Digital video analysis of health professionals’ interactions with an electronic whiteboard: A longitudinal, naturalistic study of changes to user interactions

Rasmus Rasmussen, Andre Kushniruk

Computer Science, Department of Communication, Business and Information Technologies, Roskilde University, Denmark

School of Health Information Science, University of Victoria, British Columbia, Canada

ABSTRACT

As hospital departments continue to introduce electronic whiteboards in real clinical settings a range of human factor issues have emerged and it has become clear that there is a need for improved methods for designing and testing these systems. In this study, we employed a longitudinal and naturalistic method in the usability evaluation of an electronic whiteboard system. The goal of the evaluation was to explore the extent to which usability issues experienced by users change as they gain more experience with the system. In addition, the paper explores the use of a new approach to collection and analysis of continuous digital video recordings of naturalistic “live” user interactions. The method developed and employed in the study included recording the users’ interactions with system during actual use using screen-capturing software and analyzing these recordings for usability issues. In this paper we describe and discuss both the method and the results of the evaluation. We found that the electronic whiteboard system contains system-related usability issues that did not change over time as the clinicians collectively gained more experience with the system. Furthermore, we also found user-related issues that seemed to change as the users gained more experience and we discuss the underlying reasons for these changes. We also found that the method used in the study has certain advantages over traditional usability evaluation methods, including the ability to collect analyze live user data over time. However, challenges and drawbacks to using the method (including the time taken for analysis and logistical issues in doing live recordings) should be considered before utilizing a similar approach. In conclusion we summarize our findings and call for an increased focus on longitudinal and naturalistic evaluations of health information systems and encourage others to apply and refine the method utilized in this study.

1. Introduction

As hospitals worldwide move towards increased automation, a wide variety of information systems are becoming deployed in settings such as emergency departments. One such type of information system, the electronic whiteboard (EW), is being increasingly deployed as a replacement for the ubiquitous dry-erase whiteboards traditionally used as the central coordination and communication tool in emergency departments and hospital departments in general [1–5]. Previous research has shown that EW systems have certain advantages over the traditional dry-erase whiteboards. These advantages include distributed access to and updating of patient information, the ability to store and retrieve information for future use, integration with other clinical information systems, live tracking of patients and decreased mental workload for clinicians [1,2,6,7].

However, as we move towards greater use of such health information technologies in real clinical settings, a wide range of human factor issues have emerged and it has become clear that there is a need for improved methods for designing and testing EWs that are to be integrated into complex work practices in settings such as the emergency department (ED) [8]. For example, Wong et al. [9] found the need for a number of enhancements to an EW after it was deployed as well as the need to conduct workflow review meetings to ensure adoption. A comparative study of manual whiteboards and EW systems [3] used interviews and observations of clinicians using a EW to identify issues related to the need for flexibility, the need for local customization by clinicians and the incorporation of new and emerging needs into EWs. In another line of research Riley et al. [10] have demonstrated how implementation of EWs can lead to inadvertent changes in power, work activities and professional control in clinical practice.
In this paper we describe our work in conducting naturalistic video-based analyses of user interactions with a new EW system that has been deployed in two emergency departments at Danish hospitals. The goals of the study are twofold. The first to goal is to identify specific usability problems, potential inefficiencies and workflow issues associated with the clinicians’ actual daily use of the EW system. In relation to this, we are also interested in understanding if there would be differences in human factor and workflow issues in departments that have adapted to the same EW sometime after deployment as compared to a more recent deployment. The second goal of the study is to determine the practicality of the approach we have developed and used for evaluating the EW system in this study. This approach was designed to augment and complement an initial participatory design approach described by Rasmussen et al. [11] and Fleron et al. [12].

Methods which have been previously used to evaluate the use of EWs have included approaches ranging from surveys and interviews given to EW users [10] to observation of users [13] to collecting and analyzing static digital photographs taken of EW screens [2,3]. While these methods are highly effective for capturing and documenting the work practices surrounding an EW system, e.g. group interactions or viewing the display from a distance, we find that there are some limitations to such methods. Among others, these limitations include difficulty in collecting long-term data regarding how the system responds to a wide range of user interactions in real settings. Also, we find that some of the methods previously used are somewhat obtrusive in nature and are thereby at risk of biasing the final results due to the Hawthorne effect [14–16].

