USABILITY CHALLENGES OF UPGRADING A WORD PROCESSOR USER INTERFACE

L. MOETI, L. DE WET AND T. BEELDERS

ABSTRACT

This study evaluated the difficulty, in terms of usability, of upgrading from one version of a word processor user interface (Microsoft Word 2003) to another (2007). Laboratory-based usability testing involving 23 participants, who had never used Microsoft Word 2007 before, was conducted. All participants used the two versions in a repeated-measures experimental design. A Tobii 1750 Eye Tracker was used for screen recordings during testing. Participants were also required to fill in user satisfaction questionnaires. Results from the usability test showed that, in terms of usability, migrating from an original interface to a completely changed user interface is not easy.

Keywords: Human-Computer Interaction (HCI), Usability, Eye tracker, Word processor, Microsoft Word.

1. INTRODUCTION

The term 'word processor' was first used in the 1970's by IBM to refer to electric typewriter-like equipment that could record words on a magnetic tape (Wohl, 2005). This equipment was solely sold for word processing. It permanently stored word processing software in memory (Eissenberg, 1992). Nowadays, the term 'word processor' is used to refer to a computerized text editing system. Richard Broke, a programmer, wrote the first version of Microsoft (MS) Word for DOS-based IBM personal computers (Haigh, 2006). MS Word was the first word processor to use the graphical user interface (Wohl, 2005).

Currently there are many flavours of word processors. These include Word Perfect, Open Office and Google Docs, among others. Of all the office computing systems, word processors are the most commonly used (Rogers et al., 1994). Since their inception, word processors have been successfully integrated in the daily computing lives of many people. However, according to Mark, Lewis and Carroll (1990) word processors are among the most difficult systems to learn and master. Even though each word processor has a help facility, users do not usually know how to properly phrase questions for the help facility. Hence they get very little assistance from the help facilities. Users are, in general, very reluctant to read the explanations from the help facility. They rather prefer to find solutions by themselves.

In a study carried out by Beelders et al. (2008), it was concluded that users who are familiar with an older word processor user interface find it easier to use a slightly modified word processor user interface, provided the user
interface layout remains the same. According to Wenzel (2007), users who are familiar with the earlier versions of MS Word interfaces would find it difficult to get used to the MS Word 2007 user interface, the reason probably being that Microsoft has completely changed the user interface in MS Word 2007. With the release of this latest version of MS Word, a number of usability questions arise: What has been the impact of these 'wholesale' changes on the usability of MS Word? In particular, what has been the impact of the introduction of the ribbon menu style (in which all the related menu items are grouped together) on the usability of MS Word? How easy or difficult is it for users who are familiar with MS Word 2003 to get used to the new MS Word interface? What are the usability challenges associated with upgrading from MS Word 2003 to MS Word 2007? It is in light of these questions that this study was undertaken. Usability was addressed by investigating three attributes, namely effectiveness, efficiency and user satisfaction.

2. HYPOTHESES

In order to realize the objectives of the study, the following hypotheses were formulated and tested:

H$_{0_1}$. Completely changing the word processor user interface will not result in a significant difference in word processor effectiveness.

H$_{0_2}$. Completely changing the word processor user interface will not result in a significant difference in word processor efficiency.

H$_{0_3}$. Users will not prefer to work with the older user interface in comparison to the completely changed user interface.

3. METHODS

Participants

Twenty five test participants, consisting of 8 males and 17 females, were recruited. The data of two of these participants were not included in the analysis as they had already had some exposure to MS Word 2007, which was undesirable for the conditions of the study. Test participants ranged from those who had used MS Word 2003 for one year to those who had used it for more than five years. None of them had any prior knowledge of MS Word 2007. Twelve of the participants were employees of the University of the Free State, while 4 were employed elsewhere. The rest were students (both from the University of the Free State (UFS) and the nearby Central University of Technology (CUT)). All participants had used more than one of the products in the MS Office 2003 suite. The products in the MS Office 2003 suite have almost similar user interfaces. Each participant was paid an incentive of R50-00.
Experimental Design

This study adopted a within-participants experimental design, since the objective was to evaluate the ease with which the participants were able to use the new user interface (MS Word 2007). The experiment had two conditions: MS Word 2003 user interface and MS Word 2007 user interface. The participants experienced both conditions in a one way repeated measures experimental design. That is, each participant carried out identical tasks with both MS Word 2003 and MS Word 2007. In order to curb the effects of participants retaining task knowledge from MS Word 2003 to MS Word 2007, the order of tasks for MS Word 2007 was rearranged.

