Örebro University School of Business
Department of Informatics

Master thesis

Supervisor's name: Professor Gunnar O. Klein
Examiner's name: Dr. Shang Gao
Semester: VT/2016

Evaluating the usability of 'Journalen': An Electronic Health Records System for Patients in Sweden

Authors

Name: BUKENYA Charles

Email: bukenyacharles@rocketmail.com

Name: IBRAHIM Al-Hassany

Email: ibha81@yahoo.com

ABSTRACT

There is growing adoption of Electronic Health Records (EHR) systems to promote a more patient centered healthcare. There is great potential to utilize EHRs to improve quality of care, limit the rising costs in healthcare and improve healthcare outcomes. However Usability has proven a major challenge for achieving the intended goals for EHR systems. Usability has been defined as how useful, usable, and satisfying a system is for the intended users to accomplish goals in the work domain by performing certain sequences of tasks. There is a need to ensure that EHR systems meet the usability standards in order to achieve the expectations from users and healthcare stakeholders which requires continuous monitoring and evaluation of EHR systems to ensure they meet the required usability standards. This study is an evaluation of 'Journalen', a Swedish EHR system that allows users to access and view their health records online. Following the TURF framework, a systematic usability evaluation has been conducted by testing users on the system through controlled usability Lab tests and an expert review. TURF (Task, user, representation and function) is a framework for defining, evaluating and measuring usability objectively. The major findings show that to a greater extent, 'Journalen' meets the usability requirements according to the TURF framework. The users have shown a high level of satisfaction with 'Journalen' as measured by high user SUS (system usability scale) scores on the system. On the ease of use referred to as 'Usableness' in this study as evaluated through the Task performance evaluations, results show a high degree of Usableness for 'Journalen'. On the Usefulness, the expert review has shown major usability issues like lack of minimalistic representation of data, poor language use and unmatched user expectations that need to be addressed for 'Journalen to meet optimal usefulness.

Keywords: *Usability, Electronic Health Records (EHR), Evaluation, Design Heuristics, Health records, TURF, User Centered Design (UCD).*

1.0 INTRODUCTION

1.1 A brief background of Usability in EHR.

With the growing influence of IT in the medical field, there is a focus to utilize Information systems to promote a patient centered health care (Myreteg, 2015). One of such endeavors has been the rise in Electronic Health Records (EHR) systems for patients. ISO defines (EHR) as: "*a repository of information regarding the health of a subject of care in computer processable form, stored and transmitted securely, and accessible by multiple authorized users. It has a commonly agreed logical information model which is independent of EHR systems. Its primary purpose is the support of continuing, efficient and quality integrated health care and it contains information which is retrospective, concurrent and prospective*" (ISO/TR 20514, 2005).

According to Menachemi and Brooks (2006) "EHRs for patients are essential for improving compliance with formularies and dosing guidelines, medical error reduction, improving satisfaction for patients, improving drug administration, improving preventive care and increasing medical research potential". There is also added security for the medical records. In 2005 Hurricane Katrina destroyed an uncounted number of medical record files in the US (Williams et al, 2008). Effective use of EHRs also has the potential for reducing the per capita cost of health care for countries (American Medical Association, 2014).

All the above, coupled with other factors have led to an increased interest in implementing EHR systems especially in developed countries in order to improve the health outcomes for patients.

In Sweden access to health records is a right. The National eHealth Strategy provides that patients should be able to access part of their health records also via the Internet. Thus, all county councils or healthcare regions in Sweden have decided to make EHRs available online for all patients by 2017 (Scandurra, 2015).

While there is consensus in informatics research on the benefits and potential of EHRs for improving quality of care for patients, the Usability standards of such systems has been a major point of concern (Johnson et al 2011). Usability is how useful, usable, and satisfying a system is for the users to achieve intended goals

by conducting certain sequences of tasks on the system (Zhang & Walji, 2011). Usability describes how easy it is for users to accurately and efficiently accomplish a task while using a system (Johnson et al 2011). In essence, a system with good usability is easy to use and effective. It is intuitive, forgiving of mistakes and allows one to perform necessary tasks quickly, efficiently and with minimum of mental effort (Belden et al, 2009).

While Usability is important in all information systems, it is critical for EHRs and health systems in general because of the dire consequences an error could mean for the life of a patient. For example if a patient misunderstands a medical prescription due to poor data representation and takes an overdose, a life could be lost (Belden et al, 2009). According to ISO/IEC (2011), usability is one of the 8 key software quality requirements.

EHRs that are not designed with usability in mind can hinder progress and may not be able to meet the intended outcomes (Zhang & Walji, 2011). Karsh, 2004 also emphasizes that in order for EHRs to achieve their goals, they must be designed, developed, and implemented with a focus on usability and safe use (Karsh, 2004).

1.2 Evaluating Usability in EHRs.

Given the criticality of Usability in EHRs, developers need to ensure that systems developed always meet established standards for Usability. However this is not always the case. A 2014 NCCD (American National Center for Cognitive Informatics & Decision Making in Healthcare) survey of 11 EHR software companies found that 5/11 did not employ a single Usability expert and had no implementation of any standard usability evaluation processes in their development process. This shocking revelation even highlights more why there is a need for rigorous usability evaluation of EHR systems.

According to a 2013 Swedish E-Health systems usability study, it was recommended that "*E-health systems must be managed, evaluated, supervised and continually optimized in relation to Usability*…" (Scandurra et al 2014).
According to NCCD (American National Center for Cognitive Informatics & Decision Making in Healthcare), Usability evaluation should be integrated within the development processes of EHRs through a User Centered Design (UCD) approach. "*In this iterative approach to design, the user is a major part of the*

process from first to last. Designers and engineers don't simply make assumptions about how users are likely to use a product, they use scenarios, create use cases and test their predictions with actual users, with formative and summative assessment techniques among others", NCCD (2015).

Even after implementation, EHRs should be continually evaluated for Usability to establish how they perform against the usability standards. This is because user expectations of systems can change over time. For example a floppy disk icon means saving a document but this may not be clear for a young user who never used Floppy disks. Therefore Usability evaluation should be part of the whole product life cycle.

> *"Usability is the result of careful design and evaluation throughout Product development". (Belden et al, 2009)*

There are various methods that can be used to evaluate usability in EHR. Johnson et al, 2011, has summarized the various methods for Evaluating Usability in EHR as seen in table 1 below. These methods however can be applied based on evaluation context and criteria which can be conceptualized through a Usability framework. The methods listed by Johnson et al, 2011 in table 1 below form part of the recommended methods and tools for Usability evaluation suggested under TURF framework for Usability evaluation in EHR which is used in this study.

Table 1: *Usability evaluation methods. (*Johnson et al, 2011*).*

	Description of Method	**Advantages**	**Disadvantages**
Heuristic Evaluation	Usability experts evaluate a system using a set of design principles/guidelines.	Low cost and addresses both local (interfaces) and global (system) usability problems.	Requires usability experts to complete the analysis, and it may overlook some usability problems that could only be found with user input.
Cognitive Walkthrough	Experts imitate users stepping through the interface to carry out typical tasks. Finds mismatches between users' and designers' conceptualization of a task.	Focuses on ease of learning for first time users.	Does not determine all problems with an interface. Requires expertise.
Controlled user testing	Users test the system performing representative tasks using verbal protocols. Testing gathers information about the users' performance, includes post-tests of usability and observations made by the evaluator.	Performance measurements can be obtained in addition to verbal protocol information. Quantitative results are easy to compare.	Detailed planning is required prior to running these tests. Requires experts to run the tests in controlled laboratory.
Remote Evaluation	System records events as user works through the tasks and collects results of questionnaires. Includes asynchronous and synchronous approaches.	Accurate performance measures can be obtained. Data can be ready for analysis from questionnaires.	Software can be costly.
Usability Questionnaires	Questionnaires that measure efficiency, satisfaction, learnability, system usefulness, information quality and many other measures.	Questionnaires are easy to administer online and provide written feedback and scores. Many are reliable and validated.	May not be specific to EHR systems; may only focus on assessing overall usability.
Predictive Modeling	Determines user goals to complete a task, operators to perform the goal, methods to accomplish the goal, and selection rules to reach the goal. It is part of the cognitive task analysis.	Calculates the time to reach the goal. Includes Key-Stroke level models.	Very time intensive and requires usability expertise.
Failure Modes and Effects Analysis (FMEA)	Analyzes human reliability, identifies potential failure modes, can be used to study human errors based on tasks and functions.	Cost effective test and can determine errors by severity, permits descriptive information on different types of errors.	Depends on expertise of analyst, can be time consuming to analyze.
Critical Incident Technique	Identifies and determines design flaw via self-report.	Cost effective. A method of gathering data that can be analyzed for trends. Helpful for rare events, has high face validity, provides information on types of errors.	Dependent on users' verbal reports.
Subjective Workload Assessment Technique (SWAT)	Evaluates workload measuring time load, mental effort, and psychological stress.	Most frequently cited in workload literature; theoretically grounded.	Scale must be normalized for each subject by means of a card sorting technique, large amount of subject preparation and training. Subjective rating techniques that uses three levels: low, medium, high. Low sensitivity for low mental workloads.
Simplified Subjective Workload Assessment Technique (SSWAT)	Evaluates cognitive/workload measuring time load, mental effort, and psychological stress.	Theoretically grounded. Correlated well with original SWAT.	Needs validation in a medical environment.
NASA Task Load Index	Evaluates workload measuring mental demand, physical demand, temporal demand, performance, effort, and frustration. Measures each component subscale with 20 levels.	Uses an adjustment to normalize ratings.	Will not determine many usability issues. Scale must be normalized for each subject but less time intensive than SWAT.