Due to the goals of this study and the described limitations of the methods used previously in studies of EW system, we therefore chose to employ a method whereby user interactions with the EW system were recorded in a continuous, naturalistic and unobtrusive manner. By continuously recording live user interactions over time with applications such as EWs, a large and rich data set can be collected that can be used to assess usability problems and help describe adoption issues when such applications are deployed in real clinical and emergency settings over time. Our previous work has shown that although laboratory testing of healthcare applications applying usability methods is needed, it is not sufficient for ensuring the safety and effectiveness of healthcare applications deployed in complex settings [17]. In the area of video analysis of user interactions, previous work has been published about video and screen recording and resultant analysis of healthcare professional interactions in the context of usability testing [18] and the extension of usability testing to more realistic simulations, termed “clinical simulations” [17]. However, there has been less work describing effective approaches to the naturalistic analysis of video recorded user interactions’ with systems and patients once a system has been deployed in a real clinical setting. To address this issue, we describe, alongside the results of the performed evaluation, the approach we have developed and applied to collecting and analyzing large data sets of continuous live screen recordings of user interactions with an EW system over time.

2. Methods

We conducted a qualitative longitudinal and naturalistic study of the ED clinicians’ use of the EW system following the procedure described later in this section. The healthcare region and ED management approved the study prior to it being conducted. Since the recordings would contain personal data regarding patients at the ED the study was also registered and approved by the Danish Data Protection Agency and followed the guidelines outlined in their directive. Clinicians on duty during recording sessions were briefed on the study during morning meetings and throughout the days if there were any questions or concerns.

2.1. Setting

The two emergency departments (ED1 & ED2) where this study was carried out are both relatively newly established departments at two larger hospitals in the Danish healthcare region of Zealand. Both departments participated in a larger project of developing, implementing and evaluating the EWs during the time of the study. The two departments are similar to each other in terms of their organizational structure and the tasks that they perform in the hospital. Both departments have nurses and a number of chief physicians employed directly and have resident physicians attached on an on-call basis. On average there are 20 nurses, 6 physicians and 5 medical secretaries on duty during a normal day shift at ED1 and for ED2 the numbers are 16, 9 and 5 respectively.

In regards to this study the two departments differ in terms of how much experience each department as a whole has had with the EW. Data from the first emergency department (ED1) was collected approximately 1.5 years after the EW had been installed in this department. We will henceforth refer to this department as ED1LATE. At the second emergency department (ED2) the EW system was installed more recently. We collected data at this department on two occasions: (a) approximately 1.5 months after the EW was installed and (b) approximately 5.5 months after the EW was installed. We will henceforth refer to ED2 as either ED2EARLY or ED2MID depending on the point in time referred.

2.2. The electronic whiteboard

The EW system analyzed in this study is a web-based application installed on a central server. The ED clinicians use the information stored and displayed by the EW system for internal coordination of work tasks and communication between staff members. The information displayed by the EW system includes the patients’ medical problems, triage levels, lab results, plans for further treatment and what department the patients will be transferred to in case they need hospitalization. Besides the above information the EW system also displays the patients’ first names, their age and their location in the department as well as the name of the nurse and physician currently responsible for attending to a specific patient. Finally, the EW system has the possibility of automatically retrieving and displaying information from other clinical IT systems, e.g. laboratory, radiology and patient monitoring systems. This option was implemented at ED1LATE.

Fig. 1 shows a screenshot of the EW system as it is configured at ED1LATE. While the EW system can be configured individually to match unique work practices at different departments the main functionality and purpose of the system is the same at both departments involved in this study.

The system is accessible from all devices connected to the same network as the central server, which affords flexible and distributed access to the system. Clinicians can thereby access the EW system from multiple access points throughout the department, e.g. workstations, laptops and mobile devices. The information displayed by the EW system is stored in a central underlying database and is thereby shared by all access points. However, in order to ease the clinicians’ interactions with the stored information it is possible to create specific filters or views in the system according to different parameters of the information. Both EDs involved in this study had defined specific views according to the possible locations of the patients, e.g. waiting room, the fast track section or other sections of the departments. In other words, a specific view corresponds to a specific section of the department or another
location within the department and the patients shown in this view are indicated to currently be located in this section or location. The interface buttons for changing between views are indicated by the red rectangle in Fig. 1.

The main access points and most prominent artifacts of the system are the 52-in. touch sensitive wide-screen displays located in central command rooms throughout the departments. Clinicians from all professions use these displays for updating, retrieving and discussing patient information. Other important access points to the EW system include the workstations used by the secretaries, the triage nurses and coordinating nurses. From these access points the secretaries, the coordinating nurses and triage nurses enter new patients as they are reported to the department and distribute them between the different sections of the departments as they arrive. During a normal day shift these access points have only one or two primary users while the wide-screen displays may have multiple users accessing the system.

Access to the EW through the wide-screen access points is protected by a login mechanism whereby users identify themselves and unlock the system by scanning a personal chip. Login at the personal access points, e.g. triage nurses workstation and other personal work stations is handled by the work station login mechanism and users do not need to specifically login to the EW system at these access points.