Procedure/Usability Testing

The usability test was conducted in the usability laboratory of the Department of Computer Science and Informatics at the UFS. The equipment consisted of a Tobii 1750 Eye Tracker and a web camera. The stimuli were MS Word 2003 and MS Word 2007 user interfaces. Upon acceptance of recruitment each participant had to sign a consent form. Before the usability test commenced each participant was briefed verbally about the usability test. The rights of the participant, as outlined in the consent form, were further reiterated. The participants were encouraged to express their thoughts and comments aloud as they progressed with the tasks. The eyes of each participant were calibrated for the eye tracker before the test commenced. A list of 25 tasks to perform on MS Word 2003, printed on paper, was then handed to each participant. After having completed the MS Word 2003 tasks, each participant was handed a list of 25 identical (but rearranged) tasks to perform on MS Word 2007. The tasks had to be performed in chronological order as presented. Participants were re-assured that they could skip any task that they found too difficult to do, perhaps after some trials. However, if they skipped a task they were not allowed to return to it at a later stage. Additionally, they were informed that they could leave at any stage during the usability test. Thus, although desirable, it was not a pre-requisite to complete all tasks. As each participant carried out the tasks, the researcher made observations on how the participant performed with the user interface. These observations were noted on a piece of paper bearing the participant’s reference number. Screen recordings of the activities that the participant performed on the screen were taken using the Tobii 1750 Eye Tracker. Participants were given a success level score, ranging from zero to one, for each task. A score was given for each completed step of a task that required multiple steps. All steps were awarded equal scores. A score was computed as follows:

\[
\text{Score} = \frac{\text{number of correct steps}}{\text{minimum number of steps required to complete a task}}
\]

For example, each completed step of a task that required four steps was awarded a score of 25% (0.25).
Upon completion of the usability test each participant was required to complete the user satisfaction questionnaire. The questionnaire consisted of 15 close-ended questions and 1 open-ended question. With the close-ended questions, the participants were required to indicate, on a 5-point Likert scale, the extent to which they agreed or disagreed with the statements that were given on the questionnaire. They were then asked to, optionally, give their opinions (which they felt were not covered by the 15 close-ended statements) about the MS Word 2007 user interface.

4. RESULTS

Effectiveness

Number of tasks completed

Each participant was awarded a score out of 25, based on the number of tasks completed. The scores included the success level each participant achieved on tasks that required multiple steps for their successful completion. On average, each participant completed more tasks with MS Word 2003 than they did with MS Word 2007. The average scores for MS Word 2003 and MS Word 2007 were 21.2 and 16.6 respectively.

In order to determine the significance of the difference (between the versions) in the number of tasks completed, a one-way repeated measures analysis of variance, at 0.05 significance level, was performed on the number of tasks (scores) completed by each participant on both versions. MS Word version was taken as a within-participants factor. Results of the analysis are shown in Table 1.

Table 1: Number of tasks completed ANOVA table

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>454.4022</td>
<td>22</td>
<td>20.65464</td>
<td>2.790255</td>
<td>0.005888</td>
<td>2.04777</td>
</tr>
<tr>
<td>Word versions</td>
<td>237.3967</td>
<td>1</td>
<td>237.3967</td>
<td>32.07015</td>
<td>0.000011</td>
<td>4.300949</td>
</tr>
<tr>
<td>Error</td>
<td>162.8533</td>
<td>22</td>
<td>7.402421</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>854.6522</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The analysis shows a significant difference in the number of tasks completed ($F = 32.07$, df = 1, $p < 0.05$) per MS Word version with more tasks being completed with MS Word 2003 than with MS Word 2007. Thus, the null hypothesis $H_0$ (completely changing the word processor user interface will not result in significant change in the word processor effectiveness), was rejected. Since the participants were already familiar with the MS Word 2003 user interface it was, therefore, easy for them to complete more tasks. As
opposed to the MS Word 2003 menu style, the MS Word 2007 ribbon presents
the user with too many menu options simultaneously. This could have
confused the first time users (participants). The complete rearrangement of
the menu in the MS Word 2007 interface resulted in participants struggling to
locate appropriate menu items. This led to participants often abandoning the
tasks. Hence, participants completed fewer tasks with MS Word 2007 than
they did with MS Word 2003.

