity of the organism; it cannot be directly assessed. In a usability evaluation, the abstract term demand of the environment means fulfilling a task correctly. Consequently, when use errors occur, demand has not been met, and mental workload may have been too high or too low. This may be evaluated by analyzing the distribution and characteristic of attention in use error–related tasks.

Methods such as observations, questionnaires, and interviews are used to gain insight into the usability of an interface, but the focus is mainly on the graphical UI on a screen [9,11-15]. However, users gain most information through visual perception [16], and the short-term memory has only a limited capacity [17,18]. Therefore, it is difficult to understand the causes of use errors using traditional methods only.

Eye tracking provides a first-person perspective of the user and continuous localization of the gaze point. According to Hoang Duc et al [19], "tracking eye movements has the potential to provide a more direct measure of where attention is deployed since the direction of gaze is generally considered to be tightly coupled to the orienting of attention." Furthermore, "when people attend to a particular spatial location, there is greater neural processing in portions of the visual cortex corresponding to that location" [20]. Eye tracking thus allows objective feedback to find perception problems [21,22] and gain valuable insights into hotspots in attention distribution on the UI. This information can be used for both qualitative and quantitative evaluation of the usability of the UI. As a result, in recent years eye tracking has increasingly become a method for testing attention and improving or evaluating the features of UIs. Examples are web and print advertisements [23,24] and graphical representations like x-ray images of patients [25]. More complex subjects of the investigations include graphical UIs such as computer tomography interfaces [26] or spacecraft displays [27,28]. Further, there are single studies where eye tracking is used to evaluate highly interactive UIs of tangible products like smart TVs [29], smartwatches [12,21], or medical devices [30,31].

Most studies used a remote eye-tracking system where the stimulus is presented on a screen and participants are asked to sit still in front of a desk. Aside from this setup, mobile eye tracking with minimally invasive head-mounted systems provides a degree of freedom in movability. This promises natural user behavior in the testing of tangible medical devices [30].

In a first step of the eye-tracking data analysis, the raw gaze point data are classified into three events: fixation (nearly no eye movement), saccade (fast eye movement), and blink (closed eye). Since classified gaze data contain no semantic information on the looked-at objects or features, a second step of areas of interest mapping is needed. In this step, the single fixation events are manually assigned to the specific looked-at objects or UI features. As a result, data can subsequently be analyzed object-related in terms of durations of single fixations or cumulative dwell times (DTs) on an object or feature for a particular task. Fixation duration (FD), describing a property of visual attention per unit, is associated with the processing depth, which when increased leads to longer fixations [32-35], and with the rate of information extraction [23,35,36]. DT, describing the sum of visual attention related to specific objects or features, is associated with the length of the information extraction [28,37]. Thus, these measurements represent attention and, in the context of handling tasks, mental workload as a subset of attention in two dimensions.

Using the benefits of mobile eye tracking, we aimed to gain a deeper understanding of the challenges and differences in user cognition and thus identify obstacles to the user-friendliness of single UI features of a patient assistance device intended for home use in peritoneal dialysis (PD) therapy. This paper describes, to our knowledge, the first benchmark tests of two different UI designs based on physiological measurements using mobile eye tracking. The underlying research questions of this work are as follows:

- RQ1: Do slight differences in the design of the UI of a patient assistant device lead to differences in the effectiveness of use?
- RQ2: What are the differences in visual perception between two UI designs of a patient assistant device related to single task-relevant UI features in use error–related handling tasks?

Methods

Summary

The aim of this work was to gain a deeper understanding of the challenges in user cognition and thus the obstacles to user-friendliness of single UI features of a medical device. Therefore, a quasi-experimental study was conducted for data collection with the medical device with the intention of being as realistic as possible and representing the intended use. As a result, naïve representatives of the user group of patients (young adults and seniors) were recruited, and the study was conducted in the intended environment.

Stimuli

The stimuli of the study were prototypes with two different UI designs (D1 and D2, see Figure 1) of a medical device system.

The system consists of medical device, inlet for guiding and manipulating a bag system with dialysis fluid, and catheter, which is connected to the patient in the real therapy application. The most important interface features of the medical device are the buttons for manipulating the bag system and the lever for moving the inlet inside the device. The inlet has functions for fixing, clamping, and opening a predetermined breaking point feature inside the bag lines. The medical device system supports PD handling and is aimed at adults aged 18 and older. The stimuli provide acoustic (click sounds), haptic (positioning by stops), and visual (clear states and observation windows) feedback. Both prototypes support the same functionalities and require the same handling steps. At the top level, appearance of the UI designs was neutral in a monochrome design, as shown in Figure 1, to eliminate the effects of different coloring as an additional variable.