Subjective Mental Effort Questionnaire (SMEQ)	Subjective measure of mental effort.	Contains one scale with nine labels that measures subjective mental effort after each task completed. Time limited, easy to use.	Requires analysis by usability experts to interpret results.

While all the above methods can be useful for evaluating usability in an EHR, it is upon the discretion of an evaluator to determine the suitable method depending on the context. A Usability framework can help an Evaluator select the suitable methods for Evaluating the Usability of an EHR system. One such framework that is widely used in EHR Usability Evaluation is the TURF Framework by Zhang & Walji (2011). Another is the NIST EHR usability protocol. The NIST EHR usability protocol framework is tailored for mostly pre-development planning and is more inclined to ensuring patient safety in EHRs. The TURF framework is very specific for evaluating Usability in EHR and is used in this study as described later in the methods section of this paper.

1.3 The case study: 'Journalen'

'Journalen' is an Electronic Records Systems for patients designed to enable patients in Sweden to electronically access their health records from the regional government health care centers. The system is developed by Inera.

Inera is an organization co-owned by Sweden's regional governments that coordinates the county and regional common e-health work and develops services for the benefit of residents, care staff and decision makers. Since a few years back Inera worked to develop new technology in the form of standardized integration health profiles, which allows for a standardized access by patients to part of their health records from healthcare centers in counties / regions, municipalities or private care providers. The entire IT architecture was tested with great success.

In 2011, Inera conducted a preliminary study for an online patient's records system "Journal på nätet". It examined among others, patients' and their families' expectations and needs when it comes to getting access to online health records. The feasibility study also viewed caregiver attitudes and concerns, as well as legal issues. The results of the feasibility study showed that patients were interested to access their health records through an online EHR. Despite Sweden's early adoption of EHRs in the healthcare givers' side, patients accessed their medical records through a print version medical journal.

This led to, in 2015, the development of 'Journalen', which showcases comprehensive medical records to patients online. This Health Record System is developed on the basis of experience and knowledge from the pilot study (Journal på nätet) mentioned above and an early initiative (Sustains project) in Uppsala County Council which was the first to give patients online access to their health records in Sweden on a large scale.

'Journalen' has now been launched in 12 out of the 23 counties in Sweden and it is planned to cover all regions by 2017. There has been various academic studies about the Cost-benefit, technology acceptance, participation and value creation of 'Journalen', but there has not been a Usability Study to understand how the system meets the standards of Usability for EHR systems.

However since this project is in the rollout phase, it is important to understand how 'Journalen' meets the usability requirements through a systematic Usability study. This is important to ensure that 'Journalen' will meet user expectations, promotes a user centered health care which are major goals for such systems. The study is also useful for highlighting usability improvement areas for 'Journalen' project developers.

1.4 Research Question and Objectives.

The main objective of this study is to carry out a systematic usability evaluation to establish the level of usability of 'Journalen'. The study is conducted on basis of the TURF framework recommendations for Usability Evaluation of an EHR.

Other sub-objectives include;
-Determine user satisfaction of 'Journalen' through a System Usability Scale evaluation. -Evaluate the 'Usableness' of 'Journalen' through user task performance evaluation.
-Determine the usefulness of 'Journalen' through an expert review heuristic evaluation.
-Identify Usability improvement areas for 'Journalen'.
-Contribute research knowledge on Usability Evaluation in EHRs.

The main research question for this study is: *To what extent does 'Journalen' meet the requirements for EHR Usability through evaluation of user satisfaction, task performance and usefulness according to the TURF usability framework.*

2.0 METHODOLOGY

2.1 Research approach

This study is mainly based on summative evaluation techniques (Belden et al, 2009). Summative evaluation is carried out at the end of a product development to validate how it meets the standards (Belden et al, 2009). The other type of evaluation approach is formative and is carried out during the design and development process in support of defining the application, understanding the user and user workflow, and making iterative improvements to the product (Belden et al, 2009). But since 'Journalen' has already gone through the development phase and is in the implementation phase, a summative approach was selected. "*Summative techniques involve but are not limited to Expert Reviews, Performance Testing, Risk Assessment and one-on one usability testing*" (Belden et al, 2009).

2.2 Conceptual framework

In order to synthesize and scope this study, a framework has been used to conduct this study in a systematic manner. The TURF framework for EHR usability has been used. TURF which stands for (Task, User, Representation, and Function) is a unified theory/framework developed by University of Texas School of biomedical informatics NCCD center under the (SHARPC project). The framework has been widely adopted for evaluating Usability in EHR. TURF is a modern and very specific framework for Evaluating Usability in EHR which makes it a suitable framework for this study.

Other tools like Nielsen's mathematical model for finding usability problems and De Lone and Mc Lone's model for Information systems success have been used before for Evaluating usability in EHR systems but are not specific for EHR usability rather for Information Systems in general. The TURF framework has been used in designing the evaluation tools (see appendices 1 &2) and selecting evaluation methods in this study. Figure 1 is a diagrammatic representation of the TURF framework.
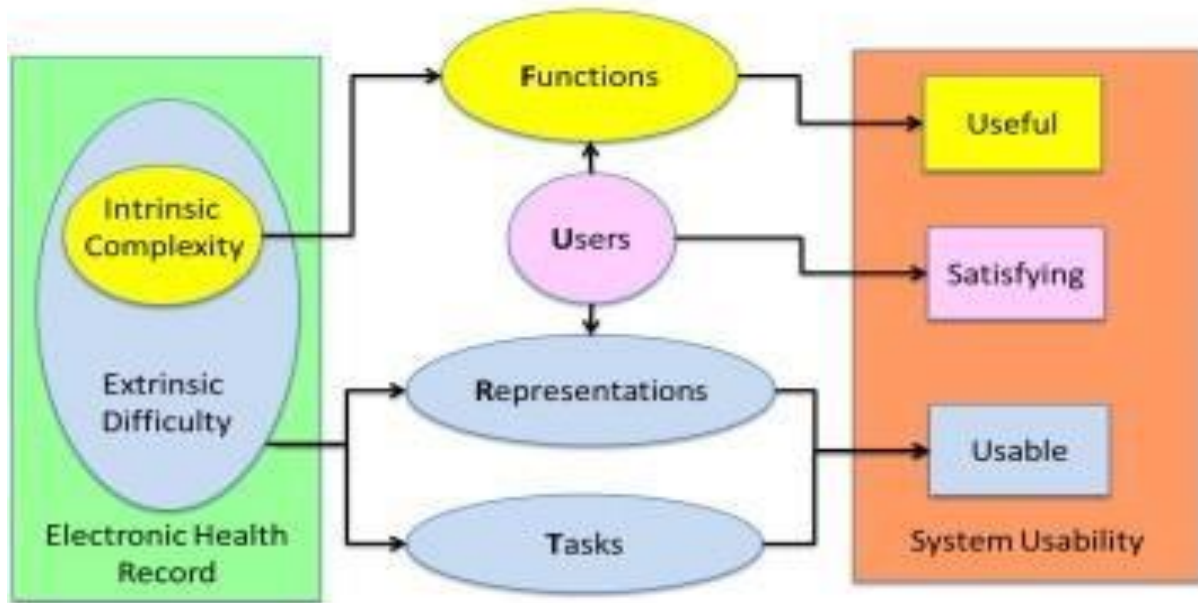
Figure 1. The TURF framework of EHR usability. (Zhang & Walji, 2011).

TURF defines usability as "how useful, usable, and satisfying a system is for intended users to accomplish goals in a work domain by performing sequences of tasks". TURF suggests tools for evaluation of all the three dimensions for usability as mentioned above. i.e. (Usefulness, 'Usableness', and Satisfaction). These tools and methods are among those mentioned by Johnson et al (2011) in Table 1 above. TURF provides a conceptualization and definition of Usability but also combines both new and existing Usability evaluation tools to be able to evaluate Usability in EHR. These recommendations by TURF have been used in the conduct of this study. Table 2 below shows the evaluation methods used in this study as recommended by the TURF framework.