2.3. Materials

The materials used for data collection in this study were selected for their inexpensiveness in order to assess the usefulness of a low cost approach to collecting live user data. The main components included three 4 Gb flash drives with the free version of the screen-capturing software HyperCam installed as well as three 2 Tb external hard drives and three 16 Gb flash drives for storage of resultant digital screen recordings. To support the analysis of the digital video files the free f5 coding tool was used in order to allow timestamping and annotating of the digital videos.

2.4. Procedure

User interactions with the EWs were captured using the screen-capturing software installed on flash drives as described above. The resulting digital movie files were stored on either an external hard drive or one of three 16 Gb flash drives. Using this setup there was no need for installing any software on the machines where the recordings were performed and no need to take up local storage space during the recordings.

User interactions with the EW were captured over a period of 5 days between 10 AM and 4 PM each day. This period was specifically chosen because experience showed that this was often the busiest time of day at the two EDs and therefore should produce the highest number of interactions with the EW system. The recordings were performed at three different access points in the two EDs. At ED1 the coordinating and triage nurses have their own workstations in separate locations in the department. Because these access points to the system were deemed to be regularly and frequently used recordings were performed here each day in the 5-day session. Interactions with the system through the wide-screen displays rotated between two of the command rooms in the department. At ED2 the same person holds the role of coordinating and triage nurse. At this department the coordinating/triage nurse shares two workstations with a secretary so recordings rotated between these two machines. Recordings of interactions with the system through the wide-screen displays rotated between the different command rooms of the department.

To capture how user interactions with the system changed over time recordings were performed initially at ED1LATE, then at ED2EARLY and 5 months later at ED2MID following the procedure described above.

2.5. Data analysis

Initially, each digital video file was viewed and logged by the first-author using a predefined taxonomy of on-screen activities developed by the authors – see Fig. 2. Using this taxonomy on-screen activity was recorded and entered via a word-processor as entries in separate log files. In this context a log entry is thus defined as any on-screen activity recorded by the screen-capturing software. This includes user initiated activity as well as system initiated activity, e.g. error messages presented to the user. As a minimum each entry in the log files contains an auto-generated timestamp (corresponding to system-user interactions observed on the digital video of the EW screens), an activity indicator, a
name for the activity and a description of the activity. In instances of user initiated activity the entries also contain an indicator of whether or not the task was completed and the number of steps used in completing the task or before aborting the task. When usability issues were discovered we marked the entries with an indicator and coded the issues using the intentionally broad coding categories found in Table 1. We also provided a description of the issue including whether or not the user solves the issue. Following the logging process each log file was perused and any activities of interest were coded for further analysis. In coding the log files we also looked for relationships between the different entries, e.g. the different entries that make up specific work patterns.

We then summarized the total number of entries and average number of entries for each recording session at each access point. For each usability issue, we tabulated the total number of times the issue occurred in total and calculated the number of times each issue occurred per hour. To control for variations in the total number of activities logged during each recording session we also calculated the percent-wise distribution of each issue relative to the total number of activities recorded at each type of access point for ED2EARLY, ED2MID, and ED1LATE respectively.

The initial logging of the digital video files was performed solely by the first-author due to restrictions imposed by the Danish Data Protection Agency’s directive. After having logged the digital video files both authors were involved in the coding of the log files and the following analysis work. All codings were discussed amongst the authors to mitigate biases in the analysis and any differences in coding were resolved through discussion.

3. Results

The result of the recordings was 166.5 h of video data (i.e. continuous digital video recordings of all recorded user interactions with the EW) divided between 45 digital video files each lasting an average of 3 h and 41 min. No audio was recorded due to the lengthy continuous nature of the recordings and the clinicians’ concern for disclosing sensitive information regarding patients and colleagues. Also, due to the applied method’s reliance on using screen-capturing software there were no recordings of activities that did not occur on-screen, e.g. users viewing the EW display without interacting with the system or the usage of paper-based artifacts. As such each video file contained only the continuous screen recordings of the clinicians’ on-screen interactions with the EW system. Also, even though the video files clearly showed the full screen of the EW displays, the mouse cursor and user interaction, for the purpose of this study we did not log cursor movements, count keystrokes or mouse clicks (even though this would be possible during the analysis of the digital video files).

Using the taxonomy described in Section 2, a log corresponding to each digital video file was created in a text file for further analysis. The process of reviewing the video and creating the logs files took approximately 200 h including roughly 20 h of analysis. A total of 2863 entries were logged producing an average of 64 entries per log file. Table 2 shows basic data for the recordings including total number of entries and average number of entries logged at each type of access point at the three EDs. Due to the naturalistic approach used in this study the results are naturally influenced by the work practices and EW usage patterns that exist within the EDs where the recordings were performed, e.g. fluctuating periods of low or high usage. Also, due to technical difficulties a few of the recordings were ended prematurely resulting in shorter periods of screen recording than had been planned. The resulting log files are therefore not directly comparable in terms of number of entries – see Table 2. To counter for differences in the number of total entries per recorded sessions, rather than directly comparing

![Fig. 2. The taxonomy used for logging on-screen activities.](image-url)
frequencies of usability issues from the different sessions, we calculate the percentage-wise distribution of the identified usability issues relative to the total number of log-entries per session whenever comparing distribution of usability issues across the three recording sessions.