Figure 1. Illustration of user interface designs D1 (left) and D2 (right) including features lever and buttons 1-3. Additional parts for the therapy handling with the medical device are bag lines and catheter (standard parts used in therapy) and inlets D1 and D2, compatible with their respective designs.
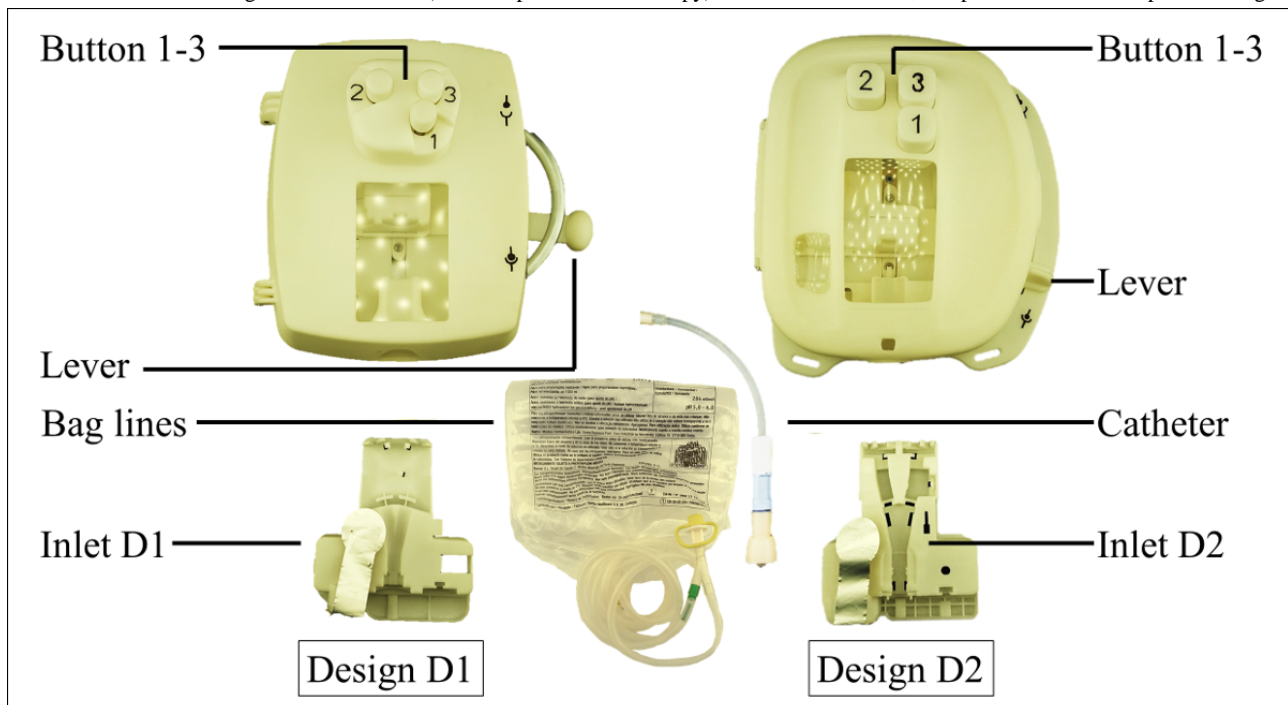
Recruitment and Data Exclusion

A total of 25 participants (18 men and 7 women, average 50.2 years, range 24 to 90 years) were recruited and participated in this study. The sample was recruited from a retirement home (10 men and 5 women, average 74.0 years, range 67 to 90 years) and from university (8 men and 2 women, average 25.1 years, range 24 to 26 years). In the PD patient population in Europe, 52% are younger than age 65 years [5]. Due to potential technical challenges with the eye-tracking technology related to the physiology of the eye area, which is especially relevant for seniors as reported by Bojko [38], more participants were invited than analyzed in the final analysis. All participants were in good physical and mental condition and assessed the suitability of study participation themselves. No participant was familiar with PD therapy or mobile eye tracking. All participants

had normal or corrected vision with contact or corrective lenses that could be connected to the mobile eye-tracking system.

One senior left the study prematurely after the first handling cycle and was therefore excluded from the analysis. For five seniors and one young adult, data quality was insufficient due to measure errors by the eye tracker resulting from drooping eyelids, watery eyes, or long eyelashes. In order to achieve a counterbalance in terms of the order of use of the two designs and represent the target population characteristic of PD patients in age, the data sets of a randomly selected senior and young adult were not included in the data analysis. Thus, a total of 16 data sets with 8 data sets from each group of young adults (25.0 years on average) and seniors (73.7 years on average) could be analyzed. Four participants in each group started with D1 and four with D2, achieving a complete counterbalance. As a result, this within-subject design mitigated the effect of individuality. Consequently, measures that naturally differ from participant

to participant, such as FDs, could be compared with this balanced design of the study.

Study Procedure

When participants arrived in the test environment, they were welcomed and thanked for their participation. Before the study began, participants were asked to read information on the goal of the study, data safety, and data management. If they agreed to participate in the study, they were asked to sign the consent form. Subsequently, participants put on the mobile eye-tracking system, and the moderator conducted a 3-point calibration. Since all participants were beginners in PD therapy and in the use of the device, the moderator briefly described the disease and associated PD therapy. Next, the moderator demonstrated the handling procedure with a low-level representation of the UI, designed and built for this purpose, and the devices. After the

introduction, participants performed the handling cycle of tasks 1 through 7 in a simulated PD therapy (see Figure 2), starting either with D1 or D2 and guided by written instructions. Each instruction was printed in a neutral design on an individual sheet to test the usability of the medical device and not the instruction. There was no time limit for the fulfillment of tasks, and the moderator assisted only in cases where the study would otherwise have had to be terminated due to the use error. Subsequent to the first completed handling cycle, participants were asked to give their feedback on usability in a semistructured interview with predefined high-level questions asking for general feedback on tasks related to use errors, guiding to the root causes of handling difficulties and use errors. Starting with the handling cycle, this process was repeated for the remaining prototype of the UI design.