Table 2: *Using TURF to select evaluation methods*

TURF Usability Dimension	Evaluation methods used in this study.
Usefulness: According to TURF a system is useful if it supports the work domain where users accomplish goals for their work independent of how the system is implemented. A system is fully useful if it includes domain, and only domain, functions essential for the work, independent of implementations.	**Expert Review** Under this method suggested by the TURF framework, experts imitate users stepping through the interface to carry out typical tasks. Finds mismatches between users' and designers' conceptualization of a task. This method is effective for functionality but may not identify all flaws in the interface. This is why it is recommended for usefulness which focuses on functionality of the system. According to Nielsen (1992), expert reviews are the most effective Usability evaluation methods finding 80% of usability problems compared to 50% from regular usability testing. Expert reviews are also objective compared to user tests which may have some subjective bias, Landauer and Nielsen (1993). An Expert was selected to give reflections on the system and to identify problems. These problems were then matched with a heuristic principle violation under the 14 TURF usability heuristics. It would have been more significant to have multiple experts so as to compare results however it was not possible to conduct multiple expert reviews.
Satisfaction: In TURF, satisfaction refers to the subjective impression of how useful, usable and likeable a system is to a user	**SUS. System Usability Scale (Brooke, J. 1996)** SUS is an established general tool which consists of ten questions intended for measuring User satisfaction with Systems.(see appendix 1) TURF recommends SUS for evaluating Satisfaction. According to Sauro, J. (2015), SUS is a valid, reliable and quick tool for user satisfaction having been used in over 5000 usability studies worldwide. Users were asked to fill out a System Usability Scale after performing several tasks on the system. The SUS was useful in understanding the overall satisfaction of the users with 'Journalen'.

'Usableness': A system is usable if it is easy to learn, efficient to use, and error-tolerant.	**Task Performance Evaluation.** Task performance indicators like time spent on tasks, task completion rate, error occurrences and steps taken to complete tasks are more objective measures for evaluating usability, Landauer and Nielsen (1993). However there is no standardized Task Performance Evaluation tool. We have created a specific Task Evaluation Form for this study.(see appendix 2) 12 Tasks were prepared for users to carry out on the system. Users were asked to fill out a performance evaluation form (see appendix 2) to determine their perception of the learnability, Efficiency and error-tolerance of the system. After further attestation, only 10 tasks were used.

2.3 Data collection and Selection of participants

Data Collection was carried out through controlled Usability Laboratory User tests and an Expert review between 15th April 2016 and 15th May 2016 in Örebro, Sweden. A group of 10 users aged between 20 to 35, 5 male and 5 female were selected for the lab tests and 1 expert for the expert review. Users who were selected are Swedish speaking students at the University of Örebro because the patients' journal system (Journalen) is in Swedish. A user group of students is selected so as to get a more informed judgement of the system usability since they are used to using IT systems regularly during their study in the university. Such users can compare with other experiences using other similar systems. The choice of equal male to female participants was important to absorb a gender balanced perspective on the system. This group may not be representative of the actual users of 'Journalen' but can give an insightful review of the system usability for design improvements since the system is still in implementation phase. All participants provided consent.

A single User Test was carried out for each lab session. A user was asked to carry out designated tasks on a user task performance evaluation form (see appendix

2) representative of the system functions. The system was already opened to the home page using a test account, login and system authentication were not part of this test. The main focus was on the core functions of the system. The tasks have been designed to cover the system functions so that the test will cover as much as possible of the system. All the data used in the system was demo data created for the purpose of system testing. This is because of the ethical limitations of using real health data in system testing. Even so, it was found that the demo data was lacking in some circumstances.

The user filled out a performance evaluation form for each task done. This performance evaluation was designed to evaluate 'Usableness' of the system which is one of the dimensions of usability according to TURF. The 6 performance Evaluation questions (see table 6) are designed according to recommendations by TURF to reflect the 3 'Usableness' goals in EHR Usability. I.e. EHR systems should be easy to learn, efficient to use, and error-tolerant. All tasks were timed. Verbalizations by the user during the task were recorded. User demographics like gender and age were also recorded for further demographic analysis.

Each user completed an SUS- System Usability Scale (see appendix 1) after completing all tasks. The SUS is intended to measure overall satisfaction of the users with the system which is one of the dimensions of Usability according to TURF. All data for each individual User Test was recorded in the Turf 4.0 (see appendix 3) Usability software suite for centralized storage and further analysis.

Also one Expert Review was conducted with a medical doctor at Skebäcks Vårdcentral in Örebro, Sweden. The Expert is also a professor in eHealth informatics which made him a suitable candidate for this study. The expert has over 20 years of experience in research, development and use of different Electronic Health Records Systems in Europe and also has extensive knowledge in usability of ehealth systems. The expert conducted a cognitive walk through on the 'Journalen' system functions based on the user tasks and identified different usability problems. This session was recorded in the Turf 4.0 software Screen and Video capture tool and stored for further analysis. Also notes were taken in this session for all major issues highlighted by the expert. The expert review was conducted for evaluating usefulness of the system ('Journalen').

2.4 Data Analysis

Mixed methods are used for analysis of data in this study. Both quantitative and qualitative methods are used for analysis of collected data to reach the findings presented in the results section. The choice of each data analysis method is based on the type of data in order to derive meaningful information relevant to the research question. Since the data from the Task performance and SUS scale was mostly numerical, quantitative methods are used while for the Expert review, qualitative methods are used since the data collected in this session was mostly qualitative. Table 3 below shows the data analysis methods used for each data type.

Table 3: *Data Analysis methods.*

Data Type	Methods used for data Analysis	Tools used
Task Performance Evaluation Data	-Frequency -Summation -Average -Mode -Tabulation	-Turf 4.0 -MS Excel -MS Word
SUS data	-Mode -Median -Tabulation	-Turf 4.0 -MS Excel --MS Word
Expert Review data	-Inductive Analysis -Data coding -Summarization	-Turf 4.0 -MS Word

3.0 RESULTS

3.1 Results from the System Usability Scale (SUS)

The System Usability Scale comprises of 10 questions answered by users to evaluate their subjective satisfaction with the system ('Journalen'). It was filled out after users completed a set of designated tasks on the system. The results show female users had a slightly higher satisfaction with the system than males. Females had a higher mode on all positive questions (A, C, E, G, I). Mode is the value that appears most times in a data set. Women agreed more that they would use the system more frequently than males. On question A which asks if a user would like to use the system more frequently, the mode for women was 5 (the highest rating) while that for males was 4. Nevertheless the median rating from all users despite of gender, on all questions was very high which is positive on usability of 'Journalen'. The results from the System Usability Scale show that the users were generally very satisfied with the system performance. This is positive on the usability of 'Journalen' as shown in table 4.

KEY FOR SUS QUESTIONS

A - I think that I would like to use this system frequently
B - I found the system unnecessarily complex
C - I thought the system was easy to use
D - I think that I would need the support of a technical person to be able to use this system
E - I found the various functions in this system were well integrated
F - I thought there was too much inconsistency in this system
G - I would imagine that most people would learn to use this system very quickly
H - I found the system very cumbersome to use
I - I felt very confident using the system
J - I needed to learn a lot of things before I could get going with the system

Table 4: *Table of results from the SUS.*

	A	B	C	D	E	F	G	H	I	J
M1	1	1	5	1	4	1	5	2	4	1
M2	4	1	4	1	3	2	4	1	5	1
M3	3	1	5	1	4	1	4	1	4	1
M4	4	2	4	1	3	3	4	2	4	2
M5	4	2	4	2	4	1	5	1	5	1
F1	5	1	5	1	5	2	5	1	5	1
F2	4	2	4	2	4	2	4	2	4	2
F3	3	1	5	1	2	3	5	1	5	1
F4	3	1	5	1	4	1	5	1	4	1
F5	5	1	5	1	5	1	2	1	5	1
Mode	4	1	5	1	4	1	5	1	4	1
Mode (M)	4	1	4	1	4	1	4	1	4	1
Mode (F)	5	1	5	1	5	2	5	1	5	1
Median	4	1	5	1	4	1.5	4.5	1	4.5	1

Key for table 4:

Mode - Overall Mode, **Mode (M)** - Mode for Males, **Mode (F)** - Mode for Females

1= strongly disagree **5**= Strongly Agree **M1**-**M5**: Male Users. **F1-F5:** Female users

3.1.1 Additional comments from users

In addition to the SUS ratings, users were asked if they had general comments about the use of 'Journalen'. One female user noted that she would be very concerned about the system security and it would greatly affect her general satisfaction with the system. "*I have to be confident that my data is secure and that there cannot be any breach of my privacy*". However the SUS doesn't have any question on system security which may be a limitation in the evaluation tools. Several users noted that they would like a more visualized interface with less text as it is now. Also many users noted that while they were able to complete the tasks and read the information, they said that they were more

interested in the interpretation of the information than the information itself. Some complained that they could not make meaning out of some of the information provided by the system.

