



*Quintus will systematically review
your product or website for major
issues or barriers to use*

Usability Comparison between the UMK (ACOEMUMK.org) and ODG- Treatment (ODG-Treatment.com)

Prepared by Erika Noll Webb, Quintus Design LLC
For more information, please contact
ewebb@quintusdesign.com 303.449.5020

User Interface Usability Comparison between ACOEM's UMK (ACOEMUMK.org) and WLDI's ODG-Treatment (ODG-Treatment.com)

For business and research purposes, Reed Group, Ltd. was interested in understanding how ACOEM's Utilization Management Knowledgebase (UMK) product compared in relation to general system usability to a competitor's product, the ODG-Treatment website. Usability of a system has more to do with how a system works for a user than with how it looks. Usability is a measurement of the quality of a user's experience when he or she interacts with a system such as either of these. The quality of the interaction is comprised of a set of different factors including the ease of learning, ease and efficiency of use, memorability, errors (both frequency and severity) and, finally, the overall subjective satisfaction of using a system.

A system that is more usable reduces costs in a variety of manners, including money saved in reduced training. Increased efficiency translates to higher worker productivity. Fewer errors and less severe errors also are reflected in a company's bottom line.

For the purposes of this comparison, a scenario was constructed both to guide the testing along the lines of real-world use and to elicit similar responses from the UMK and the ODG-Treatment products. The 5-question scenario provides information that relates to one hypothetical employee's condition. The scenario is as follows:

1. An employee is diagnosed with the following ICD-9 code: 724.4. What is the condition's name?
2. The employee was diagnosed 14 weeks ago.
 - A. What diagnostic options for this employee should be considered?
 - B. Do the diagnostic options for this employee differ if the condition were diagnosed in only the last two weeks? What are the differences and how can I tell?
3.
 - A. When, along the treatment spectrum, is surgery for this condition an option to explore?
 - B. What diagnostics are required before surgery should be considered?
4. Does the tool provide a clear method of identifying and communicating the employee's diagnostic and treatment options across the range of acute and chronic stages (such as an employee diagnosed 14 weeks ago and an employee diagnosed 2 weeks ago)?
5. A printed copy of the employee's condition and care plan needs to be printed for record keeping. Is it easy for me to print what I need?

General Principles

In order to execute the scenario according to measurable events, a set of usability principles, or “rules of thumb,” were considered. The set used for this review was taken from Nielsen (http://www.useit.com/papers/heuristic/heuristic_list.html). The list was chosen for its applicability to the products compared.

- **Match between system and the real world**

The system should speak the user’s language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Systems should follow real-world conventions, making information appear in a natural and logical order.

- **Consistency and standards**

A user should not have to wonder whether different words, situations, or actions mean the same thing. Following platform conventions minimizes confusion.

- **Error prevention**

User-oriented designs eliminate error-prone conditions.

- **Recognition rather than recall**

Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

- **Flexibility and efficiency of use**

Accelerators—unseen by the novice user—may often speed up the interaction for the expert user with the result that the system caters to both inexperienced and experienced users. One result of accelerators is that the system tailors itself to each user's most frequent actions.

- **Aesthetic and minimalist design**

Dialogues should not contain information that is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

- **Help and documentation**

Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

Based on these principles, each enactment of the scenario was given a usability “score.” The score was based on the following scale:

- 4 = No usability problem**
- 3 = Cosmetic problem only**
- 2 = Minor usability problem**
- 1 = Major usability problem**
- 0 = Usability catastrophe**

The factors that influence the score include frequency and impact. Minor problems that occur frequently are given a lower (more severe) score than those that are minor and infrequent. Likewise, minor issues that cause significant increases in the time to complete the task can impact the overall effectiveness of the product and would also lead to a lower (more severe score) on this scale. Finally, problems that have only a slight impact on completion of the tasks are not rated as severely as those that potentially cause the user to make errors in the task.

Test conditions

A single tester reviewed both products using a computer with the XP operating system and Internet Explorer 6.0.2900.2180 with SP2. Times required to complete the tasks were tracked by stopwatch.

Testing Organization

Quintus Design is a consulting firm that offers services to measure and improve the quality of a user's experience when interacting with technology or a system. Core competencies include:

- * Usability
- * Surveys
- * Accessibility
- * Focus Groups
- * Training

Quintus has experience with a broad spectrum of high-tech products and services and can quickly provide relevant feedback to lower support costs, improve user efficiency, speed development time, and align with government requirements. All the consultants at Quintus have extensive training and educational backgrounds in the behavioral sciences and collectively over 50 years experience applying that expertise in high technology businesses. The result is a strategy for product design that balances a rigorous scientific approach with the pragmatic needs of the business environment. Quintus consultants have earned several awards and patents for their innovative designs.