Figure 2. Seven tasks in medical device handling cycle. User interacts manually with inlet, bag lines, catheter, and user interface features lever and buttons 1-3.
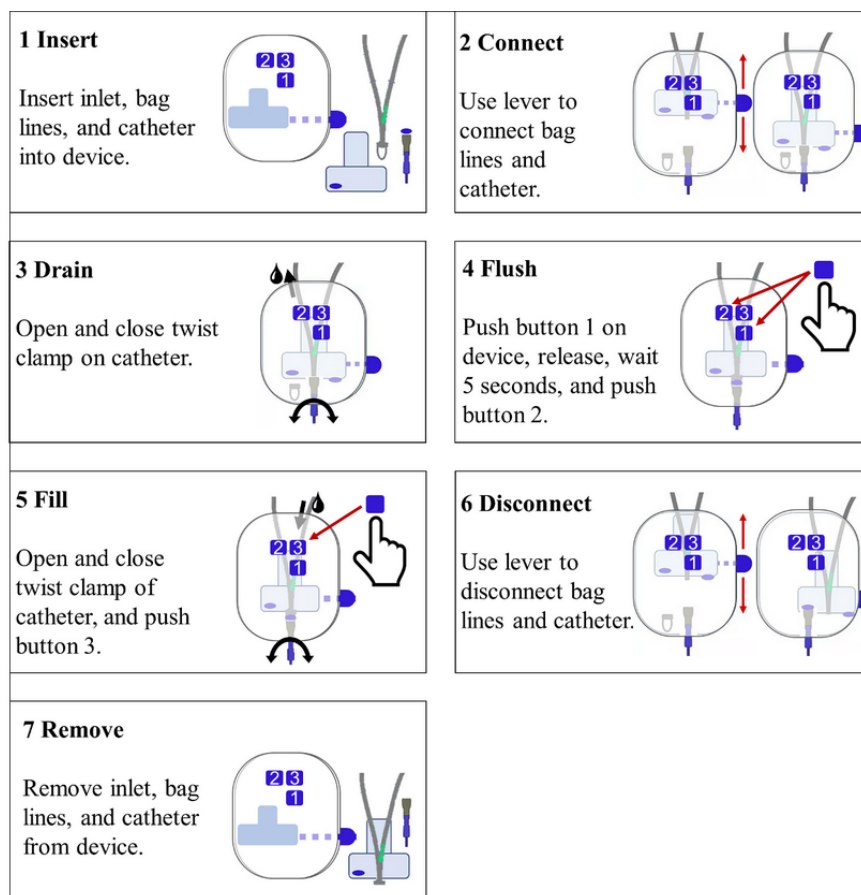
Data Analysis

In the data analysis, a 2-step approach was used. It started with the analysis of task effectivity searching for use error–related tasks. Subsequently, the in-depth gaze data analysis focused on these identified critical tasks.

For analysis of the task effectivity, the handling process of participants was observed via a live recording from the first person's perspective from the eye-tracking system. The performance in each task was evaluated by an observer. In the evaluation, two categories were distinguished according to the international standard IEC 62366-1 (2015). The first category,

safe use, is defined as "normal use without use error" [39]. The second category, use error, is defined as "user action or lack of user action while using the medical device that leads to a different result than that intended by the manufacturer or expected by the user" [39].

Gaze data were recorded with the mobile eye-tracking system, SMI Eye Tracking Glasses 2 (SensoMotoric Instruments GmbH), with a scene resolution of 1280×960 pixels (viewing angle: 60° horizontal, 46° vertical) of the front camera offering a sampling frequency of 24 Hz with the gaze point measurement having an accuracy of 0.5° over all distances. The raw gaze point data were classified into the events of fixations, saccades,

and blinks by SMI BeGaze version 3.7 (SensoMotoric Instruments GmbH). Subsequent to the areas of interest analysis, information on FD for depth and rate and DT for length of information extraction related to particular objects and features of the medical device system (Figure 1) were calculated. This information was used to understand the challenges of user cognition in use error–related tasks. Blinks were not considered in this work.

Combining the information on FD and DT, the data were analyzed with a multivariate analysis of variance (MANOVA) using SPSS Statistics 24 (IBM Corp). The MANOVA had one independent variable with two levels, D1 and D2 (see Figure 1), two dependent variables, FD and DT, both measured on a ordinal level and representing the rank of the mean

measurements for every participant for the UI features in error-related tasks.

For a better understanding of the two analyzed dimensions, length and depth of visual perception, Figure 3 combines information on the two measured parameters. Evaluating user perception of all participants as a whole, it shows the relationship between mean FD and mean DT for individual UI features of D1 compared with D2. Based on the two analyzed dimensions assigned to D1 in the coordinate origin, the mean FD and mean DT of D2 can be longer or shorter. Consequently, four different categories or patterns can be distinguished. A suggested interpretation of these patterns in terms of workload or gaze behavior is shown in Figure 3. Equations for calculating the values of the shift in both dimensions (ΔDT and ΔFD) from D1 to D2 for the diagram can be seen in Figure 4.

Figure 3. Visualization of shifts in two dimensions of the physiological gaze data measurements fixation duration and dwell time. The displayed shifts are from a Design D1 in the coordinate origin to a Design D2, presented in the middle column. In total, a distinction is made between the four categories. The right column explains the four patterns.

Figure 4. Equations for calculating the values of the shift in both dimensions of dwell time and fixation duration (ΔDT and ΔFD) from UI design D1 to D2.

$$\Delta DT = \frac{DT_{D2} - DT_{D1}}{DT_{D1}} \quad (1)$$

$$\Delta FD = \frac{FD_{D2} - FD_{D1}}{FD_{D1}} \quad (2)$$

Results

Each of the 16 participants performed 30 handling steps in the 7 tasks with both UI designs, resulting in 480 evaluated handling steps for each UI design. The results of the task performance are shown in Figure 5. Overall, 97% of the handling steps were performed correctly for D1 and 96% for D2.