Efficiency

Task Completion times

These were the times (in seconds) it took each participant to complete the
tasks with MS Word 2003 and MS Word 2007. On average, participants took
less time to complete tasks with MS Word 2003 as compared to MS Word
2007. The average times per task were 35.34 and 56 seconds for the two
versions respectively. In order to determine whether user efficiency is
compromised by upgrading from one MS Word version to the other, the time
that each participant needed to complete each task was noted. A one-way
repeated measures analysis of variance, at a 0.05 significance level, was
performed on the average times per version, with version being the within-
participants factor. The results are shown in Table 2 below:

Table 2: Time taken to complete a task ANOVA table

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tasks</td>
<td>36398.9</td>
<td>24</td>
<td>1516.621</td>
<td>3.230184</td>
<td>0.002624</td>
<td>1.983759668</td>
</tr>
<tr>
<td>Versions</td>
<td>5353.02</td>
<td>1</td>
<td>5353.02</td>
<td>11.40116</td>
<td>0.002497</td>
<td>4.259677214</td>
</tr>
<tr>
<td>Error</td>
<td>11268.37</td>
<td>24</td>
<td>469.5153</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>53020.29</td>
<td>49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There was a significant difference ($F = 11.4$, df = 1, $p < 0.05$) in task completion
times between the two MS Word versions, with participants taking longer per
task with MS Word 2007 than they did with MS Word 2003. Due to their
familiarity with the MS Word 2003 user interface, participants found it easier to
complete tasks with this version. When using MS Word 2007, participants
spent more time trying to understand the menu items. According to
Shneiderman & Plaisant (2005), a design that uses familiar terms enhances
the ability of first time users to get used to the new user interface. In upgrading
from MS Word 2003 to MS Word 2007, Microsoft did not use familiar terms for
labelling top level menu items. The MS Word 2003 top level menu items are:
File, Edit, Insert, Format, Table, Window and Help, while MS Word 2007 top
level menu items are: Home, Insert, Page layout, References, Mailings,
Review and View. Participants probably spent a lot of time trying to map the
MS Word 2007 menu items to the MS Word 2003 menu items.
The eye tracker was used to determine where the participants looked on the user interface in trying to locate appropriate menu items. However, the nature of the study could not allow for the use of the Tobii studio analysis features. As such the analysis was done manually using the captured Tobii data. Figure 1 below shows the eye gaze plot of a single participant trying to locate the 'File' menu in MS Office 2007 in order to open a file.

Figure 1: Eye gaze plot of a participant trying to locate the File menu in MS Office 2007

The participant was unaware that the Microsoft Office button ( ) in MS Word 2007 represents the 'File' menu item of the MS Word 2003. This resulted in an increase in the task completion time. Analysis of the gaze plots for the remainder of the participants indicated that they experienced the same problem with similar tasks. Hence, replacement of the 'File' menu with the Microsoft Office button had a negative effect on the efficiency of MS Word.

Average number of steps per task

As previously described, a step was considered to be a single correct selection of either a button, an icon or menu option that led to the successful completion of a task. On average, participants performed fewer steps (3.12) to complete a task with MS Word 2007 than with MS Word 2003 (3.4). Although the tasks were rearranged for MS Word 2007, identical tasks were given the same task number for analysis purposes. A one-way repeated measures analysis of variance, at a 0.05 level of significance, was performed on the average number of correct steps that the participants did per version, with version being the within-participants factor. The results are shown in Table 3 below:

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tasks</td>
<td>136.12</td>
<td>24</td>
<td>5.671667</td>
<td>7.349892</td>
<td>3.26E-06</td>
<td>1.98376</td>
</tr>
<tr>
<td>Versions</td>
<td>0.98</td>
<td>1</td>
<td>0.98</td>
<td>1.269978</td>
<td>0.270913</td>
<td>4.259677</td>
</tr>
<tr>
<td>Error</td>
<td>18.52</td>
<td>24</td>
<td>0.771567</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>155.62</td>
<td>49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There was no significant difference (F = 1.27, df = 1, p >0.05) in the number of
correct steps executed per task per version. Thus, with regard to the number of steps, a complete change of the word processor user interface may not impede on the efficiency of the word processor. Therefore, based on the number of steps taken to complete a task, $H_{a2}$ was not rejected.