Tester Unfamiliarity

The reviewer, Erika Noll Webb, had no prior exposure to the usability or user interface of either the Utilization Management Knowledgebase or Official Disability Guidelines products at the outset of this project.

Results

For each task, several measurements were taken. In addition to the 0 - 4 scale score listed above (the lower the number the less useable the interface), the total time to complete the task was measured for the standard-use case path. In addition, measurements included the number of clicks to new pages required to complete each task.

Time to complete

Figure 1 shows the mean time to complete each of the 5 tasks in the scenario. For the UMK product, the mean time to complete each task was 51.5 seconds (range of 23-86 seconds). Completing the same tasks in the ODG-Treatment product took more than 3 times as long, with a mean time to complete of 185 seconds (range of 49-304 seconds). In several cases, the information needed to complete the task did not appear to be available in the ODG-Treatment product.

Note: An apparent lack of available information may or may not indicate a user interface weakness. It may also indicate either that the product does not fulfill its business requirements or that its content falls outside of the selected scenario. Further investigation would have to be done to fully resolve the divergent explanations.

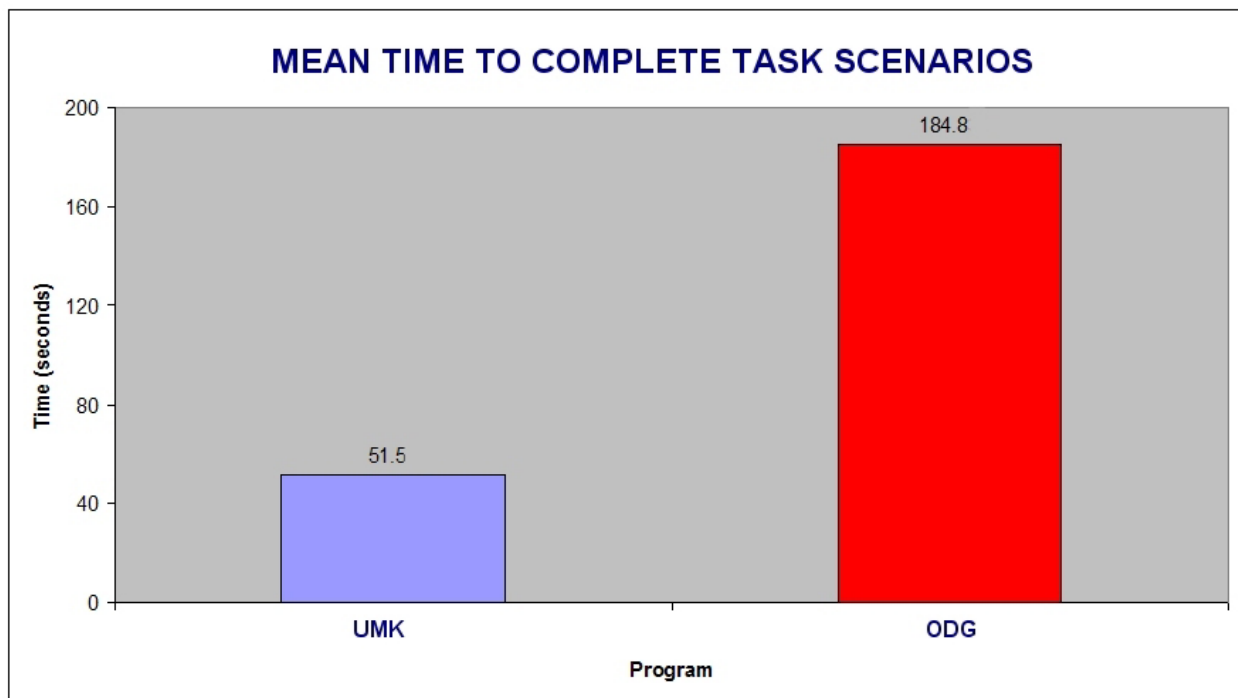


Figure 1. Mean time to complete tasks in seconds.

Number of Clicks to Complete Task

Another measurement of the ease and efficiency of a product with a task is the number of clicks that it takes a user to get to the information that he or she needs. In the comparison between the UMK and ODG-Treatment products, the difference in clicks to find the needed information was striking (see Figure 2). For the UMK product, the mean number of clicks to complete a task was 3.2 (range of 2-6). For the ODG-Treatment product, the mean number of clicks was 11.7 (range of 6-16). In several cases, the information needed to complete the task did not seem to be available in the ODG-Treatment product.