According to the results, the main challenges were in task 1 (insert), task 2 (connect), and task 6 (disconnect) for both UI designs. The remaining four handling tasks were performed without errors, except for one missing catheter closure in task 5 (fill) with D2. Observed use errors in the first task were mainly

incorrectly inserted bag lines in the inlet. Further use errors were forgetting to attach the cap of the bag lines to a safety feature on the device and folding the protective film of the inlet outwards. All use errors were discovered and corrected by participants at a later stage of the handling cycle. In task 2 (connect), use errors occurred when the lever should have been used to connect bag lines and catheter. In task 6 (disconnect), some participants forgot to operate the lever for disconnecting catheter from bag lines and for placing a new cap onto catheter. In the semistructured interview, participants mentioned difficulties positioning bag lines and catheter in the inlet, oblivion of some details in the handling from the presentation, hesitation because of fear of breaking something, and misleading

wording in the instructions for tasks 2 (connect) and 6 (disconnect). In addition, participants gave positive and negative feedback on the overall impression and experience with the device.

Multimedia Appendix 1 focuses on handling tasks with observed use errors and shows the results of the data analysis of the physiological gaze data in both dimensions. The mean values for FD are given in milliseconds and for DT in seconds. The mean FD for single UI features was between 149 and 405 milliseconds. The mean DT for single UI features was between 0.3 and 28 seconds. At task 1 (insert), there were large shifts from the UI features bag lines, inlet, and catheter to the UI features levers and buttons. While the first group had average DTs between 7 and 28 seconds, the second group had average DTs between 0.3 and 3 seconds. At task 2 (connect) and task 6 (disconnect), the DT varied from less than 1 second for the buttons to 3 seconds for bag lines and catheter. For the mean FD, clustering was not found in any of the three tasks.

A MANOVA revealed statistically significant differences between D1 and D2 for catheter (Pillai trace=.216, $F_{2,29}$=3.985,

P=.03) and for lever (Pillai trace=.348, $F_{2,25}$=6.674, P=.005) in task 1 (insert) and for inlet (Pillai trace=.22, $F_{2,25}$=3.534, P=.045) in task 6 (disconnect). All other UI features showed no statistically significant differences in the three error-related tasks.

For better understanding, Figure 6 visualizes the data presented in Multimedia Appendix 1. As shown in Figure 3, this visualization combines FD and DT as two dimensions of the gaze data. In task 1 (insert; Figure 6A), the mean DT for all task-relevant UI features is longer for D2. The bag lines show a strong category 1 pattern, while the other two UI features show little to no shift for the mean FD. For task 2 (connect; Figure 6B), three UI features show a strong category 4 pattern, while the bag lines show mainly shorter mean DTs and only a slightly longer mean FD, thus showing a weak category 2 pattern. For task 6 (disconnect; Figure 6C), the UI elements located inside the device in this task show a strong category 2 pattern, while the lever on the outside of the device shows a strong category 1 pattern.