3.2 Results from the task performance evaluation

The results below show users' perceptions towards the system 'Usableness'. An extensive analysis of user's ratings was carried out to determine how individual users and all users in general felt carrying out tasks on the system. The tasks as seen in Table 5 were selected to represent the different core functions on 'Journalen'.

Table 5: *User Tasks*

Task	Description	System Function Area
Task 1	Which doctor carried out the latest diagnose.	Diagnoses **(Diagnoser)**
Task 2	Which medicine was prescribed to you by **Henry S Johansson**	Medicines **(Läkemedel)**
Task 3	Check the results of the test **(Blod, urin eller annat vätskeprov)** that was carried out on date (14/06/2005) and read the comment.	Lab Tests and results **(Provsvar)**
Task 4	Find the issue for which you were referred to care unit **(Ortopeden 2, Skånes universitetssjukvård, Region Skåne**) in referral **(Röntgenremiss).**	Referrals **(Remisser)**
Task 5	Find the notes made by doctor (**Britt Thunblom**) between date (**20/09/2015**) and date (**01/10/2015**)	Medical Notes **(Anteckningar)**
Task 6	Find your last vaccination and read about its side effects	Vaccinations **(Vaccinationer)**
Task 7	Find all the vaccinations you took in the last year	Vaccinations **(Vaccinationer)**
Task 8	Find the dosage for medicine **(Tavegyl)**	Medicines **(Läkemedel)**
Task 9	Check your last appointment (**with whom, date and where**)	Appointments **(Besök)**

Task 10	What was the first diagnose you had? And what is the date?	Diagnoses (**Diagnoser**)

Note: The tasks that were attested are 10 from the 12 that were first designed (*See appendix 2*)

3.2.1 Task Performance Evaluation results

All task performance data has been analyzed to reach the conclusions below. First, in order to understand how the conclusions have been arrived at, a detailed sample result summary for only Task 3 is presented in order to demonstrate how the results/findings were reached at. A summary of results and findings for all the tasks later follows. For a detailed task analysis see appendix 4.

3.2.2 Summary of results for Task 3

TASK 3	FREQUENCY OF RATINGS					PERCENTAGE OF RATINGS				
TASK EVALUATION QNS	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Takes little mental effort	0	0	1	6	3	0%	0%	10%	60%	30%
Takes short time	0	0	1	5	4	0%	0%	10%	50%	40%
Took few steps	0	0	0	4	6	0%	0%	0%	40%	60%
Completed successfully	0	0	1	1	8	0%	0%	10%	10%	80%
Easy to find help	10	0	0	0	0	100%	0%	0%	0%	0%
Easy to remember	0	0	0	2	8	0%	0%	0%	20%	80%

SUMS	10	0	3	18	29	**MODE**	5
PERCENT	16.67	0.00	5.00	30.00	48.33	**AVG TIME(sec)**	41.1

Table 6: *Results summary for Task 3 performance evaluation (QNS-Questions)*

The table of results above shows the frequency of ratings for each evaluation question on Task 3 for all participants. (1 strongly disagree-5 strongly agree) and on the right are the percentages of each rating for each evaluation question for all users. The evaluation questions on the left are based on the criteria under

TURF which defines Usableness for a system task. The average time it took all users to complete task 3 was 41 seconds. The highest average time to complete a task was 62.4 seconds (Task 5) and the lowest being 15.8 seconds (Task 1).

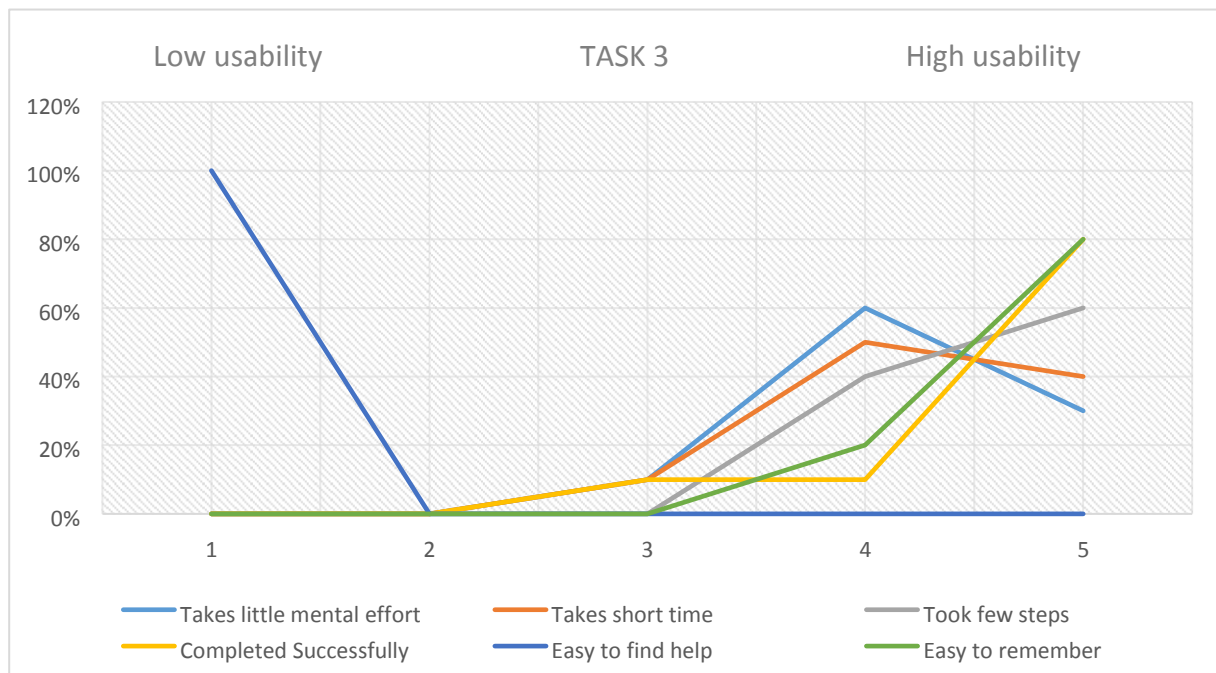The line graph below illustrates how Task 3 was rated.

Figure 2: Line graph the showing percentage distribution of the frequency of ratings for each evaluation question on Task 3. On the X Axis is the scale for rating (1- strongly disagree to 5- Strongly agree) Y Axis - Percentage of ratings for a particular scale value.

Since all the task performance evaluation questions are designed with a positive proclivity on usability (see table 6), a 1 rating means less usability and a 5 rating means high usability on that particular question. Since most users rated 5 or 4 for Task 3 above, we can therefore conclude that the usability is high for Task 3 and so is true for the system function (Finding Lab Tests and results) which Task 3 represents in the system domain. The mode for all ratings on Task 3 is 5 and this confirms a high usability rating.

However it can clearly be seen on the figure 2 above that 100% of users rated 1 for 'Easy to find help' evaluation question. This means that the users strongly disagree with the question. We found this common for all Tasks 1-10 for the

same question. Upon further review we found that the system lacks user help in all system functions. This is a major usability problem for 'Journalen' because it increases a chance for users to make errors which may affect the overall usability of the system.