MEAN NUMBER OF CLICKS TO COMPLETE TASK SCENARIOS

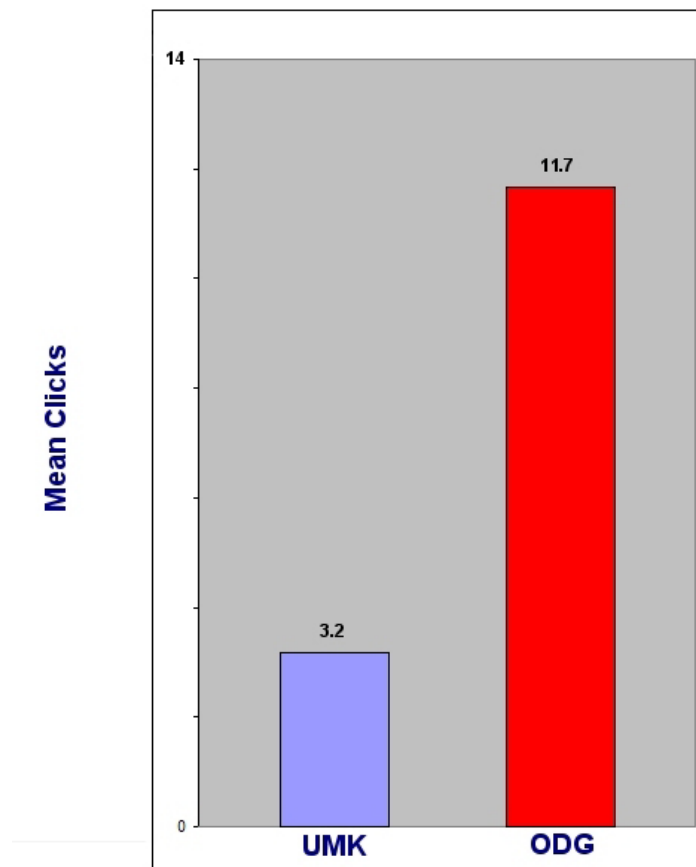


Figure 2. Mean number of clicks required to complete the task.

Usability Score

Using the scoring system described in the introduction, the reviewer scored each task according to the specified principles. In several cases, the ODG-Treatment product apparently did not include the necessary information (see Note on page 5). Assuming that these tasks are reasonable reflections of the type of information that a user would need to be successful in his or her job, the lack of this information in the ODG-Treatment product is considered a failing of the product. The usability is impacted because the user would be forced to find the necessary information from other sources, increasing the time to complete the task. In Figure 3, the UMK product had only minor usability issues with a mean usability score of 3.8 (range 3-4) while the ODG-Treatment system had a mean usability score of 0.7 (range 0-2).

Usability Score Based on Heuristic Evaluation

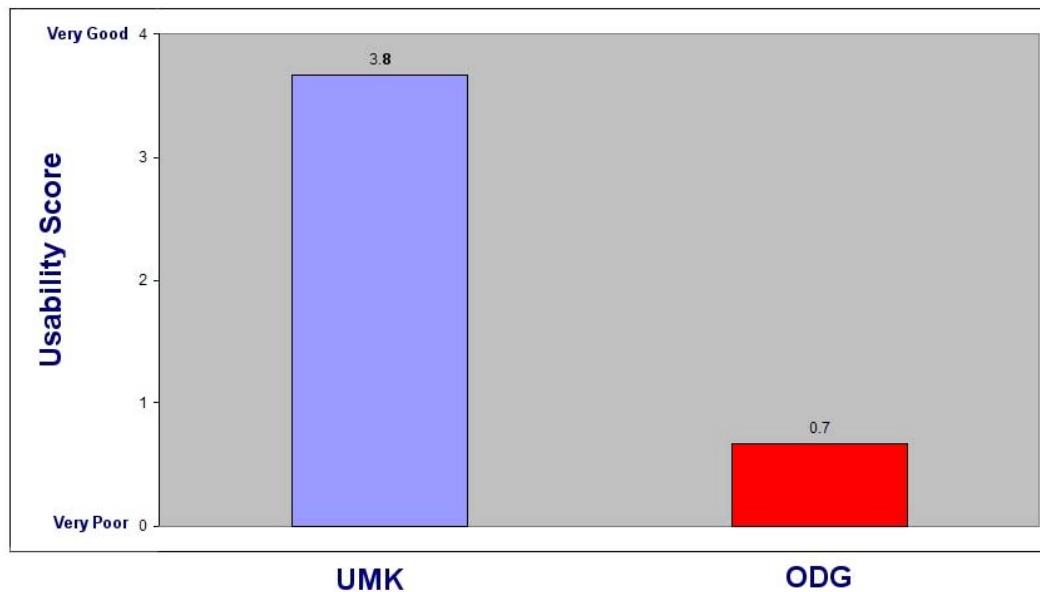


Figure 3. The mean usability score between the UMK and ODG-Treatment systems, where "4" indicates no usability problem and "0" indicates a catastrophic usability problem.