Figure 5. Comparison of task performance between D1 and D2 for all seven tasks. Evaluation in two categories, safe use and use error, according to International Electrotechnical Commission 62366-1 (2015).
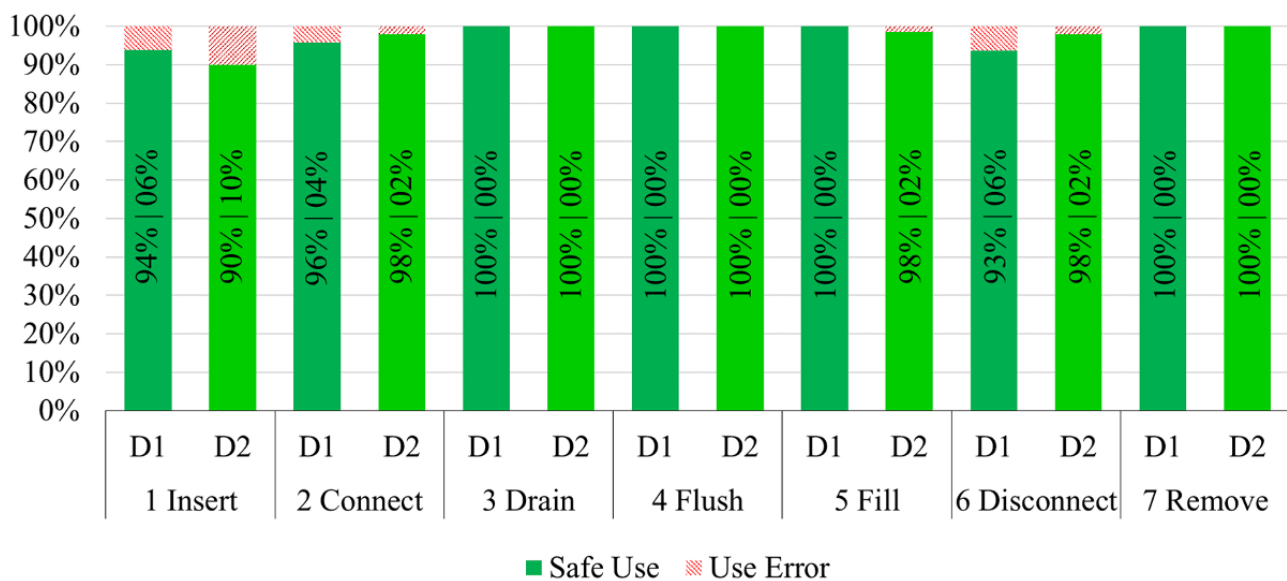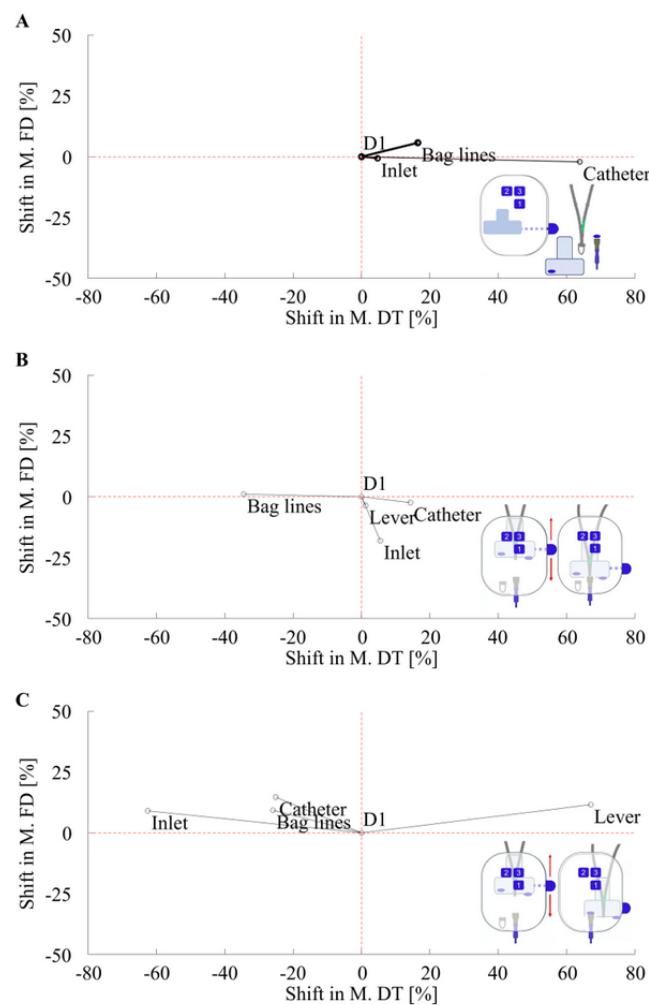
Figure 6. Shifts from D1 (in the coordinate origin) to D2 in terms of mean fixation duration (ordinate) and mean dwell time (abscissa) for task 1 insert (A), task 2 connect (B), and task 6 disconnect (C). The relevant user interface features in these three tasks are bag lines, inlet, catheter, and lever.

Discussion

Principal Findings

Task performance analysis generally showed little or no use errors in the various handling tasks for both UI designs (see Figure 5). The tasks with observed use errors were the insertion of material and connection and disconnection of bag lines and catheter. In line with the observations, participants described in semistructured interviews difficulties in the execution and in remembering of the correct handling step details in the observed use error–related tasks. Further, they reported misleading wording in the instructions as the explanation for their use errors in task 2 (connect) and task 6 (disconnect), thus providing additional information for the development of the supplementary material. In the first task, most use errors occurred when inserting the bag lines into the inlet. For this task, Figure 6A shows a category 1 pattern with longer mean DT and longer mean FD for the bag lines. Therefore, the results of gaze data analysis are consistent with the results of task performance. Gaze data shows more scrutinizing for D2 compared with D1 in order to insert bag lines and catheter into inlet and device. The longer and higher depth in visual perception indicates a higher mental workload for this task using D2.

When connecting and disconnecting the catheter, some participants missed pulling down the lever to connect bag lines and catheter again and putting a new cap on the catheter. For connecting and disconnecting bag lines and catheter, the most important interface features show category 4 patterns (task 2, Figure 6B) and category 2 patterns (task 6, Figure 6C). This visual pattern indicates more skimming behavior for task 2 and less skimming behavior for task 6. This in turn indicates more visual controls when connecting bag lines and catheter in task 2 for D2. Compared with task performance, this seems to result in slightly fewer use errors for D2 (2% vs 4%). For task 6, results indicate less visual searching associated with the relevant features inlet and catheter for D2 when a new cap is placed on the catheter. In a comparison of the two UI designs, the main difference between D1 and D2 is the position of the top window. With D2, the user can better see the inlet. This may help finding the important features while a new cap is placed on the catheter. Furthermore, the lever in task 6 shows a category 1 pattern associated with a longer and higher depth in visual perception for D2. Although the results show fewer use errors, handling the lever with D2 appears to be mentally more difficult than with D1.

When evaluating the total mental workload of the medical device system, the analyzed UI features of the medical device showed

shifts in both the mean FD and mean DT. The mean FD varied from 149 to 405 milliseconds in the critical tasks across all features (Multimedia Appendix 1). In order to be able to interpret these values, the results of three different task examples as described in the literature are compared. In a case study of a driving situation described by Velichkovsky et al [34], the values for the mean FD were between 499 and 543 milliseconds. Bojko et al [33] reported in an evaluation of drug label designs that the FD varied between 260 and 392 milliseconds. Just and Carpenter [35] observed a mean FD of 477 milliseconds observing the task of reading a scientific text. Compared with these studies, the mean FD of the handling cycle is in the same range as reading a drug label. The mean DT in the critical tasks varied in a range from 0.3 to 28.3 seconds (Multimedia Appendix 1). Especially in the first task, the insertion of the material in both UI designs required longer DT for bag lines and catheter compared with other tasks. This shows that this task requires special attention from the user. This is supported by significant differences in a MANOVA for the catheter in the considered task. The statistical analysis showed only in two other cases significant differences in the gaze data. The reason for merely three significant differences is probably because of the low level of variation in the design.

