Human factors methods in health information systems’ design and evaluation: The road to success?
Peute, L.W.P.
Chapter 2

The Significance of a Usability Evaluation of an Emerging Laboratory Order Entry System

Linda W.P. Peute, Monique W.M. Jaspers


Abstract

Objectives: To assess the usability of an emerging POE system, OM/Lab, for the computer-supported ordering of laboratory tests. We were more specifically interested in the relation of the usability problems detected in the user testing sessions with the order behaviour in terms of efficiency and errors in ordering. Methods: A cognitive walkthrough of the OM/Lab system was conducted by two analysts using four real-life scenarios for ordering laboratory tests, which were reviewed for comprehensiveness by an expert clinician. Thereafter, the OM/Lab system was evaluated on its usability in testing sessions with seven potential end users of the system performing these same four scenarios. The results of these end-user testing sessions were used to analyze the effect of usability flaws on the quality of ordering in terms of omissions, errors in orders and cancelled orders. Results: The analyses revealed a total of 33 usability problems, of which 25 problems were revealed both by the cognitive walkthrough and in the end-user sessions. These 25 usability problems indeed led to inefficient order behaviour, omissions and errors in orders and even to cancelled orders. Discussion: Our results revealed that the OM/Lab system suffered from a high number of usability flaws. The interface design flaws were, among other things, related to misallocation of buttons on the screen, incomprehensibility of button labels and feedback containing no relevant information to the user about the cause of errors made and consequences of a user’s action. Additionally, our user test session results indicated that the OM/Lab system also suffered from user interaction problems of a more socio-technical nature. These sessions revealed, among other things, that the more specific action sequences to be executed within the ordering model of the OM/Lab system did not correspond to the daily working routines of end-users and that the grouping of laboratory tests within clusters did not match the paper-based order forms. The seemingly negative effects of these usability flaws on the quality of ordering and the inefficiency of work processes during the pilot implementation finally lead to withdrawal of the OM/Lab system from clinical practice. Though the system implementation failed, our usability study brought research on (re)designing and evaluating clinical computer applications at a higher status in our institution. It is now recognized that usability evaluation studies that will support good quality of clinical practice are highly important.

Keywords: Physician Order Entry Systems; Usability Evaluation; Information Systems.


Introduction

The potential benefits of Computerized Physician Order Entry (CPOE) systems in enhancing the quality and efficiency of health care delivery are well recognized. The recognition of these benefits by health care organizations has led to numerous attempts of CPOE implementations of which however only few have succeeded [1-6]. CPOE implementations that failed their expectations have nevertheless shed light on the causes underlying these failures, among which clinicians’ discontentment with the CPOE system has been proven to be a main factor [7]. Clinicians’ adoption and usage of CPOE systems has been hampered by their poor design. A thorough understanding of the fundamental principles behind human activities is required to design systems that are usable in supporting these activities [8, 9]. Usability is thus critical to successful CPOE implementation and adoption by clinicians. As defined by the International Standards Organization, usability refers to the effectiveness, efficiency and satisfaction with which users can achieve specific sets of tasks in a particular environment [10]. User effectiveness relates to accuracy and completeness with which specified users can achieve specified goals in particular environments. User efficiency refers to the resources expended in relation to the accuracy and completeness of goals achieved. User satisfaction relates to the comfort and acceptability of the system to its users. Usability studies on CPOE have however found these factors to be highly interrelated. These studies have shown that user satisfaction with CPOE system is highly correlated with ease of use of CPOE systems (effectiveness) and with productivity and system response time (efficiency) [11, 12]. Hence, physicians would be more willing to accept CPOE if the system is fast and easy to use, that is, if the user interface behaves consistently and meets the requirements of the physicians [11, 13]. Ease of CPOE system use is influenced by the ease with which its functions can be learned and memorized by users. Since the functionality of a CPOE system is made available through its user interface, its design has a huge influence on its usability [14-18]. A cumbersome user interface may not only slow down users but introduce a new class of cognitively based errors into the process [19]. Inconveniently displayed, not easily noticeable or obscured data may cause users to make erroneous decisions or to fail to take corrective actions [20]. For the human computer interaction being effective, the tasks and procedures that the user may perform with the system need to be structured in a logical and consistent manner. This means that the system functions should correspond one to one with the goals that the user sets him self in performing his tasks and the order in which the user wants to attain these goals. Furthermore, the number of actions that a user has to perform in accomplishing a task should be minimized. Also, the presentation order of information should match the order in which a user processes this information [21]. The representational structure of displayed information indeed determines the level of cognitive complexity any given task will require [22-26] and thus information should be displayed in such a way
that human perception of it is promoted. Otherwise, inefficient order behaviour, omissions and errors in orders may occur [27].

A recent trend in evaluating health care applications on their usability is the application of cognitive analysis methods of human-computer interaction with the aim to enhance a system’s user interface. Besides, in these studies, the relation between hypothesized, actual usability problems and the occurrence and frequency of errors that result from suboptimal user interfaces is explored [19, 26]. These studies have shown that cognitive analysis methods are not only suited to characterize user-interaction flaws in a system but also to identify (the opportunities for) errors attributable to these usability problems. In evaluating CPOE systems, use of these methods may focus on analyzing the ease with which physicians can order tests, on analyzing to what extent the system functions are structured in a, for the end-user, logical and consistent manner, on analyzing whether system terminology follows terminology used in daily practice and on analyzing whether the system provides adequate feedback in case of errors [19, 20].