**Average number of errors per task**

An error was taken to be any mouse click, menu, button or icon selection that was not needed for the successful completion of the task. The average number of errors committed per task per version was 9.32 and 22.88 for MS Word 2003 and MS Word 2007 respectively. With both versions, the highest number of errors was made on the task that required participants to insert an equation. Even though the tasks were rearranged for MS Word 2007, identical tasks were given the same task number for analysis purposes.

A one-way repeated measures analysis of variance, at 0.05 significance level, was performed with version as the within-participants factor. The results are shown in Table 4:

**Table 4: Number of errors per task ANOVA table**

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>$F$</th>
<th>$P$-value</th>
<th>$F$ crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tasks</td>
<td>5304.659</td>
<td>21</td>
<td>252.6028</td>
<td>0.963667</td>
<td>0.53339</td>
<td>2.084189</td>
</tr>
<tr>
<td>Versions</td>
<td>2611.841</td>
<td>1</td>
<td>2611.841</td>
<td>9.964043</td>
<td>0.004759</td>
<td>4.324794</td>
</tr>
<tr>
<td>Error</td>
<td>5504.659</td>
<td>21</td>
<td>262.1266</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13421.16</td>
<td>43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on the analysis in Table 4 above, there was a significant difference ($F = 9.96$, df =1, $p < 0.05$) in the number of errors committed by the participants, with most errors occurring during the use of MS Word 2007. Based on the number of errors committed per task, a complete change of the word processor user interface would impede on the efficiency of the word processor user. $H_{a2}$ was thus rejected. Due to their unfamiliarity with the MS Word 2007 interface, most participants opened every top level menu item in search for an appropriate sub menu item. At times they had to exhaust the whole top level menu list in order to get to the appropriate menu item.

**Number of voluntary help references per version**

A total of 13 voluntary help references were made during the use of MS Word 2003 as opposed to 8 with MS Word 2007. These counts were relatively small to conduct meaningful statistical tests. As such no statistical analyses were performed on the number of help references per task. However, a number of
observations were made during the playback of the recorded videos. Whenever they had problems executing the task, participants rarely used the help facility, probably because system users often do not know how to appropriately phrase questions to ask in the help facility (Mark et al., 1990). Sometimes, users simply prefer not to use the help facility even though they are aware of its availability (Grossman et al., 2009).

On the MS Word 2007 user interface the help menu is represented by a small 'question mark' icon. According to Grossman et al. (2009), it is normal for users not to notice all the icons and menu options in a bloated interface like MS Word 2007. Figure 2 below shows the eye gaze of one participant who failed to notice the 'help' icon at the far top right corner on the MS Word 2007 interface.

![Figure 2: Participant's eye gaze plot in trying to locate the help icon](image)

Analysis of the eye tracker data showed that a number of participants expected to find the help menu alongside other menu items, as is the case in MS Word 2003. The task performance was thus affected. This observation is in contradiction to the conclusion by McCarthy et al. (2003) that user expectation of where the menu is placed has no effect on task performance. Another reason could be that participants could not remember what the icon represented (Ellis et al., 1995). Although participants were not explicitly told not to use the shortcuts, none of them bothered to use the 'F1' shortcut key to access the help menu. This could be due to shortcuts being difficult to learn and to master (Lane et al., 2005).

With the exception of 'number of actions per task', all efficiency metrics showed that a complete change of the word processor user interface will result in a significant change in the efficiency of the word processor. Thus, the null hypothesis $H_0$ was rejected.

**User satisfaction survey**

The responses from the user satisfaction survey are summarised in Table 5 below. This was compiled by adding the number of participants who gave a particular response to a particular survey questionnaire statement.