3.1. Log entries

Fig. 3 shows a snippet containing an entry from one of the log files. The entry reveals the timestamp for the activity recorded and then annotates the activity using the “Name” field of the taxonomy shown in Fig. 2. In this case the activity has been identified as “Updating patient information” and the description details the interactions included in this activity. After this the entry indicates that the activity was completed in five steps and that there was a usability issue during the activity. An identifier and a description of the issue detected are included as well. Table 3 provides similar examples of entries for each of the four categories of usability issues. Table 4 provides an overview of the unique issues revealed during analysis of the video files, the total number of times each specific issue occurred and a description of each issue. In some cases we grouped entries under the same issue. This was done in cases where issues were found to be of the same nature, e.g. different types of system errors.

3.2. Work patterns

As seen in Table 4 the coded work patterns are defined by a number of individual log entries directly related to each other through the task which they form part of. An example of such a work pattern is the task of updating and transferring patient information from one view of the EW to another, i.e. updating

![Fig. 3. A snippet of one of the log files.](image-url)
Table 4
The unique usability issues, their related categories and the total number of times each issue occurred.

<table>
<thead>
<tr>
<th>Coding category</th>
<th>Identifier</th>
<th>Description</th>
<th>Total number of instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>System bugs</td>
<td>Other system bugs</td>
<td>Includes disappearing patients, multiple patients, clinicians not appearing in list, information not being updated, system crashes and system opens wrong field.</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Floating pop-ups</td>
<td>Pop-ups related to one field on the EW display “wanders” around on the display.</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Jumping cursor</td>
<td>When writing in one text field within the EW system the cursor jumps spontaneously from one text field to another</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>System allows update of moved patients</td>
<td>Even though a patient has been transferred to another view the user can sometimes start an update of that patient which results in either the wrong patient being updated or an error message.</td>
<td>8</td>
</tr>
<tr>
<td>Efficiency problem</td>
<td>Complicated and long pathways</td>
<td>Some pathways to solving different tasks with the EW system require many steps to complete</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Other efficiency problems</td>
<td>Issues that occur less than five times in total or at only one location. Includes issues regarding user mistakes, unsuccessful user actions, menu items not available, not enough information provided.</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>Incorrect use of interface elements</td>
<td>The clinicians use the interface elements in an incorrect manner, e.g. trying to search with free text field or entering identical information in multiple fields.</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Menu item confusion</td>
<td>Indications that the users have trouble finding menu items when using the EW.</td>
<td>11</td>
</tr>
<tr>
<td>Error message</td>
<td>Newer information available</td>
<td>Updating information on the EW sometimes causes the system to display a message saying that newer information is available and that the current update will be cancelled.</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Browser errors</td>
<td>Generic browser errors caused by the EW application.</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Other error messages</td>
<td>Issues that occur less than five times in total or at only one location. Includes messages regarding authentication errors, duplicate rows and unavailable resources.</td>
<td>5</td>
</tr>
<tr>
<td>Work patterns</td>
<td>Login to system before interaction</td>
<td>Access to the EW through the wide screens is only possible if the clinicians login</td>
<td>179/184</td>
</tr>
<tr>
<td></td>
<td>Updating and transferring patients</td>
<td>When updating patient information and afterwards transferring the patient to another EW view the transfer should not be done until all other information has been updated</td>
<td>62/45</td>
</tr>
</tbody>
</table>

the location information for a patient corresponding to the patient being moved physically within the department. This task consists of a number of steps where the user initially updates general patient information, e.g. time of arrival, triage level, medical problem, etc. After updating this information the user updates the patient’s location to indicate that the patient is being moved to another part of the department, e.g. from arrival to a patient room in the department. This transfers the patient and related information from one filter or view of the whiteboard to another.

While coding the log files we found that, for the task of updating and transferring patient information, the users would in some instances follow an efficient work pattern corresponding to the sequence of steps described above to complete this task – see Fig. 4 for a graphical illustration. However, in other situations we observed that some users would follow what appeared to be a less efficient work pattern. For the task of updating and transferring patient information the less efficient work pattern involved updating the location information before having updated the remaining patient information. This causes the system to move the information to the indicated view of the EW system before the user has finished the complete update, which forces the user to switch to another view of the EW and restart the information update. Fig. 5 shows how the inefficient work pattern adds two additional steps to the sequence of steps needed to complete the update and transfer patient information task.