A variety of cognitive analysis methods and techniques are used to evaluate a system’s design, among which the cognitive walkthrough and the think aloud technique [28, 29]. Cognitive walkthrough is a usability inspection method, based on evaluation of computer systems by an analyst, whereas the think aloud technique has been applied in usability testing sessions, based on system evaluations by true or potential end-users of a system. Cognitive walkthrough applies general principles from cognitive psychology and is to be conducted by an analyst who simulates the cognitive processes and user actions needed to accomplish a certain computer-supported task [30, 31]. The cognitive walkthrough method is conducted by use of predefined scenarios and focuses on the ease with which action-sequences can be executed in the system. The think aloud method formally belongs to the verbal report methods and stems from the field of cognitive psychology. It was specifically developed to gather information on the cognitive behaviour of humans performing tasks. This technique has proven to be of high value for evaluating a system’s design on usability flaws and is therefore frequently used to gather information about a system’s usability in testing computer systems with potential end-users. Users ‘interact’ with a system or prototype interface during recorded usability sessions according to a predetermined set of scenarios while verbalizing their thoughts [32].

Several studies have found that approximately 1/3 of a system’ usability problems are identified by all methods. Different methods have a unique complementary value in discovering usability flaws [33, 34]. Usability inspection methods focus on how intuitive an application is from a general human cognitive perspective, whereas usability testing methods focus on whether the application is compatible with user’s activity and with the cognitive aspects of that activity in particular [35, 36]. It is therefore highly recommended to combine both usability inspection and testing methods in a system usability study.
The study presented in this paper is concerned with the usability evaluation of a working prototype CPOE system for ordering laboratory tests (OM/Lab) that in a pilot was to be implemented at the outpatient neurology clinic and the outpatient laboratory of the Academic Medical Center (AMC) – Amsterdam. If the OM/Lab pilot implementation would prove to be a success, the system would be implemented at other clinical departments of the AMC.

We were asked to evaluate the prototype OM/Lab system. The main objective of this evaluation was to assess whether the OM/Lab system complied with the user requirements and to reveal potential usability flaws in the system. Besides these aims we were also interested in the relation of the usability problems detected in the user testing sessions with end-users’ order behaviour in terms of omissions, factual errors in orders and cancelled orders.

OM/Lab system background

The planning for the development and the implementation of the OM/Lab system followed a standard software design cycle. From both departments to be concerned in the pilot implementation of OM/Lab (the outpatient laboratory and the outpatient neurology clinic) the head of the department was involved in the project as representative and as OM/Lab acceptant. During the software prototype phase the request to evaluate the OM/Lab system was brought in for the choice of the 'one-page' human computer interaction model implemented in OM/Lab. Use of this 'one-page' model, offering users a total overview of the entire tasks to be executed in the system, proved to lead to better understanding by users of their progress in the system as opposed to a ‘page browsing interaction’ model [37]. This ‘one-page’ model had therefore been much discussed by the OM/Lab project team as an interaction model not only for ordering laboratory tests, but also for ordering of other tests such as radiology images. It was only after the prototype phase that the project implementation team asked us to evaluate the OM/Lab prototype system. So we were not involved in any earlier phase of OM/Lab development.

The ‘one-page’ user interface of the OM/Lab system (see Figure 1) consists of a six-step model with six bars containing drop down panes, one for each of the following: (1) Entering information identifying the requester, (2) Selecting laboratory protocol based order sets, (3) Ordering individual laboratory tests, (4) Entering patient data, (5) Entering a research question and (6) Sending the order. In step 1 the requesting clinician enters personal information, such as his name, and telephone number. Step 2 ‘Selecting a protocol’ offers the user the opportunity to request laboratory tests by choosing a protocol based order set for a specific neurological disorder. For example the protocol 'TIA' Trans Ischemic Attack defines 14 laboratory tests that should be ordered by the physician to assure if the patient has had a TIA. By choosing the TIA protocol the user does not have to select all these 14 tests individually by searching for each of these tests in step 3 and selecting these tests one by one by clicking on it with the mouse. Instead, after the selection of a specific protocol, tests are automatically selected as a set. After the user has chosen a protocol, he may add more tests or deselect tests from the
protocol pre-selected test set. As described earlier, ordering tests by predefined order sets (protocols) may save time and may improve the quality of ordering. The OM/Lab order entry prototype offered the following 5 protocols: Cardio Vascular Accident (CVA), Young Stroke (YS), Trans Ischemic Attack (TIA), Poly Neuropathy (PN) and Carpal Tunnel Syndrome (CTS).

In Step 3 ‘Ordering laboratory tests’ users may select the required laboratory tests one by one (see Figure 1). Users may otherwise choose between different kinds of materials; blood, urine portion, collected urine and liquor, after which the system displays the relevant tests accordingly. These sheets are further subdivided in several tab sheets: general, hormones, vitamins, toxicology. Within a sheet the tests are alphabetically ordered according to types of lab tests such as ‘chemistry’, ‘haematology’, ‘coagulation’, ‘blood bank’ or ‘purchase tube’. Tests which are not commonly requested are only made visible after the user has clicked on the ‘extra tests’ button. In step 3 a clinician may specify a specific time and date for the order to be carried out and may indicate whether the order is /should be carried out by the clinician himself/the department nurse or by the laboratory personnel. Step 4 'Enter clinical information' is implemented as a type-in field in which a clinician may enter additional information concerning the patient. A user is however not allowed to skip step 4 if the laboratory personnel needs additional patient information concerning the analysis of a specific test ordered by the physician in step 3. For example, laboratory personnel may need to know whether a patient is a diabetic or not in analyzing a specific test. Step 5 'Enter research question' will only be activated by the system after a specific test is ordered in step 3 for which laboratory personnel needs to know the physician’s reason for selecting that particular test, such as requests for not frequently ordered and expensive tests. Step 6 'Sending the order' provides an overview of all tests as selected in step 3. In this step, physicians may review their selection of lab tests after which they can accept and send the order to the laboratory. Once the order is sent, a form is printed with a code, to be scanned at the laboratory. The patient receives this form, and brings the form to the laboratory.