3.2.3 Summary of results for all Task Performance Evaluation ratings

The combo bar graph below shows the summary of ratings for all tasks based on the percentage of sums of frequencies for ratings on each task.
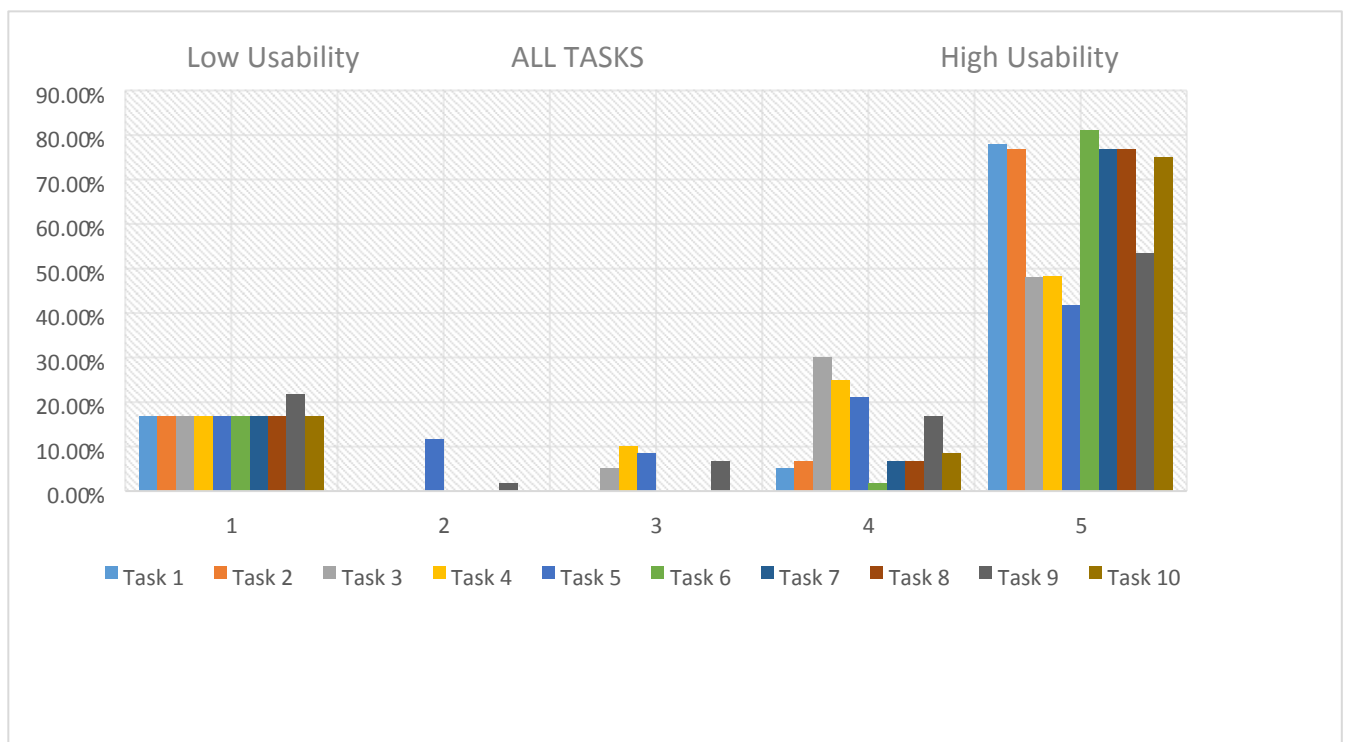

Figure 3: Percentage of sums of frequencies for ratings on each task. On the X Axis is the scale of rating (1- strongly disagree to 5- Strongly agree). Y Axis is the percentage of ratings for a particular rating scale value for each task.

It can clearly be seen on the graph above that for most tasks, the majority of ratings were 5 which shows a high usability performance for the System. ('Journalen'). Since usability concerns with accurately and efficiently accomplishing a task while using a system, the high rating on individual tasks can translate into overall Usability of the system. Task 3, 4, 5 and 9 however can be identified as with lower usability compared to task 1, 2, 6, 7, 8 and 10. This corresponds well with the average time for task completion in table 7 below. The

ones with lower usability also took longer to complete while the ones with higher usability took less average time to complete for users. Some can argue that some tasks are complex than others and therefore the users will have different perceptions about them, nevertheless the very essence of usability is that all tasks should be easy to accomplish on the system. Therefore the argument of complex tasks is an argument of poor usability design.

However it can also clearly be seen that the percentage of 1 ratings is almost equal for all tasks. This skewedness in data is because 100% of all users found that it was not easy to find help for all tasks as already mentioned above. Though users were able to complete most tasks successfully, they said the system lacked any user help information like captions on features to describe what the user is supposed to do, there are no bread crumbs or any other user navigation tracking to help the user see their path and navigate backwards etc.

Table 7: Average time in seconds taken to complete tasks.

Name	Task1	Task2	Task3	Task4	Task5	Task6	Task7	Task8	Task9	Task10	AVERAGE TIME PER USER(SECS)
M1	20	45	51	8	17	44	5	25	55	5	27.5
M2	11	11	63	43	121	22	16	40	32	36	39.5
M3	13	11	16	18	53	13	9	9	20	17	17.9
M4	7	16	37	57	48	11	7	7	51	14	25.5
M5	46	39	62	79	126	26	55	30	122	74	65.9
F1	12	23	28	55	40	10	30	6	16	10	23
F2	14	23	50	85	112	15	36	6	27	10	37.8
F3	7	14	28	25	43	13	7	5	58	9	20.9
F4	23	25	28	42	37	19	42	14	55	28	31.3
F5	5	11	48	41	27	16	26	10	30	20	23.4
AVG TIME/TASK(SECS)	15.8	21.8	41.1	45.3	62.4	18.9	23.3	15.2	46.6	22.3	
						Average time for Males(sec)					35.26
						Average time for Females(sec)					27.28

One interesting finding from the task performance evaluation is that women on average spent 8 seconds less on a task. This may explain the finding from the SUS results above which found that women were overall slightly more satisfied with

the system than men. Nielsen (1992) has said that performance result will always have an impact on the subjective satisfaction of users with the system.

3.3 Expert review results

Below are the results from the expert review. 22 Usability problems were identified that need improvement in the system. These are matched to a Heuristic violation. Design heuristics are principles that must be considered in order to build systems with good usability. The problems highlight areas that need a critical analysis so that the system can achieve the highest degree of Usefulness for the intended purpose. The problems below may not cover all available due to the nature of this study. It requires more time to conduct an exhaustive expert review on the system which was not possible with the scope of this study.

Table 8: *Usability problems identified through the expert review.*

No	Problem Description	Problem Location	Heuristic Violation
1	The presentation of the dosage may not be very clear for users. A minimalistic representation like 1-0-1-0 where each digit represents a time point of the day can be used.	Medicines	04 Minimalism
2	The name of the person who records a diagnose in the system is shown instead of the doctor who carried out the diagnose himself. This may be confusing for the patients.	Diagnoses	03 Match
3	**'Per os'.** This Latin term which means to be taken orally is used in the system **(Administrationssätt)** on the medicines prescriptions but may be not understood by patients. Simple direct term like oral should be used.	Medicine	12 Language
4	**(Ändamål)**. Wrong term used in prescription and the information shown is for dosage.	Medicines	12 Language
5	Data loading on the timeline shifts position from where the user is currently viewing which may makes a user lose control	Timeline	13 Control
6	There is no user help in any form to support the user in the usage of the system.	General	14 Help

7	The naming of the notes and structure could be improved for easy visibility.	Notes	02 Visibility
8	There is no information about the recipient of the referral. Only the sender is shown.	Referrals	03 Match
9	It takes more steps to reach some information. It should be presented directly to avoid many steps for the patients.	Timeline	01 Consistency
10	The information shown in referrals could be more informative. For example the costs for the referred treatment so that patients can make informed decisions.	Referrals	06 Feedback
11	The substances contained in the medicine are not shown. Users may want to know this information.	Medicines	06 Feedback
12	Under Laboratory tests, more information is given than may be useful for the user	Tests	05 Memory
13	Presentation with tables and graphics may be more effective than the current textual representation for easy readability especially for older patients	General	04 Minimalism
14	There are no overlay captions to explain what the different functions mean which may lead to user errors.	General	09 Prevent errors
15	**'Förpackning'**. Not clear what the word means in the system context. No sample data was shown.	Medicines	12 Language
16	Using of too much medical language and too much information that the user may not understand under diagnoses.	Diagnoses	12 Language
17	The purpose for the prescription should be more structured. It is poorly described in the system.	Medicines	04 Minimalism
18	The heading of the medicine prescriptions in the timeline shouldn't be the name of the medicine but the name of the prescription.	Medicines	03 Match
19	There is no note for prescribed medicine.	Notes	03 Match
20	Wrong naming of a system function **(Alternativ).** This may not be very meaningful to the user to access other system settings.	Home	03 Match

| 21 | The diagnose name is not shown on timeline which may be misleading. | Diagnose | 02 Visibility |
| 22 | The information shown on the timeline may not be sufficient enough for a user to make system action. | Overview | 03 Match |

Note: Some of the Usability problems identified in table 8 above could be due to the poor quality of the demo data that was available for user testing.