Another example of a work pattern, which the users performed in both an efficient and inefficient manner, was whether or not they remembered to log in to the EW system before initiating activities that included updating patient information. Following the efficient work pattern the users would initially log in to the system, allow the system to unlock and then start performing the intended work task. Fig. 6 shows a graphical illustration of this work pattern. However, in other situations the users would forget to log in and simply initiate the intended activity. This would lead the system to display an error message asking the user to log in before changes were allowed and thereby force the users to reinitiate the intended activity after having logged in. Fig. 7 shows how this work pattern adds two steps to the sequence of steps needed to initiate an activity with the EW system.

In our analysis of the found work patterns we were interested in uncovering how the users efficiency with the EW system changed as the staff members of the EDs collectively gained more experience with the system. We therefore noted all instances of efficient and inefficient work patterns for the both the update
and transfer patient information and log in tasks. In the following we describe how we compared efficient and inefficient work patterns to reveal how increasing departmental experience affected the users’ work patterns with the EW system.

3.3. Comparing issue distribution

In order to determine how time since implementation had affected the usability issues described above, we compared the percentage-wise distribution of the different issues relative to the total number of log-entries per session. As indicated by the sparkline illustrations in Table 5, the identified usability issues generally occur infrequently during use of the EW system, which, as will be discussed, points to the need to perform longer continuous recording of interactions to identify problems that might be important to identify but occur infrequently. Also, a number of the issues occur at only one type of access point. This is either due to the nature of the access points or the tasks performed at the access point. Furthermore, Table 5 shows that the system-related issues, e.g. system bugs, error messages and system-related efficiency issues, were not affected by increasing departmental experience with the system.

The results show that these issues occur unsystematically and there does not appear to be any pattern in how often they occur across the three recordings sessions. In contrast to this, the sparkline illustrations in Table 5 indicate that a number of the user-related issues do seem to be affected by increasing departmental experience with the EW system. This is especially true for the work patterns identified in the analysis and described above. The complicated and long pathways issues also seem to be affected.

![Fig. 5. The inefficient work pattern for the update and transfer patient information task.](image1)

![Fig. 6. The efficient work pattern for logging into the EW system.](image2)

![Fig. 7. The inefficient work pattern for logging into the EW system.](image3)
by increasing departmental experience with the EW system. However, we touch upon these results elsewhere [19] and will therefore not discuss them here.

In order to uncover how often the efficient patterns were followed compared to the inefficient patterns, we calculated the percent-wise distribution for both the efficient pattern and inefficient patterns relative to the total number of work pattern instances recorded across the different types of access points. As indicated by the sparkline illustrations in Table 5, our results show a tendency for the users of the EW system to follow the efficient login work pattern less as the combined departmental experience increases. In our results the users of the EW system at ED2_{EARLY} followed the efficient login work pattern in 65% percent of all instances where a user logs in and uses the system. For ED2_{MID} this number dropped to 60.92% and at ED1_{LATE} the same number is 9.20%. This indicates that the experienced users at ED1 have a tendency to be less efficient than the less experienced users at ED2 with regards to this work pattern. Also, this indicates that the users at ED2 have a tendency to forget to login more often as they gain more experience with the EW.

Table 5
Sparkline illustrations for the percent-wise distribution of the found issues relative to the total number of log-entries per session.

<table>
<thead>
<tr>
<th>Coding category</th>
<th>Issue identifier</th>
<th>Access point type</th>
<th>Wide screen</th>
<th>ED2_{EARLY}</th>
<th>ED2_{MID}</th>
<th>ED1_{LATE}</th>
<th>Work stations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other system bugs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Floating pop-ups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jumping cursor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System allows update of moved patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complicated and long pathways</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency problem</td>
<td>Other efficiency problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incorrect use of interface elements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Menu item confusion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Newer information available</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error message</td>
<td>Browser errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other error errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Login to system before interaction (efficient)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Login to system before interaction (inefficient)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6
Sparkline illustrations of the percent-wise distribution of efficient and inefficient work patterns relative to the total number of work patterns instances.

<table>
<thead>
<tr>
<th>Work pattern</th>
<th>ED2_{EARLY}</th>
<th>ED2_{MID}</th>
<th>ED1_{LATE}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Login to system before interaction (Efficient)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Login to system before interaction (Inefficient)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Updating and transferring patients (Efficient)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Updating and transferring patients (Inefficient)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For the update and transfer work pattern the situation is reversed. As indicated by the sparklines in Table 6, our results show a tendency for the users of the EW system to follow the more efficient work pattern for this task as the department experience increased. In our results the users at ED2EARLY would follow the efficient work pattern in 35% of all instances while at ED2MID the users would follow the efficient work pattern in 47.46% of all instances of this work pattern. And at ED1LATE the users would follow the efficient work pattern in 96.43% of all instances of this work pattern. These results indicate that the experienced users at ED1 have a tendency to be more efficient than the less experienced users at ED2 with regards to this work pattern. However, the results also show that the users at ED2 seem to become more efficient with increased experience with the EW system. In the following we will discuss these findings and related them to previous and similar work.