Based on the results of this study, benchmarking D1 and D2 showed the following. Inserting the material seemed to be challenging for both UI designs in general. Therefore, the guiding material (manual and quick starting guide) and training should focus on this task. The lever of D1 seemed to result in lower mental workload. It has a more dominant appearance compared with D2, where the lever is integrated into the housing for protection in case of a fall. The UI design D2 of the inlet seems to be easier to perceive visually. The higher position of the top window in D2 shows a positive impact on the task connecting and disconnecting bag lines and catheter.

Analysis of two dimensions of visual perception using eye tracking provided a detailed picture of the length and depth of the visual perception and therefore the challenges in user cognition and ease of use. Results highlighted the differences in information extraction for different UI features in single tasks. This information helped human factors engineering to focus the development on the critical UI features. Following this work, a summative study evaluated the final UI of the device. This final design and the instructions incorporated the results of this study, such as the detailed description of the insertion of the material and the coloring of the main UI features to guide the user's gaze. The summative study included patients, relatives, nurses, and physicians. They represented the later user population in the characteristic in age, preknowledge, and comorbidities. Patients had two types of comorbidities, such as arthritis and Reynaud syndrome, in addition to the renal disease with its own accompanying symptoms. The summative study confirmed the safety, efficiency, and effectiveness of use [40].

Limitations

Due to the novelty of the medical device presented in this study, there are several limitations regarding the results. First, participants were not patients in the real therapy. They were beginners who had no experience in this specific therapy or associated tasks. Furthermore, the device was not used in the real therapy application but in a simulation. These factors provide information on how forgetfulness or even dementia would influence use of the medical device in the later use by patients. Second, when the final product is used, individual training of the user is mandatory and labeling material supports the user. This support was not provided in this study. Instead, a presentation with an additional low-level representation of the UI and a neutral text of the seven tasks guided participants through the handling cycle. Consequently, the focus was on intuitive task performance and perception of information depending on the different UI designs. Third, the design of the two different top-level designs was similar due to a unicolored representation. This is not a strong contrast between the main UI functions and the rest of the medical device. As stated in Methods, this was chosen to eliminate influences of different coloring as an additional influencing factor. At the level of gaze data analysis, representation of the combination of mean FD and mean DT is the first published. Further research is needed to assess whether identified patterns apply to different usability studies with different tasks and stimuli.

Conclusion

The prototypes of the medical device system as stimuli of the study had only little differences in the single UI features. Consequently, results in the effectiveness of use revealed only marginal differences, with a maximum of 6% versus 10% use errors in task 1 (insert). Based on the two dimensions of the physiological gaze data measurements FD and DT, four distinct patterns could be distinguished between the two UI designs. A MANOVA revealed statistically significant differences in these patterns for three UI features.

Studying the impact on the usability of alternatives of different UI designs is crucial to understand which best supports the user. Traditional methods such as observation, interviews, or questionnaires tend to give feedback only at the level of the UI as a whole. Furthermore, when it comes to reporting usability issues or first impressions of the medical device during interviews or questionnaires, several challenges arise. Test participants may forget to report their impressions or adapt their answers to social expectations [30,41]. This makes it difficult to identify the root causes of usability problems and thus the necessary changes in UI design. In alignment with Lohmeyer et al [31] and Koester et al [30], this study showed that mobile eye tracking provides objective quantitative results based on physiological measurements related to individual UI features. These results can be used to evaluate usability in much more detail compared with traditional methods.

This information is crucial to be able to adapt the design of a product to the needs of the users. Therefore, results of usability testing must be more detailed than just a yes-or-no result of use errors. On the contrary, evaluation of each feature of the UI promises to achieve the best possible UI design by combining the best features found. This combined solution would therefore offer the highest level of usability. In this way, manufacturers can develop products that can be used even by untrained people without prior knowledge. This would allow home care to be provided not only by highly qualified nurses and caregivers,

but also by patients themselves, partners, children, or neighbors. This would contribute to removing barriers to home care and thus to a higher quality of life and normalization of everyday life, which is less dominated by illness for patients.

Acknowledgments

Conflicts of Interest

SN and JCG are employed by Peripal AG.

Multimedia Appendix 1

Analysis of eye tracking metrics for user interface features bag lines, inlet, catheter, lever, and buttons. Mean fixation duration (FD) in milliseconds and mean dwell time (DT) of the gaze in seconds for user interface designs D1 and D2. Evaluated tasks are task 1 (insert), task 2 (connect), and task 6 (disconnect). Multivariate analysis of variance analyzed the combination of FD and DT for significant differences according to the Pillai trace (p) between D1 and D2.
[DOCX File , 14 KB-Multimedia Appendix 1]