**Key:** 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree and 5 = strongly agree
Table 5: Summary of survey questionnaire responses

<table>
<thead>
<tr>
<th>Questionnaire statement(s)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I would like to use MS Word 2007 frequently</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>2. I found MS Word 2007 unnecessarily complex</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>3. I would need a support of a technical person in order to be able to use MS Word 2007</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>4. I think MS Word 2007 is easy to use</td>
<td>10</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>5. I found various functions on MS Word 2007 well integrated</td>
<td>8</td>
<td>9</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>6. Most people would learn to use MS Word 2007 quickly</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>7. I found MS Word 2007 very cumbersome to use</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>8. I needed to learn a lot of things before I could get going with MS Word 2007</td>
<td>9</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>9. The menu style of MS Word 2007 was easy to navigate</td>
<td>8</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>10. The functions on the MS Word 2007 menus were well grouped together</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>11. I found the MS Word 2007 colour scheme very pleasing to use</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>12. I found it difficult to find menu options on the MS Word 2007 interface</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>13. I found the help function of MS Word 2007 helpful</td>
<td>7</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>14. I would generally prefer MS Word 2007 to MS Word 2003</td>
<td>13</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

Sixty five percent of the participants felt that MS Word 2007 was not easy to use. This could be because they struggled to locate appropriate menu items on the MS Word 2007 user interface. Most of the participants could not recognise that doing tasks with MS Word 2007 was completely different from the way similar tasks are done in MS Word 2003. Hence most participants continued to repeat steps that would not lead to the successful completion of the tasks. Although 52% of the participants found MS Word 2007 unnecessarily complicated, 74% of them found the colour scheme of MS Word 2007 aesthetically pleasing.

In order to test the null hypothesis \( H_{0,3} \) (users will generally not prefer to work with the older user interface in comparison to the completely changed user interface), a 'goodness of fit' Chi Square analysis, at 0.05 significance level and four degrees of freedom, was performed on statement 14 of the questionnaire (refer to Table 5). This was to determine if there was any significant difference between the expected responses to the ones observed per category. The categories were taken to be the questionnaire response options (strongly disagree, disagree, neutral, agree and strongly agree). The tabulated Chi Square value at 0.05 significance level and 4 degrees of freedom is 9.49. The questionnaire statement was taken to be the null hypothesis. Since the calculated Chi Square value (12.85) is greater than the tabulated value (9.49), the null hypothesis \( H_{0,3} \) was rejected. This means that
users would generally not prefer MS Word 2007 to MS Word 2003. This could be due to the 'fear of the unknown'. Since they already knew MS Word 2003, users preferred to stay with what they were familiar with.

5. CONCLUSIONS

The primary objective of this study was to assess, relative to usability, the ease with which users can migrate from an older version of the word processor to the completely changed version. Since usability researchers differ on which attributes should constitute usability (Folmer & Bosch, 2004), only three attributes (effectiveness, efficiency and user satisfaction) were considered for this study.

Effectiveness was measured by the number of tasks completed per version. The hypothesis \( H_{e1} \) was that there would be no significant difference in the number of tasks completed per version. The null hypothesis was rejected as it was found that there was actually a significant difference in the ability to complete tasks per MS Word version, with participants completing more tasks with MS Word 2003 than they did with MS Word 2007. Since the user interface has been completely changed and rearranged, users took more time trying to locate appropriate menu items. Some users resorted to perusing the whole menu on the user interface in order to locate what would seem to be an appropriate menu item. With tasks that required multiple steps, the amount of search time further increased. In a number of cases the users gave up the search for appropriate menu items. In terms of the number of tasks completed per version, the conclusion is that it is more difficult for users to get used to the word processor user interface that has been completely changed. Thus, completely changing the word processor user interface compromises its effectiveness. This is consistent with Colazzo et al. (2008) in which they concluded that people who are familiar with the old MS Office user interface would experience more problems in performing tasks with the new interface.