The bar graph below summarizes the heuristic violations that were recorded for the above usability problems in table 8.
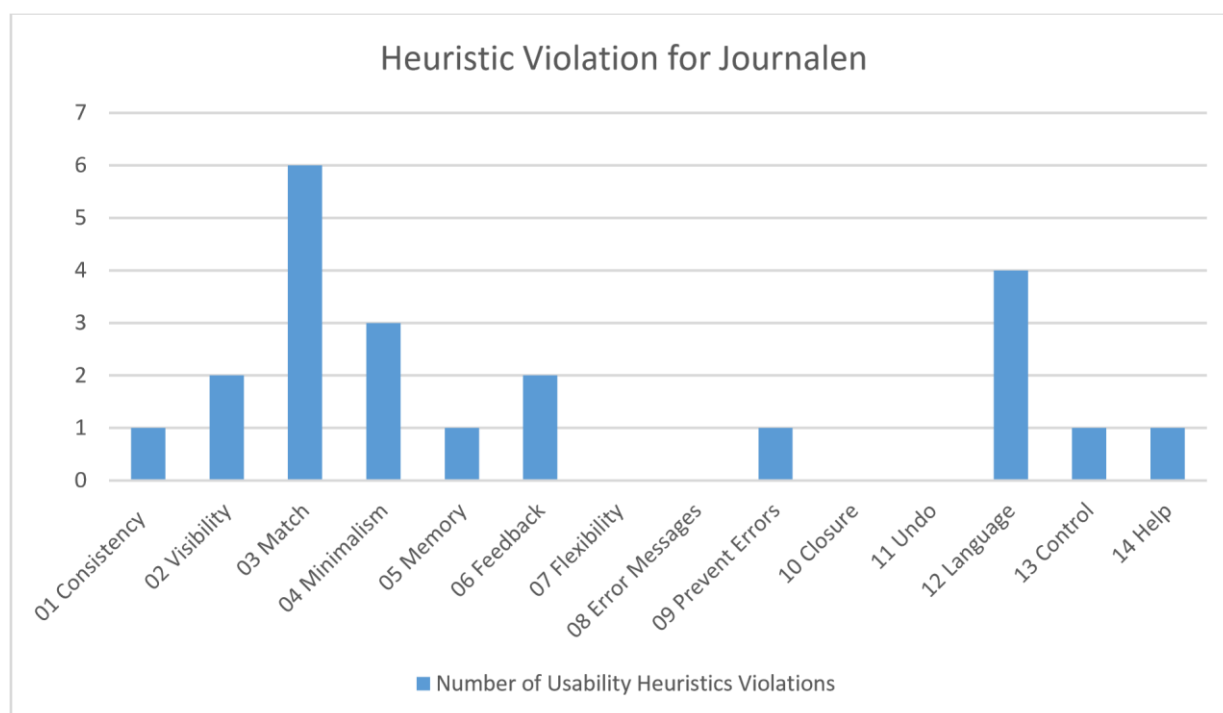
Figure 4: Shows the number of heuristic violations.

The results from the expert highlight various usability challenges for 'Journalen'. The expert identified that some of the functions may need further refinement to meet user expectations. This is especially with the use of language in naming of system functions and the data itself. Some names were found confusing while others complicated for the user to understand. The system is also only available in Swedish language unlike most public systems in Sweden. However the expert

recommends the system and thinks it is useful compared to the old print based system. But he notes that more can be done in terms of Usability to make the system more useful to the users.

4.0 DISCUSSION

This section is a critical reflection on the findings according to the three dimensions of usability according to TURF. This is to reflect on the significance of the findings on the usability of 'Journalen' and the research field of usability evaluation in general.

4.1 Satisfaction

According to the results as presented above, the system has achieved a high level of satisfaction among the users tested. The female users were found to be slightly more satisfied with the system. While the sample group for this study was small and may not be statistically significant to substantiate this claim, it corresponds to a prior study of eHealth trends in Europe which found that young women are the most active internet health users (Kummervold et al 2008). While satisfaction is used as a measure for usability according to TURF, it can still be debatable if it is a good measure for usability. It is important to understand that satisfaction is subjective and may be influenced by user experience, task complexity and use context and thus may be a measure of perceived usability rather than real usability of a system (Frøkjær et al 2000). However Nielsen and Levy (1994) have measured that there is a strong positive correlation (r=.53) between satisfaction and usability performance. They have said that when users have an easier time using a system, they tend to rate it better in satisfaction. The findings of this study confirm Nielsen and Levy (1994) assertion as the high task performance results also correlate with high user satisfaction ratings for 'Journalen'.

 In previous studies time saving, memorability, learnability, flexibility, ease of use etc. have been identified as major factors for user satisfaction with EHRs, Ozok et al (2014) which could also be the same factors for this study as all these were covered in the SUS which was used to measure satisfaction. However perceived security has come up as another factor that may affect user satisfaction with

EHRs as pointed out by one user. A cross examination of usability evaluation methods and frameworks finds little or no attention to perceived security on EHR usability. However this study has not investigated in detail the question of perceived security on usability of 'Journalen' and EHRs in general as it was mentioned by only one user and thus future studies will need to investigate this further.

4.2 'Usableness'

According to TURF, a system is usable if it is easy to learn, easy to use, and error tolerant. This can be measured by Task performance evaluation. The results show a high degree of 'Usableness' as expressed through high user task performance evaluation ratings on 'Journalen'. Since the tasks have been based on demo data, it is likely that different results could be achieved if users were tested on their real data, however due to ethical and legal limitations, it wasn't possible to conduct this. Worth noting is that demo data allowed users to be more task oriented, rational and honest rather than being emotionally attached which could make the results more accurate. 'Usableness' as a dimension of Usability is a new term used in TURF, other frameworks use effectiveness and efficiency. However they are all intended to measure ease of use. According to Nielsen and Levy (1994) ease of use or in this case 'Usableness' is a performance measure compared to Satisfaction which is a preference measure. However they argue that performance indicators will have an effect on the preference indicators. Findings from this study prove this. For example in this study, users who spent less time on tasks had higher satisfaction ratings for the system than those who took more time on tasks. While measuring 'Usableness' is straight forward with variables like time taken to complete tasks, task completion, error occurrences, path deviation, steps taken to complete tasks, there is a big debate in literature as to how many users can give an accurate result. In their mathematical model for finding usability problems, Landauer and Nielsen (1993) argued that at 5 test users, the cost benefit ratio begins to fall. Virzi, (1992) claims that 5 users had found 80% of problems in his Usability studies. Landauer and Nielsen (1993) argue that usability lab tests are tedious and could blow up the budget if a lot of test users are involved. These claims were highly criticized by Faulkner (2003) who argued that 5 participants were very few. He argues *"It is advisable to run the maximum number of participants that schedules, budgets,*

and availability allow... the more powerful argument for implementing software usability testing, then ,is not that it can be done cheaply with, say, 5 test users, but that the implications of missing usability problems are severe enough to warrant investment in fully valid test practices". In response Nielsen (2006) has stated that for Quantitative Usability studies, 20 test users give a reasonably tight confidence interval with a margin of error at +/- 19%. At 10 users he estimates the margin of error at +/- 27%. As an argument for this study, 10 users were the manageable number within the timeframe and scope of this study, thus selected.

4.3 Usefulness

The expert review has been used as a measure for this dimension of usability. The results from this review have shown potential usability problems that need improvement. There is consensus among Usability researchers that Experts can find far more usability problems than regular users. Hollingsed & Novick (2007). Nielsen (1992) has introduced the concept of 'double specialists'. These are experts with experience in Usability design but also the work domain of the system. For example experts in Usability and financial systems, Usability and Health systems etc.   He argues that these are more accurate between 81%-90% compared to regular usability experts' 74%-84%, Nielsen (1992). He says this is because double specialists have more specific knowledge on the work domain of the system under evaluation. The expert used in this study is a double specialist being a medical practitioner and an Informatics professor. Nonetheless it would have been more significant if multiple expert reviews were conducted for comparative reasons.

Also double specialists are in a position to go further beyond the ergonomics and look at the Usefulness of the system functions. Usefulness concerns with system functions being able to create value for users. Zhang & Walji, 2014 has stated that "*If the functionality or utility of an application is not useful, whether it is usable or not is irrelevant".* While the expert review results focused more on challenges other than the strengths however the expert commended the system and says it is useful and a step in the right direction.

5.0 CONCLUSION

In this paper, a systematic evaluation of 'Journalen' is conducted in order to evaluate the Usability of the system. This has been conducted according to the TURF framework for usability evaluation. The major findings show that to a greater extent, 'Journalen' meets the usability requirements according to the TURF framework. The users have shown a high level of satisfaction with 'Journalen' as measured by high user SUS scores on the system. While gender is not the focus of the study, however this study has given an insight into gender perspectives as females expressed more satisfaction with 'Journalen'. On the 'Usableness' as evaluated through the Task performance evaluations, results show a high degree of Usableness for 'Journalen'. On the Usefulness, the expert review has shown major usability issues like lack of minimalistic representation of data, poor language use and unmatched user expectations that need to be addressed for 'Journalen to meet optimal usefulness.

Due to the limitations in the quality of the demo data used and size of the sample group tested, the results from this study may not be very generalizable but provide an insight into the usability experience of 'Journalen'. Further representative studies can be conducted to get a broader understanding of the usability of 'Journalen' for example among elderly, minorities and migrants.

The contribution of this study is that it provides a methodical way of conducting 'Usability Evaluation in EHR' for such future studies and highlights improvements areas that need to be addressed for better Usability of 'Journalen' and EHR systems in general.

ACKNOWLEDGEMENT

REFERENCES

American Medical Association (2014). Improving Care: Priorities to Improve Electronic Health Record Usability. CCJ: 14-0462: PDF: 9/14.