References

1. Rygh E, Arild E, Johnsen E, Rumpsfeld M. Choosing to live with home dialysis-patients' experiences and potential for telemedicine support: a qualitative study. BMC Nephrol 2012 Mar 19;13:13 [FREE Full text] [doi: 10.1186/1471-2369-13-13] [Medline: 22429705]
2. Brown EA, Johansson L, Farrington K, Gallagher H, Sensky T, Gordon F, et al. Broadening Options for Long-term Dialysis in the Elderly (BOLDE): differences in quality of life on peritoneal dialysis compared to haemodialysis for older patients. Nephrol Dial Transplant 2010 Nov;25(11):3755-3763 [FREE Full text] [doi: 10.1093/ndt/gfq212] [Medline: 20400451]
3. Alloatti S, Manes M, Paternoster G, Gaiter AM, Molino A, Rosati C. Peritoneal dialysis compared with hemodialysis in the treatment of end-stage renal disease. J Nephrol 2000;13(5):331-342. [Medline: 11063136]
4. Devoe DJ, Wong B, James MT, Ravani P, Oliver MJ, Barnieh L, et al. Patient education and peritoneal dialysis modality selection: a systematic review and meta-analysis. Am J Kidney Dis 2016 Sep;68(3):422-433. [doi: 10.1053/j.ajkd.2016.02.053] [Medline: 27125246]
5. ERA-EDTA Registry Annual Report 2016. Amsterdam: European Renal Association–European Dialysis and Transplant Association URL: https://www.era-edta-reg.org/files/annualreports/pdf/AnnRep2016.pdf [accessed 2020-05-05]
6. Wilson SE, Ouédraogo CT, Prince L, Ouédraogo A, Hess SY, Rouamba N, et al. Caregiver recognition of childhood diarrhea, care seeking behaviors and home treatment practices in rural Burkina Faso: a cross-sectional survey. PLoS One 2012;7(3):e33273 [FREE Full text] [doi: 10.1371/journal.pone.0033273] [Medline: 22428006]
7. Jones R, Jordan S. The implementation of crisis resolution home treatment teams in wales: results of the national survey 2007-2008. Open Nurs J 2010 Feb 18;4:9-19 [FREE Full text] [doi: 10.2174/1874434601004010009] [Medline: 20502646]
8. Dijkstra NE, Sino CGM, Heerdink ER, Schuurmans MJ. Development of eHOME, a mobile instrument for reporting, monitoring, and consulting drug-related problems in home care: human-centered design study. JMIR Hum Factors 2018 Mar 07;5(1):e10 [FREE Full text] [doi: 10.2196/humanfactors.8319] [Medline: 29514771]
9. Alnosayan N, Chatterjee S, Alluhaidan A, Lee E, Houston FL. Design and usability of a heart failure mhealth system: a pilot study. JMIR Hum Factors 2017 Mar 24;4(1):e9 [FREE Full text] [doi: 10.2196/humanfactors.6481] [Medline: 28341615]
10. Kantowitz BH. Attention and mental workload. 2000 Presented at: Proceedings of the 15th Triennial Congress IEA 2000/Human Factors and Ergonomics Society Annual Meeting; 2000; San Diego. [doi: 10.1177/154193120004402121]
11. Beatty AL, Magnusson SL, Fortney JC, Sayre GG, Whooley MA. VA FitHeart, a mobile app for cardiac rehabilitation: usability study. JMIR Hum Factors 2018 Jan 15;5(1):e3 [FREE Full text] [doi: 10.2196/humanfactors.8017] [Medline: 29335235]
12. Liang J, Xian D, Liu X, Fu J, Zhang X, Tang B, et al. Usability study of mainstream wearable fitness devices: feature analysis and system usability scale evaluation. JMIR Mhealth Uhealth 2018 Nov 08;6(11):e11066 [FREE Full text] [doi: 10.2196/11066] [Medline: 30409767]
13. George D, Hassali MA, Hss A. Usability testing of a mobile app to report medication errors anonymously: mixed-methods approach. JMIR Hum Factors 2018 Dec 21;5(4):e12232 [FREE Full text] [doi: 10.2196/12232] [Medline: 30578216]
14. Rotondi AJ, Spring MR, Hanusa BH, Eack SM, Haas GL. Designing eHealth applications to reduce cognitive effort for persons with severe mental illness: page complexity, navigation simplicity, and comprehensibility. JMIR Hum Factors 2017 Jan 05;4(1):e1 [FREE Full text] [doi: 10.2196/humanfactors.6221] [Medline: 28057610]

15. Medicines and Healthcare Products Regulatory Agency. 2017. Human factors and usability engineering—guidance for medical devices including drug-device combination products URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/645862/HumanFactors_Medical-Devices_v1.0.pdf [accessed 2020-05-04]

16. Rosenblum LD. See What I'm Saying: The Extraordinary Powers of our Five Senses. New York: WW Norton; 2010.

17. Miller GA. The magical number seven plus or minus two: some limits on our capacity for processing information. Psychol Rev 1956 Mar;63(2):81-97. [Medline: 13310704]

18. Peterson LR, Peterson MJ. Short-term retention of individual verbal items. J Exp Psychol 1959 Sep;58:193-198. [doi: 10.1037/h0049234] [Medline: 14432252]

19. Duc AH, Bays P, Husain M. Eye movements as a probe of attention. Prog Brain Res 2008;171:403-411. [doi: 10.1016/S0079-6123(08)00659-6] [Medline: 18718333]

20. Anderson J. Cognitive Psychology and its Implications. 6th Edition. New York: Worth; 2004.

21. Wu Y, Cheng J, Kang X. Study of smart watch interface usability evaluation based on eye-tracking. In: Marcus A, editor. Design, User Experience, and Usability: Technological Contexts, Vol. 9748. Springer International Publishing: Cham; 2016:98-109.