Efficiency was measured by the time taken to complete tasks, number of actions per task, number of errors per task, number of voluntary help references per task and level of success per task. The hypothesis \( H_{e2} \) was that efficiency would not be compromised by completely changing the word processor user interface. With regard to task completion rates the null hypothesis was rejected as there was a significant difference in the times taken to complete tasks per word processor version. Users took significantly longer to complete similar tasks with MS Word 2007 than they did with MS Word 2003. Users spent a lot of time trying to understand the new menu layout; more time than to actually perform the tasks. Often the users had to navigate through the full/complete menu system in order to get to the appropriate submenu. In terms of the task completion rates, it can be concluded that it is significantly more difficult to migrate from one version of the word processor to a completely changed version. Thus, completely changing the word processor user interface has a negative impact on its
The number of errors committed by the users while performing the task differed significantly between the two versions. More errors were committed with MS Word 2007 than with MS Word 2003. An explanation for this could be that the users were already familiar with completing tasks with MS Word 2003; hence fewer errors were committed during the use of MS Word 2003. However, the significant number of errors can be attributed to the 'trial and error' approach that many participants adopted when selecting menu items in MS Word 2007. Participants resorted to navigating the full-complete menu system in search for an appropriate submenu. The menu labelling on the ribbon did little to assist the users to locate appropriate menu items. This high error rate indicates that it was difficult to upgrade to a completely changed word processor interface. It can, therefore, be concluded that a complete change of the word processor user interface would result in an increase in the number of errors committed by the users. Thus, based on the number of errors committed per task per version, the hypothesis ($H_{o.2}$) was rejected. Regarding the number of voluntary help references, users made more help references while using MS Word 2003 than they did when using MS Word 2007. A lot of users struggled to locate the MS Word 2007 help menu. They looked for the help menu alongside other menu items, as is the case with other systems including MS Word 2003. In MS Word 2007 the help menu is represented by a small 'help' icon (atitis) on the far top-right corner of the screen. It is arguably the smallest of all the icons in the MS Word 2007 user interface. Another reason why users could not use the help menu may be that they did not know what to ask in the help menu (Mark et al., 1990). In terms of the average number of actions required to complete a single task, there was no significant difference between the two versions. This indicates that a complete change of the word processor user interface (including the menu style) may not necessarily hamper the performance of the system. Relative to the number of actions needed to execute a task, Microsoft's claim of improved usability in MS Word 2007 was confirmed to some extent. Hence, based on the average number of actions required to complete a single task, hypothesis ($H_{o.2}$) was not rejected.

**User satisfaction**

In a user satisfaction survey that was carried out at the end of the usability testing sessions a significant number of participants indicated that they would prefer to use MS Word 2003 as compared to MS Word 2007. However, it should also be noted that a significant number said that they would like to use MS Word 2007 more frequently (refer to questionnaire statement number 1 in Table 5). This could perhaps be caused by reasons that fall outside the usability scope. Results from the questionnaire further revealed that users would not prefer a completely changed word processor user interface, probably because of the amount of time and effort required to learn a completely changed interface. Therefore, hypothesis $H_{o.3}$ (users would generally not prefer the old user interface) was rejected.
6. RECOMMENDATIONS FOR FUTURE RESEARCH

The conditions under which the study was conducted were chosen so as to test the ease with which users could migrate from the well-know interface of MS Word 2003 to the completely overhauled interface of MS Word 2007. Therefore, it was imperative that the participants had not used MS Word 2007 before so that they had no chance to become accustomed to the new look of the interface. It was found that migration to a completely new interface poses some difficulties for users. However, this does not necessarily mean that the new interface is not usable. Long-term exposure to the interface could yield very different results than those found in this paper. Therefore, possible further research could be to allow some medium to prolonged exposure to the system and then re-test these participants on the MS Word 2007 interface. The task times, errors and other measures can be captured and compared to those achieved using MS Word 2003. In this way it will be possible to analyse the long term effect of the interface, particularly in terms of its learnability and other overall usability.

The following are the suggestions for future research in the area of word processor usability:

- Microsoft has introduced a ribbon menu style. Until now this style is only being used by Microsoft. Since the ribbon menu style is a new menu style altogether, it could be interesting to investigate its usability relative to the already existing menu styles.

- Investigating the impact of the gradual introduction of changes, as opposed to a complete changeover of the word processor interface on the usability thereof.

- Investigation of the effect of gender on the usability of the word processor.

7. REFERENCES