Methods

In evaluating the “one-page” human computer interaction model of the OM/Lab system two analysts conducted a cognitive walkthrough. We specifically chose to apply this usability inspection method as it focuses on the identification of goals and of subsequent actions to be performed to accomplish tasks within the system, thereby analysing whether the users’ background knowledge would enable completion of these tasks. As the focus of our study was on evaluating whether the one-page model was intuitive in use from a human cognitive perception, we felt that applying the cognitive walkthrough method would be more informative than the application of a heuristic evaluation method, in which the focus is on evaluating a system’s design using a set of general heuristics or principles.
The cognitive walkthrough consists of a preparation phase and an execution phase. In the preparation phase, the analyst constructs a correct action sequence for the task to be computer-supported. The user-background of the end-users is then defined on the basis of their computer experience. Thereafter, the execution phase is started. For each action in the action sequence the analyst answers four questions: 1) Will the user try to achieve the right effect (goal setting), 2) Will the user notice that the correct action is available, 3) Will the user associate the correct action with the effect trying to be achieved, and 4) If the correct action is performed, will the user notice that progress is being made towards completion of the task? With the description of the user's computer experience in mind, the evaluator analyses whether each question leads to success or failure by stepping through the system. After all actions have been evaluated, the usability problems are merged into a list.

In this study, the cognitive walkthrough was likewise used to define a framework of sub-goals and related action-sequences per sub-goal to be performed in the system. An illustration of a goal and subsequent actions to complete the task of selecting non-frequently ordered laboratory tests in the OM/Lab system is shown in Table 1. The analysts independently used this framework to identify potential usability problems related to the action-sequences to be executed in the OM/Lab system. Each of these potential usability problems was then coded according to ‘Norman’s theory of action’ into ‘goal problems’ (the user tries to accomplish the wrong thing) or ‘action problems’ (the user...
would like to perform the correct action but does not know how) [38]. Subsequently all potential usability problems were discussed and given a severity classification, ranging from 1 (minor problem) to 5 (usability catastrophe). Our purpose of rating potential usability problems in terms of severity was to help us communicate prioritized usability flaws that would need fixing by the system designers.

Table 1. Example of a goal-action sequence in the OM/Lab system.

<table>
<thead>
<tr>
<th>System’s state: Subtask (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Goal</strong></td>
</tr>
<tr>
<td><strong>Goal Action</strong></td>
</tr>
<tr>
<td><strong>Action</strong></td>
</tr>
<tr>
<td><strong>System response</strong></td>
</tr>
</tbody>
</table>

**Coded Potential Disparities:**

| Goal problem 1) | The function of the ‘extra tests’ button is not clearly defined in the system |
| Action problem 2) | End users might not notice the button ‘extra tests’ due to its invisibility on the screen. |
| Action problem 3) | After a user has clicked on the ‘extra tests’ button, the tests added are not visible on the screen. |

A think aloud usability test session starts with a preparation phase to familiarize the user with thinking aloud. Thereafter, the actual user test sessions start in which users then verbalize their thoughts in performing the task supported by the system. These sessions are preferably videotaped [31]. After all user sessions have been completed, the evaluator produces a complete list of the usability problems detected through review of the video-tapes and verbal protocols. In this study, these user sessions took place in the clinical environment of the potential end users and lasted 40 to 50 minutes. Sessions were videotaped; screen sequences of the OM/Lab system were captured for subsequent analysis. The verbal protocols of end users were transcribed and physicians’ utterances were analysed for usability problems encountered in these testing sessions. Besides, the user sessions were also analyzed to characterize the ways in which these usability problems negatively influenced the clinicians’ task behaviour in terms of omissions in orders, errors in orders or cancelled orders.
Subjects and materials

Four real life clinical scenarios, which were reviewed by an expert clinician for comprehensiveness, were constructed and were used in both the cognitive walkthrough and end-user testing sessions to assess the usability of the OM/Lab system. The scenarios were designed in such a manner that they overall covered all action-sequences to potentially be followed by end-users ordering tests by use of OM/Lab. Standard practice laboratory protocols were used as input for three of the four scenarios. Seven potential end users of the OM/Lab order entry system, three neurologists and four neurologists in training, were asked to perform these four simulated laboratory-ordering tasks supported by the OM/Lab system. The test users differed in age and computer experience. They all worked with the paper-based laboratory ordering forms in use at the outpatient neurology clinic. None of the end users had prior experience with a system for ordering laboratory tests; they had not been involved in the OM/Lab system requirement analysis phase, nor in its development phase.

Results

Cognitive walkthrough: potential usability problems

The cognitive walkthrough identified six subtasks and 29 associated actions that needed to be executed by a user in order to enter and send an order. The six subtasks included: (1) Check name and phone number of requester, (2) Select a protocol, (3) Select laboratory tests, (4) Enter patient clinical data, (5) Enter reason for ordering laboratory tests, and (6) Send the order to the outpatient clinical laboratory. These subtasks represent the six states of the 6-step model of the OM/Lab system (see Figure 1). For example, an action associated to the subtask of selecting laboratory tests is: click on the drop down menu ‘material’. In executing a relatively simple order, that is, ordering one laboratory test with the help of the system, only four of these six subtasks are to be performed and 13 actions. In executing a relatively complex order all six subtasks of the system need to be performed and all 29 associated actions, of which some are to be performed several times.

The in-depth cognitive walkthrough analysis of the OM/Lab user interface revealed a total of 25 potential usability problems associated with actions to be performed in executing the 6-step model. Table 2 provides the results if the cognitive walkthrough analysis, the severity of problems detected and the number of encounters by end-users in the Think Aloud user testing.
### Table 2. Results of the cognitive walkthrough (verified by the user testing sessions).