22. Groen M, Noyes J. Using eye tracking to evaluate usability of user interfaces: is it warranted? IFAC Proceedings Volumes 2010;43(13):489-493. [doi: 10.3182/20100831-4-FR-2021.00086]

23. Wedel M, Pieters R. A review of eye-tracking research in marketing. In: Review of Marketing Research, Vol. 4. Bingley: Emerald Group Publishing Limited; 2008:123-147.

24. Chynal P, Sobecki J, Szymanski J. Remote usability evaluation using eye tracking enhanced with intelligent data analysis. In: Hutchison D, Kanade T, Kittler J, Kleinberg JM, Mattern F, editors. Design, User Experience, and Usability. Design Philosophy, Methods, and Tools. Vol. 8012. Springer: Berlin; 2013:212-221.

25. Harezlak K, Kasprowski P. Application of eye tracking in medicine: a survey, research issues and challenges. Comput Med Imaging Graph 2018 Apr;65:176-190. [doi: 10.1016/j.compmedimag.2017.04.006] [Medline: 28606763]

26. Erol Barkana D, Açık A, Duru DG, Duru AD. Improvement of design of a surgical interface using an eye tracking device. Theor Biol Med Model 2014 May 7;11 Suppl 1:S4 [FREE Full text] [doi: 10.1186/1742-4682-11-S1-S4] [Medline: 25080176]

27. Guo Q, Xue C, Lin Y, Niu Y, Chen M. A study for human-machine interface design of spacecraft display & control device based on eye-tracking experiments. In: Harris D, editor. Engineering Psychology and Cognitive Ergonomics: Cognition and Design. Vol. 10276. Springer International Publishing: Cham; 2017:211-221.

28. Dill E, Young S. Analysis of eye-tracking data with regards to the complexity of flight deck information automation and management—inattentional blindness, system state awareness, and EFB usage. 2015 Presented at: 15th AIAA Aviation Technology, Integration, and Operations Conference; 2015; Dallas. [doi: 10.2514/6.2015-2901]

29. Obrist M, Bernhaupt R, Beck E, Tscheligi M. Focusing on elderly: an iTV usability evaluation study with eye-tracking. In: Cesar P, Chorianopoulos K, Jensen JF, editors. Interactive TV: A Shared Experience. Vol. 4471. Springer: Berlin; 2007:66-75.

30. Koester T, Brøsted JE, Jakobsen JJ, Malmros HP, Andreasen NK. The use of eye-tracking in usability testing of medical devices. Proc Int Symp Human Factors Ergonom Health Care 2017 May 15;6(1):192-199. [doi: 10.1177/2327857917061042]

31. Lohmeyer Q, Schneider A, Jordi C, Lange J, Meboldt M. Toward a new age of patient centricity? The application of eye-tracking to the development of connected self-injection systems. Expert Opin Drug Deliv 2019 Feb;16(2):163-175. [doi: 10.1080/17425247.2019.1563070] [Medline: 30577710]

32. Godfroid A, Spino LA. Reconceptualizing reactivity of think-alouds and eye tracking: absence of evidence is not evidence of absence. Language Learning 2015 Oct 08;65(4):896-928. [doi: 10.1111/lang.12136]

33. Bojko A, Gaddy C, Lew G, Quinn A, Israelski E. Evaluation of drug label designs using eye tracking. Proc Human Factors Ergonom Soc 2016 Nov 05;49(11):1033-1037. [doi: 10.1177/154193120504901107]

34. Velichkovsky B, Dornhoefer S, Pannasch S, Unema P. Visual fixations and level of attentional processing. 2000 Presented at: Proceedings of the symposium on eye tracking research & applications - ETRA '00; 2000; Palm Beach Gardens p. 79-85. [doi: 10.1145/355017.355029]

35. Just MA, Carpenter PA. A theory of reading: from eye fixations to comprehension. Psychol Rev 1980 Jul;87(4):329-354. [Medline: 7413885]

36. Kim JH, Zhao X, Du W. Assessing the performance of visual identification tasks using time window-based eye inter-fixation duration. Int J Industr Ergonom 2018 Mar;64:15-22. [doi: 10.1016/j.ergon.2017.09.002]

37. Sarter NB, Mumaw RJ, Wickens CD. Pilots' monitoring strategies and performance on automated flight decks: an empirical study combining behavioral and eye-tracking data. Hum Factors 2007 Jun;49(3):347-357. [doi: 10.1518/001872007X196685] [Medline: 17552302]

38. Bojko A. Eye Tracking the User Experience: A Practical Guide to Research. New York: Rosenfeld Media; 2013.

39. International Electrotechnical Commission. IEC 62366-1:2015. Medical devices—Part 1: application of usability engineering to medical devices. 2015. URL: https://www.iso.org/standard/63179.html [accessed 2020-05-04]

40. Hess S, Dubach M, Meboldt M, Foggensteiner L. Evaluating patient safety and ease of use of a novel connection-assist device for peritoneal dialysis. Patient Prefer Adherence 2019;13:1785-1790 [FREE Full text] [doi: 10.2147/PPA.S218663] [Medline: 31754299]

XSL·FO
RenderX

41.   Wiklund M, Kendler J, Strochlic A. Usability Testing of Medical Devices. Boca Raton: CRC Press; 2016.

Abbreviations

DT: dwell time
FD: fixation duration
MANOVA: multivariate analysis of variance
PD: peritoneal dialysis
UI: user interface