4. Discussion

The research described in this paper has aimed to explore the extent to which usability issues experienced by users change as they gain more experience with the electronic whiteboard system. In addition, the paper explored the use of a new approach to collection and analysis of continuous digital video recordings of naturalistic “live” user interactions. As will be discussed below, using the method applied it was found that many of the usability issues identified did not change (e.g. lessen) over time, but rather continued even as their users gained more experience over time. This has important implications for considering the need for identifying usability problems prior to widespread system release, as training and experience may not mitigate serious usability problems. Along these lines (as will be discussed below) the results also indicate a need to conduct new forms of data collection and analysis are needed in order to ensure that otherwise long-lasting usability issues (that if not identified will continue after release) are detected and mitigated prior to widespread release. As our study has indicated this could include the use of continuous recordings of live user interactions with systems such as electronic whiteboards. This is especially the case with complex healthcare information systems, as the whiteboard described in this paper had usability issues and associated inefficient work patterns that were not detected prior to release using conventional testing methods, but that were identified through continuous live recording conducted in the study. Furthermore, it is expected that the methodological approach could be used for evaluating and optimizing other types of healthcare information systems (e.g. electronic health records, computerized provider order entry systems and other emerging information technologies).

4.1. Usability issues and work pattern changes

Through our analysis of user interactions with the EW system we found a range of different categories of usability issues and within each of these categories we found a number of specific issues. A number of the issues found were related to the design and technical implementation of the EW system, e.g. system bugs, error message and system-related efficiency issues. Other issues were related to the users of the EW system, e.g. inefficient work patterns and user-related efficiency issues. Although the usability issues did not occur frequently it is interesting to note that the same issues occurred at both EDs and at all three intervals since implementation, i.e. ED2EARLY, ED2MID and ED1LATE. As Table 5 indicates the system-related usability issues found do not appear to change over time as the users of the EW system gain more experience with the system. Kjeldskov et al. [20] found the same pattern in their longitudinal usability study of an electronic patient record (EPR) system. In this study the authors found that as the users of the EPR system gain more experience with the system the usability issues they face are similar to problems they uncovered as novices both in terms of type and severity.

However, from Table 6 we find that the work patterns uncovered in the analysis do indeed seem to be affected by the users gaining more experience with the EW system. In the following we will discuss these changes to work patterns and relate our findings to similar studies. Also, we will discuss the methodology used in this study including the limitations and challenges associated with using this method.

In our results we found work patterns that seem to be affected by how much time has passed since the system was implemented at the ED. One example of such a work pattern is related to the task of updating and transferring patient information from one view of the EW to another. In our results we found that the clinicians at ED2EARLY would follow an efficient work pattern in only 35% of the time when completing this task. At ED2MID this number had increased to 47.46% and for ED1LATE the same number was 96.43%. This indicates that there is a tendency in our results for the clinicians to become more efficient with increasing departmental experience with the EW system.

In their study Kjeldskov et al. [20] report a similar trend. Even though this study finds that users do not become significantly more efficient in completing tasks as they go from novices to experts the authors note that experts were faster in simple data entry tasks, e.g. typing in patient values [19]. These tasks are similar to the tasks that compose the work pattern of updating and transferring patient data in the EW system. Vaughan and Dillon [21] report on a similar trend in their longitudinal study of readers of a web-based newspaper. Here, the authors find that with time the newspaper readers become increasingly efficient as they gain more experience with the different versions of the newspaper involved in the study. This is positive because it indicates that despite having issues in the initial stages of usage the users of the EW system have the ability to overcome these issues and learn to use the system in an efficient manner.

Interestingly, our results also show that in some cases the users of the EW system became less efficient with increasing departmental experience. An example of this is the work pattern of logging in to the EW system before starting an update of the displayed information. In this instance we found that the clinicians at ED2EARLY and ED2MID were more efficient than the clinicians at ED1LATE despite the department having had the EW system implemented more recently. As [22] finds it is crucial that the login mechanism for any healthcare information system is well designed to fit into the “highly nomadic, dynamic, interrupted, and cooperative work in hospitals” [22]. The login mechanism for the EW system resembles the silent login mechanism that [22] presents as one possible login mechanism for systems similar to the EW making the process of logging in to the system relatively easy. Also, a previous study of the EW system has shown that the chip reader login mechanism does not have a negative impact on the systems usability [23]. Therefore, it is interesting to note that our results show a tendency for the clinicians to forget to login more frequently the more experience they gain with the EW system. A possible explanation of this trend could lie in the novelty value of the system at the two EDs and how the procedure of logging to the system fits into the clinicians’ perception of their work tasks with the EW.