Belden, J. L., Grayson, R., & Barnes, J. (2009). *Defining and testing EMR usability: Principles and proposed methods of EMR usability evaluation and rating*. Healthcare Information and Management Systems Society (HIMSS).

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4-7.

Faulkner, L. (2003). Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers*, *35*(3), 379-383.

Frøkjær, E., Hertzum, M., & Hornbæk, K. (2000, April). Measuring usability: are effectiveness, efficiency, and satisfaction really correlated? In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 345-352). ACM.

Hollingsed, T., & Novick, D. G. (2007, October). Usability inspection methods after 15 years of research and practice. In *Proceedings of the 25th annual ACM international conference on Design of communication* (pp. 249-255). ACM.

ISO/IEC. (2011). Systems and software engineering – Software product Quality Requirements and Evaluation (SQuaRE) – Software product quality and system quality in use models. (CD 25010.3), First Edition March 1, 2011.

ISO/IEC. 9241-14 (1998) Ergonomic requirements for office work with visual display terminals (VDT)s Part 14 Menu dialogues, ISO/IEC 9241-14: 1998 (E), 1998.

ISO/TR 20514 (2005), "Health Informatics – Electronic health record – Definition, scope, and context", ISO, Geneva, Switzerland.

Johnson CM, Johnston D, Crowley PK, et al. (2011) EHR Usability Toolkit: A Background Report on Usability and Electronic Health Records (Prepared by Westat under Contract No. 290-0900023I-7). AHRQ Publication No. 11-0084-EF. Rockville, MD: Agency for Healthcare Research and Quality. August 2011

Karsh, B.T. (2004). Beyond usability: designing effective technology implementation systems to promote patient safety. Qual Saf Health Care, 2004. 13(5): p. 388-94.
Kummervold, P., Chronaki, C., Lausen, B., Prokosch, H. U., Rasmussen, J., Santana, S., & Wangberg, S. (2008). E-Health trends in Europe 2005-2007: a population-based survey. Journal of medical Internet research, 10(4), e42.

Landauer, T. K., & Nielsen, J. (1993).A mathematical model of the finding of usabilityproblems.Interchi'93, ACM Computer–Human Interface Special Interest Group.

Menachemi, N. and. Brooks, R. (2006). Reviewing the Benefits and Costs of Electronic Health Records and Associated Patient Safety Technologies. Journal of Medical Systems. June 2006, Volume 30, Issue 3, pp 159-168

Myreteg, G. (2015) Cost-benefit evaluation of e-health services: acceptance and value creation are interactive forces. *Health Systems*, 4(3), 204-21

Nielsen, J. (1992). Finding usability problems through heuristic evaluation. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 373-380). ACM.

Nielsen, J. (2006). Quantitative studies: How many users to test. *Alert box, June*, *26*, 2006.
Nielsen, J., & Levy, J. (1994). Measuring usability: preference vs. performance. *Communications of the ACM*, *37*(4), 66-75.

Ozok, A. A., Wu, H., Garrido, M., Pronovost, P. J., & Gurses, A. P. (2014). Usability and perceived usefulness of personal health records for preventive health care: A case study focusing on patients' and primary care providers' perspectives. *Applied ergonomics*, *45*(3), 613-628.

Sauro, J. (2015). Measuring usability with the system usability scale (SUS). 2011. *Verfügbar unter: http://www. Measuringusability. Com/sus.php [22.04. 2013]*.

Scandurra, I., Hägglund, M., Persson, A., Åhlfeldt, R.-M. (2014). Disturbing or Facilitating? On the Usability of Swedish eHealth Systems 2013 in eHealth – For Continuity of Care, eds. C. Lovis et al. Stud Health Technol Inform. 2014; 205:221-5

Scandurra, I., Jansson, A., Forsberg-Fransson, M-L. & Ålander, T. (2015). Is 'online patient access to health records' a good reform? – Opinions from Swedish healthcare professionals differ. Accepted 2015-08-20. Procedia Computer Science. PROCS6824. PII: S18770509 (15)02749-0 doi 10.1016/j.procs.2015.08.614

Virzi, R. A. (1992). Refining the test phase of usability evaluation: How many subjects is enough? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *34*(4), 457-468.

Williams, F., & Boren, S. (2008). The role of the electronic medical record (EMR) in care delivery development in developing countries: a systematic review. *Journal of Innovation in Health Informatics*, *16*(2), 139-145.

Zhang. J, and Walji. M, (2011). TURF: Toward a unified framework of EHR usability. Journal of Biomedical Informatics. Volume 44, Issue 6, December 2011, Pages 1056–1067

Zhang. J, and Walji. M, (2014). Better EHR. Usability, workflow & cognitive support in electronic health records. First Ed: National Center for Cognitive Informatics and Decision Making in Healthcare; 2014.

APPENDICES

Appendix 1: The system usability scale

User ID: ………   Age: ……………….   Gender: …………………. Date: _ _ _ _ /_ _ /_ _

System Usability Scale

© Digital Equipment Corporation, 1986.

	Strongly disagree				Strongly agree
1. I think that I would like to use this system frequently	1	2	3	4	5
2. I found the system unnecessarily complex	1	2	3	4	5
3. I thought the system was easy to use	1	2	3	4	5
4. I think that I would need the support of a technical person to be able to use this system	1	2	3	4	5
5. I found the various functions in this system were well integrated	1	2	3	4	5
6. I thought there was too much inconsistency in this system	1	2	3	4	5
7. I would imagine that most people would learn to use this system very quickly	1	2	3	4	5
8. I found the system very cumbersome to use	1	2	3	4	5
9. I felt very confident using the system	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system	1	2	3	4	5

Appendix 2: The task performance evaluation form

User ID: ………   Age: ………………..   Gender: ……………….. Date: _ _ _ _ /_ _ /_ _

Task Name	Task description	Task result	Time taken (Seconds)	Evaluation rating- Circle suitable (1-Strongly disagree 5-Strongly agree)
Task1	Which doctor carried out the latest diagnose.			-Takes little mental effort 1 2 3 4 5 -Takes short time 1 2 3 4 5 –Took few steps 1 2 3 4 5 -Completed successfully 1 2 3 4 5 -Easy to find help 1 2 3 4 5 -Easy to remember 1 2 3 4 5
Task 2	Which medicine was bed to yo 5 Johanss			-Takes little mental effort 1 2 3 4 5 -Takes short time 1 2 3 4 5 –Took few steps 1 2 3 4 5 -Completed successfully 1 2 3 4 5 -Easy to find help 1 2 3 4 5 -Easy to remember 1 2 3 4 5
Task 3	Check the results of the test (Blod, urin eller annat vätskeprov) that was carried out on date (14/06/2005) and read the comment.			-Takes little mental effort 1 2 3 4 5 -Takes short time 1 2 3 4 5 –Took few steps 1 2 3 4 5 -Completed successfully 1 2 3 4 5 -Easy to find help 1 2 3 4 5 -Easy to remember 1 2 3 4 5
Task 4	Find the issue for which you were referred to care unit (Ortopeden 2, Skånes Region Skåne) universitetssjukvård, in referral (Röntgenremiss).			-Takes little mental effort 1 2 3 4 5 -Takes short time 1 2 3 4 5 –Took few steps 1 2 3 4 5 -Completed successfully 1 2 3 4 5 -Easy to find help 1 2 3 4 5 -Easy to remember 1 2 3 4 5
Task 5	Find the notes made by doctor (Britt Thunblom) between date (20/09/2015) and date (01/10/2015)			-Takes little mental effort 1 2 3 4 5 -Takes short time 1 2 3 4 5 –Took few steps 1 2 3 4 5 -Completed successfully 1 2 3 4 5 -Easy to find help 1 2 3 4 5 -Easy to remember 1 2 3 4 5
Task 6	Find your last vaccination and read about its side effects			-Takes little mental effort 1 2 3 4 5 -Takes short time 1 2 3 4 5 –Took few steps 1 2 3 4 5 -Completed successfully 1 2 3 4 5 -Easy to find help 1 2 3 4 5 -Easy to remember 1 2 3 4 5

Task 7	Find all the vaccinations you took in the last year			-Takes little mental effort -Takes short time –Took few steps -Completed successfully -Easy to find help -Easy to remember	1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
Task 8	Find the dosage for medicine (Tavegyl)			-Takes little mental effort -Takes short time –Took few steps -Completed successfully -Easy to find help -Easy to remember	1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
Task 9	last appointment with Check your and (whom, date where)			-Takes little mental effort -Takes short time –Took few steps -Completed successfully -Easy to find help -Easy to remember	1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
Task 10	What was tfirst diagnose you h ? And what is date?the			-Takes little mental effort -Takes short time –Took few steps -Completed successfully -Easy to find help -Easy to remember	1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
Task 11	Give access to your journal to another user.			-Takes little mental effort -Takes short time –Took few steps -Completed successfully -Easy to find help -Easy to remember	1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5
Task 12	Check your journal overview and see how many activities you had under 2006, which month and what sort of activities they were.			-Takes little mental effort -Takes short time –Took few steps -Completed successfully -Easy to find help -Easy to remember	1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5

Additional comments:

Note: The last two tasks were not used after finding they were not achievable or useful in the test environment.