<table>
<thead>
<tr>
<th>CW Code</th>
<th>Characterization potential usability problem</th>
<th>Severity</th>
<th># Users</th>
<th># Encounters.</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A.202</td>
<td>Colour and location of the term ‘change’ on a bar may confuse users</td>
<td>1</td>
<td>4 / 7 *</td>
<td>Mentioned 5 times</td>
<td>Inefficient order behaviour</td>
</tr>
<tr>
<td>1. A.203</td>
<td>The required action sequence for opening step 1 might be difficult for end-users to comprehend.</td>
<td>2</td>
<td>3 / 7</td>
<td>3 / 28 **</td>
<td>Inefficient order behaviour</td>
</tr>
<tr>
<td>1. C.203</td>
<td>An inexperienced user may not know if nor how he can change the telephone number.</td>
<td>1</td>
<td>2 / 7</td>
<td>3 / 7</td>
<td>Errors of omission</td>
</tr>
<tr>
<td>2. A.203</td>
<td>Inflexibility of selecting more than one protocol. Cancellation of a selected protocol set in the system is not possible. This may lead to inefficient order behaviour.</td>
<td>3</td>
<td>3 / 7</td>
<td>4 / 7</td>
<td>Errors of omission Cancelled orders</td>
</tr>
<tr>
<td>3. A.201</td>
<td>Visibility and location of dropdown pane for selecting specific laboratory material.</td>
<td>2</td>
<td>2 / 7</td>
<td>3 / 28</td>
<td>Errors of omission</td>
</tr>
<tr>
<td>3. B.102</td>
<td>The understandability and naming of labels of laboratory tests may result in inefficient search strategies.</td>
<td>4</td>
<td>4 / 7</td>
<td>Mentioned 6 times</td>
<td>Factual errors Errors of omission</td>
</tr>
<tr>
<td>3. B.201</td>
<td>The chosen subgroups for laboratory tests on the screen may lead to difficulties in locating tests as it is in a substantially different format as the paper laboratory form.</td>
<td>2</td>
<td>2 / 7</td>
<td>13 / 28</td>
<td>Inefficient order behaviour Factual errors</td>
</tr>
<tr>
<td>3. B.202</td>
<td>Locating and memorizing particular lab tests might be difficult for end-users due to their invisibility, by use of the scrollbar.</td>
<td>4</td>
<td>7 / 7</td>
<td>298 / 28</td>
<td>Inefficient order behaviour</td>
</tr>
<tr>
<td>3. D.202</td>
<td>The location of labels on tabloids and different naming of tabloids might be difficult for end-users to notice when searching for tests in the system.</td>
<td>2</td>
<td>3 / 7</td>
<td>12 / 28</td>
<td>Inefficient order behaviour</td>
</tr>
<tr>
<td>3. E.103</td>
<td>The function of ‘Extra Tests’ button is not clear and might not be understood by end-users.</td>
<td>3</td>
<td>3 / 7</td>
<td>3 / 7</td>
<td>Errors of omission</td>
</tr>
<tr>
<td>3. E.202</td>
<td>The button ‘Extra Tests’ on the screen might be difficult for end-users to notice due to its invisibility and unclear label.</td>
<td>4</td>
<td>7 / 7</td>
<td>4 / 7</td>
<td>Inefficient order behaviour Errors of omission</td>
</tr>
<tr>
<td>3. G.101</td>
<td>End-users might deselect tests both in cito as in the general sheet due to the display of the filled in selection boxes in both sheets after selection of a test only in the general sheet.</td>
<td>5</td>
<td>7 / 7</td>
<td>13 / 14</td>
<td>Inefficient order behaviour Factual errors</td>
</tr>
<tr>
<td>3. G.103</td>
<td>End-users might unsuccessfully try to order a test in cito as the action sequence required by the end-user is tedious.</td>
<td>4</td>
<td>7 / 7</td>
<td>Mentioned 4 times</td>
<td>Factual errors</td>
</tr>
<tr>
<td>3. J.203</td>
<td>End-users might be confused due to unclear error messages providing no feedback</td>
<td>4</td>
<td>6 / 7</td>
<td>14 / 28</td>
<td>Inefficient order</td>
</tr>
<tr>
<td></td>
<td>Concerning the user action.</td>
<td></td>
<td>Behaviour</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. K.203</td>
<td>The fill-in box for entering a date for the execution of the order might not be noticed by end-users because of its screen allocation and missing error message if no information is entered.</td>
<td>3</td>
<td>‘All 7 end-users did not notice nor filled in the date’</td>
<td>Errors of omission</td>
<td></td>
</tr>
<tr>
<td>3. L.103</td>
<td>The fill-in box for entering a time to execute the order might not be noticed by end-users because of its screen allocation and missing error message if no information is entered.</td>
<td>3</td>
<td>‘All 7 end-users did not notice nor filled in the time to execute the order’</td>
<td>Errors of omission</td>
<td></td>
</tr>
<tr>
<td>3. M.103</td>
<td>End-users might not indicate who is to perform the order, for this is no standard option on the known paper ordering form and the system does not signal the user to make a selection.</td>
<td>2</td>
<td>2 / 7</td>
<td>2 / 7</td>
<td>Errors of omission</td>
</tr>
<tr>
<td>4. A.103</td>
<td>The action sequence required by an end-user to fill in clinical information is hard to understand and might lead to inefficient order behaviour.</td>
<td>2</td>
<td>7 / 7</td>
<td>26 / 28</td>
<td>Inefficient order behaviour</td>
</tr>
<tr>
<td>4. B.103</td>
<td>End-users may not go on to step 5 in the system if the end-user does not fill in information in step 4 when the system request additional patient information for a specific test.</td>
<td>3</td>
<td>2 / 7</td>
<td>3 / 28</td>
<td>Inefficient order behaviour</td>
</tr>
<tr>
<td>5. A.102</td>
<td>Inefficient action sequence in step 5 leads to too many actions to be performed by end-users.</td>
<td>5</td>
<td>5 / 7</td>
<td>16 / 21</td>
<td>Inefficient order behaviour</td>
</tr>
<tr>
<td>5. A.103</td>
<td>The user does not associate the goal ‘filling in research question’ with the correct action to complete the task in the system.</td>
<td>5</td>
<td>4 / 7</td>
<td>11 / 21</td>
<td>Errors of omission</td>
</tr>
<tr>
<td>5. C.104</td>
<td>No consistent system reaction to go to next step in the system, if required information has not yet been entered by the user.</td>
<td>5</td>
<td>7 / 7</td>
<td>16 / 21</td>
<td>Inefficient order behaviour</td>
</tr>
<tr>
<td>5. D.101</td>
<td>The go-on button might be clicked by end-users before finishing the ordering step, due to its invisibility and confusing label terminology.</td>
<td>2</td>
<td>4 / 7</td>
<td>11 / 21</td>
<td>Inefficient order behaviour</td>
</tr>
<tr>
<td>6. A.202</td>
<td>User might not be able to change selected tests in the system, for the system correct action sequence to fulfil this goal might be difficult to see for end-users.</td>
<td>2</td>
<td>3 / 7</td>
<td>5 / 28</td>
<td>Factual errors</td>
</tr>
<tr>
<td>6. C.103</td>
<td>The button start over’ might be incorrectly associated with the goal to add another order for the same patient, due to its location and label terminology.</td>
<td>2</td>
<td>1 / 7</td>
<td>1 / 28</td>
<td>Cancelled orders</td>
</tr>
</tbody>
</table>