As witnessed by the high number of login interactions, EW usage is characterized by frequent and rapid interactions see – see Table 4. This is distinct for the nomadic work practices commonly found in hospital departments [22]. In a situation where nomadic work practices influence and characterize the way the clinician’s use the EW the concept of logging on might not fit well.
with their everyday work with the EW system. In other words the idea of first logging into the system might not be present in the clinician's minds when using the system to carry out their tasks and this would not change as the clinician's gain more experience with the system.

Furthermore, as the clinicians move further away from the point in time when the system was implemented the novelty of the system and thereby the explicit recollection of how to use the system might wear off. As [24] report in an analysis of users' ability to learn and retain the functionalities of a web-based information retrieval interface, novice users were found to have the ability to learn how to use an interface relatively easily through hands-on training. However, it was also found that the test subjects were unable to retain what they had learnt in a second test session 4 weeks later and performed worse than during the first session [24]. In their discussion [24] argue that with time the users forget some of the functionalities of the interface. Following this line of thought we argue that because the EW system and associated training is fresher in the minds of the novice users at ED2EARLY and ED2MID they tend to follow the efficient login work pattern and ED2MID patients were unable to retain what they had learnt in a second test session 4 weeks later and performed worse than during the first session [24]. In their discussion [24] argue that with time the users forget some of the functionalities of the interface. Following this line of thought we argue that because the EW system and associated training is fresher in the minds of the novice users at ED2EARLY and ED2MID they tend to follow the efficient login work pattern more often than the experienced users of ED1LATE, because they are still aware of the system and better recall their initial training. Also, our results show that problems with login mechanisms in healthcare information systems is a recurring issue that still needs to be researched and refined in order to not be a hindrance to the efficient use of systems such as the EW studied here.

4.2. Discussion of methodology

The aim of the methodology applied in this study is very similar to those of other usability evaluation methods, i.e. uncovering usability issues in the design of a given system. However, the method applied here differs from many of the more traditional usability evaluation methods due to its naturalistic and longitudinal approach. Where usability evaluations involving think-aloud protocol analysis [25,18], cognitive walkthroughs [26] or clinical simulations [17] are capable of uncovering usability issues in a relatively short time frame and under controlled conditions, the method employed here has the capability of uncovering issues that occur on a longer time scale in a naturalistic setting. This includes potentially important or serious issues that occur infrequently or only under specific condition and are therefore more likely to be detected over lengthier periods of recording user interactions. In this regard the applied method is somewhat similar to the method applied by Zheng et al. [27] in their longitudinal evaluation of an electronic health records system and provides to a certain extent the same type of results. However, the method applied here differs from [27] due to its focus on qualitatively analyzing user interactions captured on digital video instead of the quantitative analysis approach described by Zheng et al. [27].

Thus, the naturalistic nature of the method applied in this study allows researchers to uncover patterns between issues that might not have been found through other types of usability evaluations and so can be considered a method which is complementary to other approaches. Another advantage of this method is the ability to uncover long-term changes to work patterns, which could be of value to future improvements to a given system. Furthermore, the method used is relatively unobtrusive for the organization where the evaluation is carried out when compared to other naturalistic evaluation methods, e.g. in situ interviews and observations. The advantage of this is that the results of the evaluation will be less biased by the evaluators presence and thereby forego the Hawthorne effect and say-do problematic of other evaluation methods. The unobtrusive nature of the method also becomes important when conducting naturalistic evaluations in a working environment where interruptions might have a negative impact on the work being carried out, e.g. interrupting patient care in EDs.

In addition, the cost for carrying out the study was minimized by using free screen-capturing software and digital video annotation software. Overall we feel the approach could serve as an important complement to traditional usability testing methods, and that it can be carried out in a cost-effective manner. Furthermore, it will uncover errors that may not be detected from short-term rapid usability testing prior to release (i.e. issues that may only be identified over longer periods of time under real conditions of system use).

However, the method applied in the study is not without limitations. In this study we were unable to record any audio while recording interactions. Recordings of clinicians thinking-aloud while using the system during this study might have captured some utterances of relevance for the analysis of the data, however this was not practically possible for longer-term recording of user interactions in real-life busy ER departments.

Also, when applying this method it becomes the evaluator's responsibility to determine what constitutes an issue. This is different from other evaluation methods, e.g. use of think-aloud protocol analysis where it is often the user who indicates they are having a problem in using the system [25]. This disadvantage is especially relevant in situations where the issues found are related to the internal cognitive processes of the user. In such cases the evaluator might unknowingly infer issues that are non-existing for the user, e.g. confusion regarding tasks, menu item placement and the meanings of icons. It is less relevant when the issues found are related to visible aspects of the system, e.g. system bugs, complicated pathways and issues that end up invoking error messages.