Appendix 3: The Turf 4.0 Usability software interface
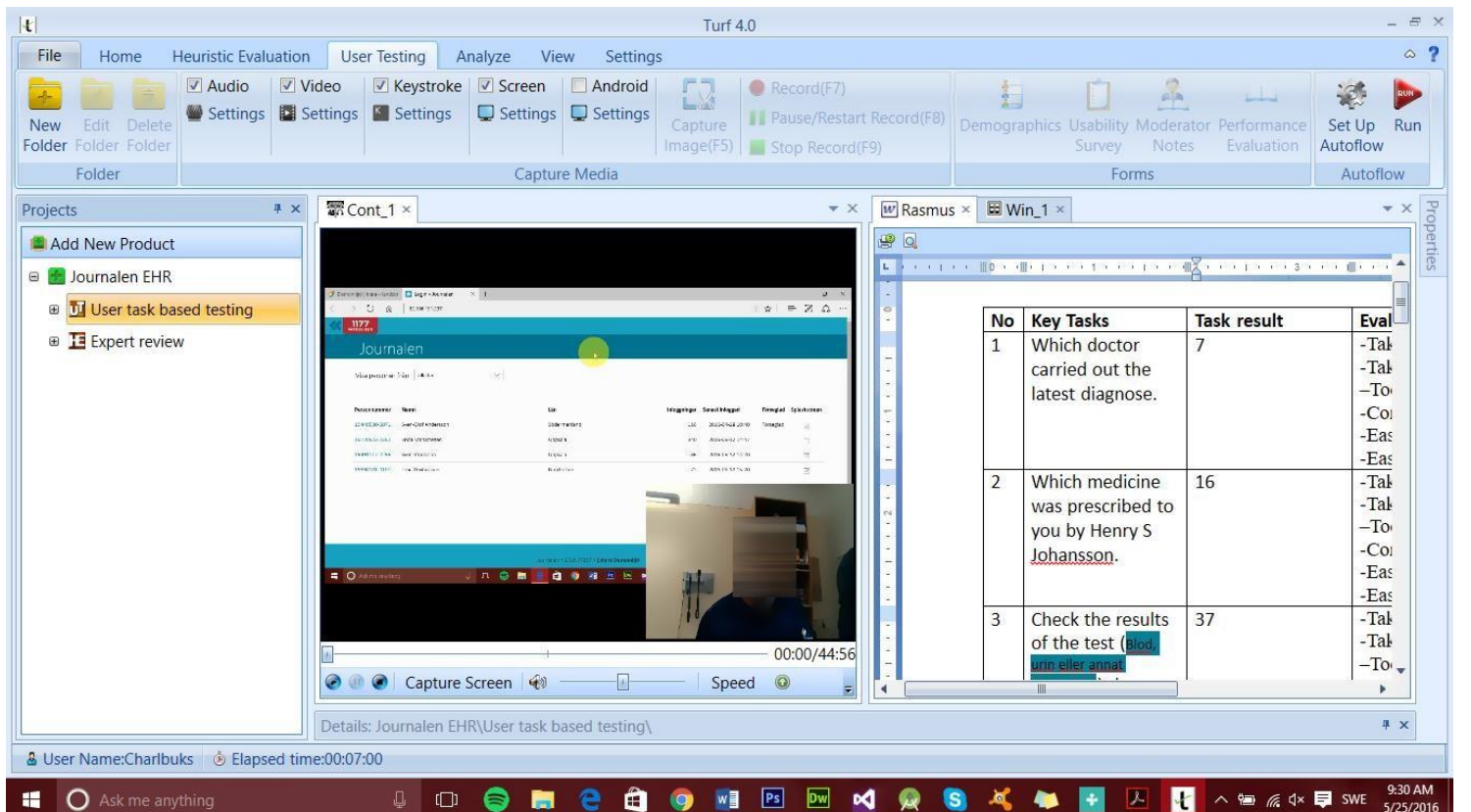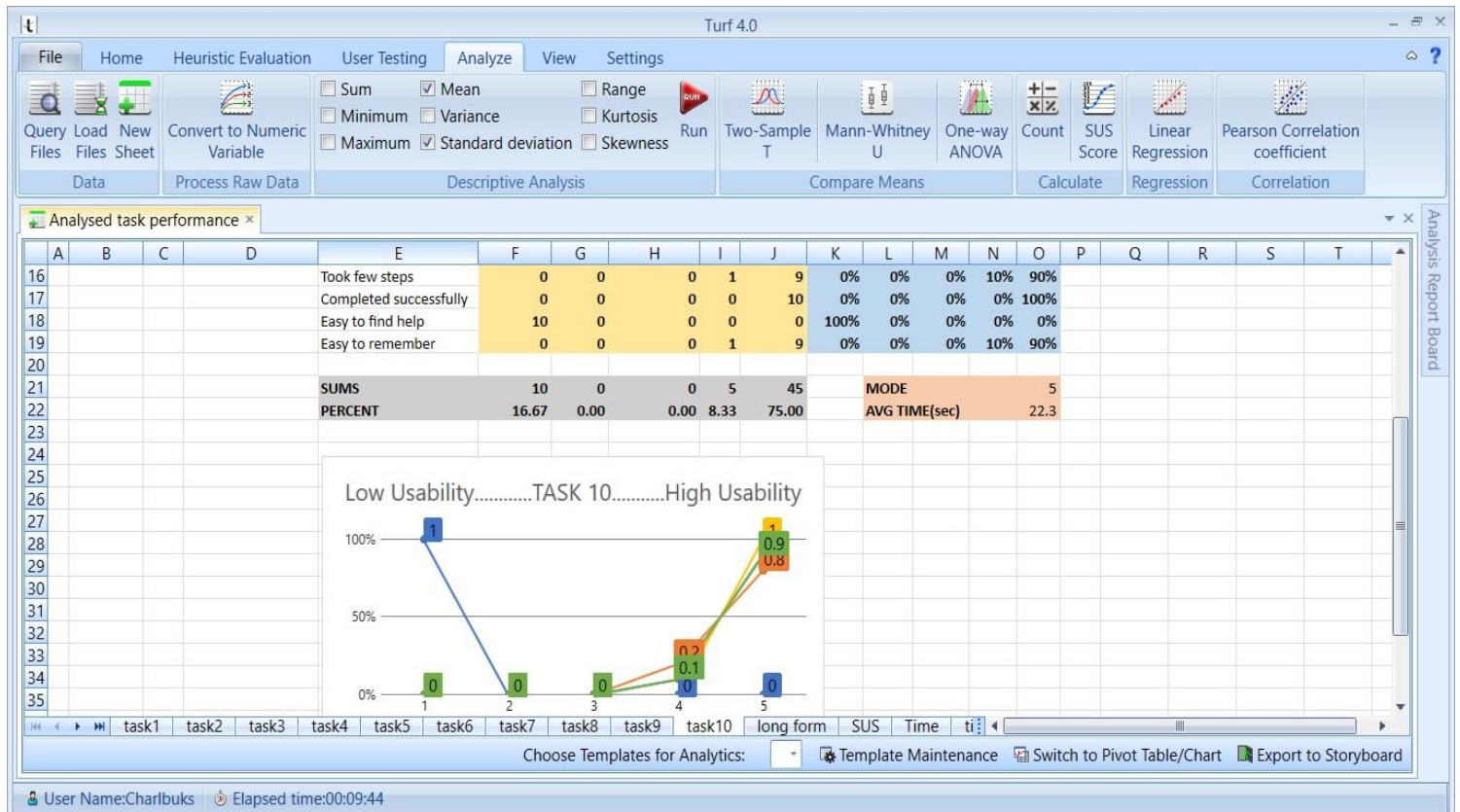
Appendix 4: Detailed results analysis for all tasks.

Attached as separate document

Appendix 5: Turf EHR Usability heuristics

1. [Consistency] Consistency and standards in design.
2. [Visibility] Visibility of system state.
3. [Match] Match between system and world.
4. [Minimalist] Minimalist design.
5. [Memory] Minimize memory load.
6. [Feedback] Informative feedback.
7. [Flexibility] Flexibility and customizability.
8. [Error Message] Good error messages.
9. [Prevent Errors] Prevent use errors.
10. [Closure] Clear closure.
11. [Undo] Reversible actions.
12. [Language] Use users' language.
13. [Control] Users are in control.
14. [Help] Help and documentation.

END