* = 4 out of 7 end-users  
** = encountered errors 3 times in 28 scenario’s
Cognitive walkthrough: results verified by think aloud sessions

All potential problems detected by the cognitive walkthrough method were indeed encountered by end-users during the think aloud usability tests. In the next section, examples of the most severe and most frequently encountered usability problems, revealed by the cognitive walkthrough and verified by the think aloud analysis, will be described per subtask.

Subtask two ‘Select a protocol’

Inflexibility of protocol selection: The cognitive walkthrough analysis characterized this usability problem as the inflexibility of reversing a protocol selection in the system. The think aloud usability test showed that two end users choose a wrong protocol and after discovering their mistake were not able to undo their selection. Instead of restarting the system, one end user deselected all of the system’s protocol based pre-selected set of 16 laboratory tests, before selecting the set of laboratory tests he had decided on. Thus, the inability to undo the selection of a protocol within a laboratory order forced users to inefficient and time-consuming ordering behaviour.

Subtask three ‘select laboratory tests’

Inability to navigate to laboratory tests in the system: The cognitive walkthrough detected a potential usability problem referred to the invisibility of laboratory tests on the screen and the time spent in selecting laboratory tests by use of the scroll-bar (see Figure 1). In performing the four scenarios, the seven end users clicked the scrollbar 298 times. Because searching of tests with use of the scrollbar arrows took additional time, all end users were dissatisfied and complained about system performance. Moreover, though all tests were alphabetically organized the users on average still needed two and a half minutes to find the test they were searching for. Five of the seven users verbalized more than once, while searching for a test in the OM/Lab system, that the laboratory tests in the program were not organized as they were used to with the paper ordering forms. They all expressed that the system was needlessly inefficient and that they needed a search function by which they could find tests more easily, for example by entering the first letter of a test or a test synonym.

Invisibility of button labels and incomprehension of button function: For ordering infrequently ordered laboratory tests a button ‘extra tests’ had to be selected in the system. After activation of this button additional laboratory tests appeared on the OM/Lab screen. The cognitive walkthrough analysis revealed that due to the location of this button on the screen and its name label, users might not notice nor comprehend the function of this button, and might therefore not find the tests additionally supplied by the system after selection of this button. During the think aloud sessions end users indeed experienced severe difficulties in ordering these laboratory tests; they did not notice this button on the screen and if they did, they did not comprehend what its function was. The button was only found in
three of the seven times it should have been selected. Most end users complained that many laboratory
tests were difficult to find and that selecting a test required too many actions by the user. Because of
this problem in ordering tests, several errors of omissions were made by the users in the usability
testing sessions.

*Appearance and information content of error messages in the system:* The cognitive walkthrough
revealed that most error messages in the system provided the user no information concerning the cause
and effects of his preceding action. In the think aloud testing sessions, in 14 of the in total 28 scenarios
that were to be performed by the seven end users, system error messages were presented to them. All
users appeared to be distressed and confused by these messages, leading to inefficient ordering
behaviour by the end-users.

*Understandability of required sequence of actions:* The cognitive walkthrough revealed that the
sequence of actions to be performed by end-users to order tests cito (tests that require results in a
shorter time frame) was inefficient and could be confusing to them. After an end-user had selected a
laboratory test on the general sheet, this test would also be given a mark and grey label on the (pink
coloured) cito sheet. However, this mark on the cito sheet was only to indicate to an end-user that he
or she had previously selected the test on the general sheet, though not in cito. If a user would however
like to indicate that this specific test needed to be performed cito, he or she would first have to deselect
the test by clicking on its box on the cito sheet, and then select the test again by clicking on the same
box once more, after which action the colour of this specific test label was changed from grey into
pink. The cognitive walkthrough analysis found this action sequence to be highly inefficient. Indeed,
during the user test-sessions none of the seven end users executed the required sequence of actions in
the OM/Lab system to order tests cito. The think aloud analysis provided insight into a different
problem though related to the understandability of required system actions. Users misinterpreted the
marked and grey labelled tests on the cito sheet and believed that all these tests were already ordered
cito, though in fact they were ordered but not yet cito. Though the system designers had deliberately
decided to use a highly noticeable colour for the cito sheet (namely pink), the mixing of colours for
labels (grey for tests ordered in the general sheet, pink for tests ordered cito) disoriented all users as
they interpreted grey marked tests on the cito sheet as if these tests were to be performed cito. They
therefore deselected those tests for which it was not their intention to have them performed cito. This
resulted in missing laboratory tests on each of the seven occasions it could have occurred.