Appendix A

The following tables describe factors that influenced the overall usability score for each task. Where the systems did not achieve a perfect "4," we state the principle on the basis of which the overall score given was reduced. The final is influenced in addition by the length of time to complete the task, the number of clicks to complete the task and whether or not the task could be completed successfully. The scores are not derived from a formula, but from expert reviewer(s) assessment of the severity of the problems encountered. The score was based on the following scale:

- 4 = No usability problem**
- 3 = Cosmetic problem only**
- 2 = Minor usability problem**
- 1 = Major usability problem**
- 0 = Usability catastrophe**

Table 1. Table 1 shows the results of the task evaluations for the ACOEM's UMK product.

Tasks	Usability Score	Principles Used in Scoring	Comments
1. An employee is diagnosed with the following ICD-9 code: 724.4. What is the condition's name?	4		No problems encountered. Task was easy to complete in few steps.
2. The employee was diagnosed <u>14 weeks ago</u> . A. What diagnostic options for this employee should be considered? B. Do the diagnostic options for this employee differ if the condition were only diagnosed in the last two weeks? What are the differences and how can I tell?	4		No problems encountered. Task was easy to complete in few steps.

Tasks	Usability Score	Principles Used in Scoring	Comments
<p>3.</p> <p>A. When in the treatment spectrum is surgery for this condition an option to explore?</p> <p>B. What diagnostics are required before surgery should be considered?</p>	<p>4</p>		<p>No problems encountered. Task was easy to complete in few steps.</p>
<p>4. Does the tool provide a clear method of identifying and communicating the employee's diagnostic and treatment options that distinguishes between an acute and a chronic condition (such as an employee diagnosed 14 weeks ago and an employee diagnosed 2 weeks ago)?</p>	<p>4</p>		<p>No problems encountered. Task was easy to complete in few steps.</p>
<p>5. A printed copy of the employee's condition and care plan need to be printed for his file. Is it easy for me to print just what I need?</p>	<p>3</p>		<p>Task was fairly easy to complete. However, the print option for this topic would have been clearer if there had been a typical printer icon instead of a PDF icon.</p>

Table 2. Table 2 describes the WLDI's ODG-Treatment product (ODG-Treatment.com).

Tasks	Usability Score	Principles Used in Scoring	Comments
<p>1. An employee is diagnosed with the following ICD-9 code: 724.4. What is the condition's name?</p>	<p>2</p>	<p>Flexibility and efficiency of use</p> <p>Match between system and the real world</p>	<p>Task required clicking through multiple pages to narrow down to the page with the needed information. The condition listed did not have the same name as was used for the same code in the UMK product. This disparity may or may not be by design.</p>
<p>2. The employee was diagnosed <u>14 weeks ago</u>.</p> <p>A. What diagnostic options for this employee should be considered?</p> <p>B. Do the diagnostic options for this employee differ if the condition were only diagnosed in the last two weeks? What are the differences and how can I tell?</p>	<p>0</p>	<p>Flexibility and efficiency of use</p> <p>Match between system and the real world</p> <p>Error prevention</p> <p>Recognition rather than recall</p>	<p>No clear distinction could be found for the timeline differences. Diagnostic options were presented in context of procedures, requiring the user to review all procedures listed for the diagnostic options that should be considered for that particular procedure.</p>
<p>3</p> <p>A. When in the treatment spectrum is surgery for this condition an option to explore?</p> <p>B. What diagnostics are required before surgery should be considered?</p>	<p>1</p>	<p>Flexibility and efficiency of use</p> <p>Match between system and the real world</p> <p>Error prevention</p> <p>Recognition rather than recall</p>	<p>It was not clear when surgery was an option to explore. Under surgery there were multiple options. Each option had to be checked to understand diagnostics that might be considered.</p>

Tasks	Usability Score	Principles Used in Scoring	Comments
<p>4. Does the tool provide a clear method of identifying and communicating the employee's diagnostic and treatment options that distinguishes between an acute and a chronic condition (such as an employee diagnosed 14 weeks ago and an employee diagnosed 2 weeks ago)?</p>	<p>0</p>	<p>Flexibility and efficiency of use</p> <p>Match between system and the real world</p> <p>Error prevention</p> <p>Recognition rather than recall</p>	<p>There is no such option with the ODG-Treatment system.</p>
<p>5. A printed copy of the employee's condition and care plan need to be printed for his file. Is it easy for me to print just what I need?</p>	<p>1</p>	<p>Flexibility and efficiency of use</p> <p>Error prevention</p> <p>Match between system and the real world</p>	<p>Potential for serious error. There is no simple system for printing only the information about this condition. User could accidentally start a print job of hundreds of pages.</p>