There are also practical considerations when applying the method. One problem that we faced during the study was that the recordings were interrupted due to technical problems or users logging in and out of the workstations and thereby disrupting the recordings. Following a regular routine of checking the recordings throughout each recording session can mitigate this issue but it could still be a potential issue when conducting recordings at multiple points as we did. Also, the method is time consuming in terms of data analysis since each video file must be viewed at least one in near real time pace. This makes the methodology less applicable for quick usability evaluations and more relevant for longer evaluations of system deployments. However, using video annotation and time stamping software during the first viewing can reduce the overall analysis time, since these tools allow easy navigation between coded parts of the videos when necessary.

Finally, there are certain challenges when applying the method described in this paper. One challenge is caused by the naturalistic nature of the method. When applying this method the evaluator does not have any influence on what the results of the evaluation will show, as is to some degree the case with other evaluation methods, e.g. usability testing involving think-aloud protocol analysis or clinical simulations where the tasks to test the system are predetermined by the investigators. In our case this meant that many of the found issues could not have been predicted using a narrowly defined coding taxonomy. We therefore chose to apply a broad coding taxonomy and set of categories of usability issues and also allow issue to emerge from the data during the analysis. Researchers that intend to apply the method should in the same way allow issues to emerge in a similar data-driven fashion. Another challenge related to the naturalistic nature of the method is the issue of how much data will be captured during the evaluation. As mentioned in the description of our results there were some differences in the number of entries between the three sessions. However, we found that the average lengths of the recordings from the three sessions were approximately the same. This indicates that interrupted recordings or other technical issues are
not responsible for the differences in number of entries. Instead, the differences may have been caused by differences in how busy the EDs have been during the three sessions or how often the selected access points were used throughout the recordings. This poses a challenge for the evaluator when selecting the points for recording and when analyzing the data from the recordings. In this case we chose to compare the percent-wise distribution of the occurring problems to overcome this challenge.

5. Conclusions

Using continuous screen recordings of user interactions with an EW system at two EDs we applied a longitudinal and naturalistic approach to studying the usability issues related to the usage of this system. Through the application of the approach and the analysis of the resulting recordings we found a wide range of system-related usability issues that did not appear to change over time as the collective experience of the users at the different EDs increased. However, we found that certain work patterns related to different tasks did in fact change as the departmental experience with the EW system increased. In one work pattern we found that the users appeared to become more efficient with the EW system, which indicated that despite potential efficiency issues in the initial stages of use users have the ability to overcome these issues and learn to use the EW efficiently.

In another work pattern we found that the users became less efficient as the department gained more and more experience with the EW system. We argue that a mismatch between this work pattern and the work practices of the ED coupled with the clinician’s possibly forgetting the functionality of the EW as they move further and further away from the initial stages of usage could cause this decrease in efficiency. This indicates that some aspects related to the EW system and the use of it, need to be rethought and possibly redesigned. For the work pattern regarding logging in to the EW system a potential solution could be to conduct a technical redesign aimed at providing the users with cues to remind them to log in before initiating activities. Another approach would be to periodically refresh the users’ training in an attempt to remind them how to use the EW system most efficiently. This would also assist in providing new users with the necessary training in a working environment characterized by a high degree of staff turnover.

Through the study we also gained experience with the application of the methodology applied. We found that the method affords a number of advantages over more traditional usability evaluation methods. We especially find that the longitudinal and naturalistic nature of the method provides researchers and usability evaluators with a tool to uncover issues that would be difficult to reveal with other methods, e.g. the ability to expose usability issue that occur very infrequently or only under very specific conditions that can be difficult to predict using more traditional evaluation methods.

However, the application of this methodology is not without limitations, disadvantages and challenges. Some of these are inherent in the methodology and can therefore not be avoided when applying the method, e.g. time-consuming initial evaluation of digital video data, increased evaluator responsibility and reduced control of the final results. These issues have to be accounted for when analyzing and reporting the results. Others issues could be mitigated by enforcing certain procedures throughout the application of the method, e.g. inspecting recordings regularly to avoid interruptions. In addition, it should be noted that the usability of electronic whiteboards may vary across settings and that usability may vary across functions used, requiring further analyses of such systems under different contexts to assess the generalizability of the findings from this study. In conclusion we encourage that more research be focused on longitudinal and naturalistic evaluations of health information systems and we encourage the use and refinement of the method described in this paper with the hope that researchers will continue to improve the systems that keep our hospital running smoothly and safely. This work will be important in order to help identify and mitigate usability issues and potentially detrimental impact of newly introduced health information systems on workflow, based on analysis of live interactions conducted in the implementation process.

Acknowledgments

The authors would like to thank and acknowledge the clinicians of Slagelse Hospital Emergency Department and Nykøbing Falster Hospital Emergency Department for their participation in the study and allowing us access to their work life. Finally, we would like to thank the healthcare region for its support of the study.

References