*Subtask 4 ‘Enter clinical patient data’*

*System flexibility:* To analyze certain lab tests, laboratory personnel needed additional patient
information. Clinicians had to indicate, for example, whether a patient was suffering from diabetes.
The cognitive walkthrough revealed however that if an end-user did not supply the additional patient
information required by the system in step four, the end-user could not go on to the next step in the
system. As the system offered no user feedback in the form of error messages on the reason for the need of entering this patient information, users indeed tried unsuccessfully to skip this task during the think aloud sessions. The end-users remarked several times “I don’t need to give additional information, I want to go on with sending the order”. Finally, they filled in “not necessary”, after which they could go-on with step five in the system. However, in the analysis of the test sessions these orders were regarded as incomplete orders, as in these cases the information the laboratory personnel (would have) needed to adequately perform tests had not been supplied by the ordering physician.

Subtask 5 ‘Enter reason for ordering laboratory tests’

Understandability of required sequence of actions/ User workload:

So that laboratory personnel could adequately process certain expensive or infrequently ordered laboratory tests, physicians needed to enter their reason for selecting these kinds of tests in the OM/Lab system. The cognitive walkthrough revealed that the system did not provide any feedback in support of the actions to be performed to successfully complete this step. Five of the seven users were indeed disoriented about the correct sequence of actions they needed to execute in entering a reason for ordering a specific laboratory test in the OM/Lab system. Users ineffectively tried to skip this step to accomplish sending the order. After on average executing the wrong action three times, four end-users finished this subtask correctly. However it seemed that this system feature was not easily learned, for an end user performed the same ‘ineffective’ action sequence several times in different task scenarios.

Think aloud analyzed usability problems

During the think aloud analysis eight specific additional usability issues arose which were not identified by the cognitive walkthrough. These problems were all revealed in the context of actions to be performed in the system that did not directly match the standard paper ordering routines. Five of these usability problems are discussed in the next paragraph.
Table 3- Usability problems additionally revealed by the think aloud analysis.

<table>
<thead>
<tr>
<th>Step 2: TA 1</th>
<th>Locating required laboratory tests after selection of the ‘extra test’ button because of reorganization of laboratory tests.</th>
<th>Does not adhere to working practices</th>
<th>Errors of omission</th>
<th>7/7*</th>
<th>12/14**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 3: TA 2</td>
<td>The end-user’s goal to send an order in the system is prematurely as opposed to the additional tasks to be performed in the system. The text on button ‘start over’, confused end-users’.</td>
<td>Does not adhere to working practices</td>
<td>Cancelled orders</td>
<td>2/7</td>
<td>3/28</td>
</tr>
<tr>
<td>Step 3: TA 3</td>
<td>Disconformities in ordering protocol based laboratory order sets. Users order additional tests caused by interpretation differences of the paper based protocols in clinical practice.</td>
<td>Does not adhere to working practices</td>
<td>Factual errors in orders by adding tests</td>
<td>5/7</td>
<td>12/21</td>
</tr>
<tr>
<td>Step 3: TA 4</td>
<td>System content mismatched laboratory test terminology as applied in clinical working practice thereby leading to the believe that some test were not included in the system though in fact they were.</td>
<td>System terminology</td>
<td>Factual errors</td>
<td>Errors of omission</td>
<td>6/7</td>
</tr>
<tr>
<td>Step 4: TA 5</td>
<td>The need to enter clinical information about the patient was not in accordance with the paper-based laboratory ordering process. In practice requesters did not supply patient clinical information on the paper order forms, this additional task in the system was not a goal the users sets himself in ordering tests in the system.</td>
<td>Does not adhere to working practices</td>
<td>Inefficient order behaviour</td>
<td>Errors of omission</td>
<td>7/7</td>
</tr>
<tr>
<td>Step 4: TA 6</td>
<td>The system did not support the users’ goal to enter additional information concerning a specific test, for example a patient’s weight or length. In practice this was implicitly solved by adding written text on the paper laboratory form next to a test to be ordered.</td>
<td>Does not adhere to working practices</td>
<td>Potential errors in practice: Inability to correctly analyze lab tests, Double orders</td>
<td>4/7</td>
<td>4/-</td>
</tr>
<tr>
<td>Step 5: TA 7</td>
<td>Forced entering of additional information concerning specific tests in step 5, after selection of a protocol based order-set, was considered as redundancy of information entering. End-users expressed that they had supplied this information by selecting a protocol.</td>
<td>Does not adhere to working practices</td>
<td>Factual errors Requested information was not supplied by end-users</td>
<td>7/7</td>
<td>21/21</td>
</tr>
<tr>
<td>Step 5: TA 8</td>
<td>Though the system was immature, users expected that in selecting a protocol based order set in step 2, all test according to the protocol would be ordered by the system, though some tests could not be ordered by use of the system.</td>
<td>Does not adhere to working practices</td>
<td>Errors of omission</td>
<td>3/7</td>
<td>9/21</td>
</tr>
</tbody>
</table>

* = 7 out of 7 end-users, ** = encountered errors 12 times in 14 scenario’s
Subtask three 'select laboratory tests' 

Difficulty in finding required laboratory tests:

After a user had clicked on the ‘extra tests’ button in step 3 of the OM/Lab system the lists of tests on the screen was extended with additional laboratory tests. However, the list of tests was then reorganized and regrouped leading to end-users experiencing difficulties in locating the correct test they had planned to order in the system. The cognitive walkthrough analysis did not reveal this reorganization of laboratory tests as a usability flaw as the analysts found the list of tests logically organized according to their alphabetic order, and after regrouping were still in alphabetic order. All seven end-users however experienced this reorganization of laboratory tests as a problem, as the list of tests was very large and some tests were after reorganization not directly visible on the screen. In reaction to this usability flaw, users skipped the selection of several tests as they claimed they could not find the test they needed. This resulted in errors of omission in 12 of the 14 orders this could have occurred during the users testing sessions. The users also commented on the tests arranged under the subgroup “blood tube” on the general sheet. They explained that, in order to parallel the paper based order forms, these tests should in fact have been grouped under a completely different tabloid that is under the tabloid ‘extensive coagulation’. This likewise contributed to the users’ experiencing severe difficulties in finding tests in the system.

Premature user goal setting of sending order in the system:

The system offered the possibility in step 3 to completely start over with the ordering process by clicking on the single button ‘start over’. Two end-users made the error to click the ‘start over’ button. The button ‘start over’ was interpreted by these users as ‘this order is ready, send the order and start a new order’. The two end-users who made this error also verbalized that they had selected the required tests and didn’t want to provide clinical information nor a research question. They were convinced that this was the next right action to send the order to the laboratory. In their clinical practice it was not customary to supply additional patient information. Also an error handling message in the system related to the button ‘start over’ in this case might have prevented 3 cancelled orders.

System terminology versus terminology used on the paper-based forms:

The names of several laboratory tests differed from the names used on the paper based laboratory forms. This was caused by the fact that the OM/Lab system automatically abstracted terms and names for lab tests from the ‘resource catalogue’ of the Laboratory Hospital Information System. Six of the seven end-users remarked that a test was not included in the system and could therefore not be ordered, though in fact the test could be ordered in the system but yet under a different name. This resulted in 10 omissions in orders.
Subtask 4 ‘Enter clinical patient data’

Self-descriptiveness of tasks in the system vs. users’ unawareness of additional system goal:

In the think aloud analysis all seven end-users verbalized that they did not understand the necessity of filling in clinical patient information in a laboratory order, as only they themselves would review the order results. Since physicians did not understand the reason for entering this extra information, they started searching for different routes in the system to send the order. However, this information was needed by the laboratory personnel for adequately analyzing tests. Since the end-users of the OM/Lab system in daily practice do not supply patient clinical information on the paper order forms, they were not willing to fill in any such information. This led to 10 occasions in which the required patient information was not provided by end-users.

Subtask 5 ‘Enter reason for ordering laboratory tests’

Redundancy of information vs. users’ unawareness of additional system goal:

Users, after having selected a specific protocol in step two of the system, were forced to add information for each of the tests ordered as a set by the protocol concerning the reason for its ordering. All seven end-users felt that they had supplied this information in step two just by selecting a protocol. Therefore, on nine occasions during the user sessions, three of the seven end-users deliberately omitted the requested information.

Discussion

In this study, we performed a usability evaluation of the OM/Lab system, an emerging CPOE laboratory system. Overall, the results of the usability evaluation revealed that the OM/Lab’s usability fell short of the designers expectations. Despite the “one-page” human computer interaction model implemented in the OM/Lab system, which seemed an intuitive model to follow in ordering laboratory test sets, our evaluation revealed a total of 33 usability flaws in the system, which all resulted in insufficient or erroneous order behaviour in the user test sessions. Though the system designers had felt that the six-step model could facilitate completion of the task of test ordering in a logical manner, many of the action-effect relations of the OM/Lab system were not transparent to the evaluators and the actual end users; in fact many of the system action sequences impeded rather than facilitated the ordering process.

As to an extent the design of the system influenced its low usability. Our usability inspection results, all verified in the user testing sessions, revealed a large number of human interaction problems in the OM/Lab interface that were of a general nature. These ‘surface’ human interaction problems of the OM/Lab system were related to inconveniently displayed buttons, incomprehensibility of button
labels, and feedback containing no information to the user about the cause of the error made and the consequences of a user’s action. Likewise the inflexibility of protocol selection and the cumbersome OM/Lab navigation structure for searching tests to be ordered were brought forth by conceptual design aspects deliberately chosen by the system designers. The poor perceptual cues of the interface and error messages without clarifying information content resulted in the users making faulty inferences regarding next actions required and hence triggered the users in applying inappropriate task strategies. These “surface” problems indicate that their existed mismatches between the designers’ and the users’ conceptualizations of the general task hierarchy (implemented as the six-step model in the OM/Lab system) of ordering lab tests. Yet, these problems could have been easily fixed before the OM/Lab pilot implementation effort.

Moreover, the analysis of the test sessions with end-users showed that the more specific action sequences to be performed within the general task hierarchy eq. six-step model of the OM/Lab system did not correspond to the daily working routines of end-users. These usability problems, of a more socio-technical nature, dealt with the order in which our end-users normally process information in their daily working practice in contrast to the way they had to process this information in the OM/Lab system. Further, in addition to the problem of the invisibility of laboratory tests on the OM/Lab screen, revealed in the usability inspection, certain laboratory tests were clustered in the system in such a way that these categories did not correspond with the way our clinicians tend to cluster tests in clinical practice whereas the paper-based order forms follow these clusters. This classification of tests in the system and within each of these classifications the alphabetic ordering of tests, considered logical by the system designers, did not facilitate but instead hampered end-users in searching for particular laboratory tests. A related usability problem of a socio-technical nature was the fact that users had to provide specific clinical patient information on system’s demand which did not correspond to their goal setting. End-users did not consider this information relevant for the ordering of laboratory tests. As explained earlier, during the requirement analysis phase of OM/Lab only one representative of the outpatient neurology clinic and one representative of the outpatient laboratory had been consulted about the design of the user interface of the OM/Lab system. The total lack of communication between these representatives might explain why the system lacked in its design as it caused the designers to fulfil user requirements expressed by the representative of the outpatient laboratory which demanded actions of the neurologists which thus far were not required in their daily practice of ordering laboratory tests, that is in ordering tests on paper-based forms. Presumably, the representative of the outpatient laboratory had thought to take full advantage of the OM/lab system development efforts by expressing additional information needs of laboratory personnel so as to enable them to perform tests more efficiently than in the paper-based situation. Yet, these information needs of laboratory personnel were not discussed with the neurologists who were to provide this patient information. As a
consequence the neurologists felt that they did not need to provide this information which in their opinion was redundant and irrelevant for their (own) working practices.

These more socio-technical usability aspects were not only the basic cause for omissions and errors in orders and cancelled orders in our evaluation study but also in the pilot implementation of OM/Lab. Due to time pressures these flaws were not repaired in the OM/Lab system before its pilot implementation. However, these flaws indeed brought about a disruption of work patterns during the pilot implementation, such as delays in the time it took to process orders by the laboratory personnel and the necessitated return of patients to the hospital in order to finish incomplete laboratory orders. During the implementation of OM/lab it became clear that clinicians still experienced difficulties in ordering tests. Among other things, they could not easily localize tests in the system and they did not provide the patient clinical information required by the laboratory personnel to perform laboratory tests adequately. The clinicians seemed not aware that they had to provide this information which led to a vast amount of unfinished and incomplete orders. This, in the end, contributed to rejection of the system by the neurologists as well as the outpatient laboratory personnel. Yet, most of these usability problems in OM/Lab could not have been fixed effortlessly as they were part of the basic database structure of the system, which had not been developed in-house but by a commercially based vendor.

It is well known that the domain knowledge and daily working experiences users bring to the interface may determine its success or failure [39, 40]. Ideally a user’s domain knowledge and working behaviour would be considered in an early stage of system design to analyze in more detail whether the use of certain design features, such as the order of action-sequences to be performed and related menu structures, the terminology used for labelling categories, menus, buttons, and system feedback contents do bear a resemblance to the task domain and related working routines under consideration. However due to the fact that we had not been involved in the requirement analysis phase of the OM/Lab system we did not have had sight on the neurologists’ and laboratory personnel’s daily working routines in ordering or performing laboratory tests, nor on their additional information needs or other requirements to be fulfilled by the OM/Lab system. In the context of the relatively high number and severity of the usability problems encountered in our usability evaluation, we can only but assume that if the requirements capture activity for OM/Lab would have been performed more profoundly, this would have resulted in a higher usability of the OM/Lab system.

Usability evaluation studies require the construction of task scenarios. The manner of how to construct these task scenarios however seems to be a somewhat neglected aspect in the description of how to apply usability methods; task scenarios are a fundamental prerequisite for performing usability evaluations in a structured manner but explanations of usability methods provide no explicit guidance on how to construct these scenarios [40]. Variables such as task selection and task coverage have yet been found to affect the success of usability evaluation methods; detailed task descriptions can significantly change the number and type of, both severe and less severe, usability problems found in a
system’s design [41]. In our study, we were able to construct task scenarios covering all action-sequences to potentially be followed by end-users ordering tests by using OM/Lab. This may have enhanced the discovery of usability flaws in our evaluation study. Usability evaluations of applications more complex than OM/Lab will however be more demanding in time and human resources and can probably not be evaluated on all their aspects in one test. As it is with most systems not practically possible to include all a system’s aspects in one evaluation test, it is important to thoroughly analyse and set the goals for usability evaluations and focus the construction of task scenarios the system is to be evaluated against accordingly. Usability evaluators should verify these task scenarios in a systematic way to ensure that all system aspects relevant in the context of the goal are covered by these task scenarios and will be explored in the usability assessment. The ultimate goal is to develop end versions of health information systems that map on users tasks and strategies in performing these tasks and trigger the cognitive dynamics of the user in such a way that the intended tasks can be accomplished with minimal cognitive effort [42].

This study aimed at evaluating whether the OM/Lab system fulfilled end-users’ requirements and in so doing providing the OM/Lab system developers with information on how to redesign the system. However, due to management implementation pressures and unavailability of additional resources, the AMC system implementation project leader decided that the major usability problems brought forward by our usability evaluation would not be fixed before the OM/Lab system would be implemented in a pilot phase at the outpatient neurology clinic at the AMC. After a six weeks pilot implementation period, OM/Lab was withdrawn by the project leader due to the inadmissible problems the system caused in clinical practice such as delays in the patient waiting time at the outpatient laboratory and the inability of physicians to work with the system. Its pilot-evaluation results showed that OM/Lab’s low usability directly instigated clinicians’ rejection of it. In the end, the ‘one-page’ model was aborted, and a new setting was chosen for the sequel of order management at the AMC.

The failure of the OM/Lab system was nevertheless exactly the reason for the AMC to assemble a new medical informatics research team on (re)designing and evaluating clinical computer applications in an early stage of system development. With this new approach to system development and implementation, our aim is to bring forth clinical systems that support clinicians to the optimum in their daily working routines and consequently enhance good clinical practice.

Acknowledgements

The authors thank M. Smeulers, J. de Gans, P. F. Mooijweer- Groen, T.L.F. Urbanus and P. J. M. Bakker for their cooperation in conducting this study.
REFERENCES


