INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600
RICE UNIVERSITY

IMPROVING SYSTEM KNOWLEDGE AND PERFORMANCE WITH PROACTIVE INTELLIGENT ON-LINE HELP

by

LARRY L. CORNETT

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE DOCTOR OF PHILOSOPHY

APPROVED, THESIS COMMITTEE:

David M. Lane, Associate Professor, Director Psychology

Kenneth R. Laughery, Professor Psychology

Miguel A. Quinones, Assistant Professor Psychology

H. Albert Napier, Associate Professor Administrative Science

Houston, Texas
April, 1997
ABSTRACT

Improving System Knowledge and Performance with Proactive Intelligent On-line Help

by

Larry L. Cornett

Computer software is becoming increasingly powerful and complex, but most people only utilize a small fraction of the capabilities of the software that they use to accomplish their daily tasks (Antsey, 1988; Fischer, 1987, 1993; Sutcliffe & Old, 1987). Carroll and Rosson (1987) described this problem as the production paradox. Typical methods of training and help, such as tutorials, manuals, and on-line help, fail to help intermediate-level users overcome this problem (Elkerton & Williges, 1989). The use of proactive, intelligent help systems (e.g., Coaches and Critics) may provide a more effective approach (Selker, 1994). Coach, Critic, and tutorial training were compared to determine which intelligent help design would provide the most effective instruction. Sixty-two users were trained, tested, and evaluated on advanced word processing functions during three sessions to assess the acquisition, retention, and application of efficient methods. The Coach and tutorial groups both improved their performance times and efficiency by the second test. However, the Critic group did not perform as well. These results suggest that providing users with frequent, proactive, context-specific feedback and help may encourage them to learn and utilize more efficient software methods.
ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my wife, Nicole. This would not have been possible without her years of encouragement, understanding, patience, and assistance. I hope that I can someday repay her for all of her support and long hours of help. I would also like to thank my families, both the Cornetts and the Mocks, for their interest in my research and for encouraging me to continue my efforts. Finally, I would like to express my appreciation to the members of my committee: Dr. David Lane and Dr. Al Napier for supporting my research and mentoring me throughout my years at Rice and Dr. Mickey Quiñones and Dr. Kenneth Laughery for all of their helpful comments and feedback.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Types of On-line Help</td>
<td>7</td>
</tr>
<tr>
<td>Issues</td>
<td>14</td>
</tr>
<tr>
<td>METHOD</td>
<td>18</td>
</tr>
<tr>
<td>Participants</td>
<td>18</td>
</tr>
<tr>
<td>Design</td>
<td>18</td>
</tr>
<tr>
<td>Materials</td>
<td>19</td>
</tr>
<tr>
<td>Procedure</td>
<td>22</td>
</tr>
<tr>
<td>RESULTS</td>
<td>25</td>
</tr>
<tr>
<td>Background Questionnaire</td>
<td>25</td>
</tr>
<tr>
<td>Learning Measures</td>
<td>27</td>
</tr>
<tr>
<td>Analysis of Help Use</td>
<td>36</td>
</tr>
<tr>
<td>Behavior</td>
<td>38</td>
</tr>
<tr>
<td>Reactions to Training</td>
<td>44</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>48</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>60</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>68</td>
</tr>
<tr>
<td>APPENDIX B</td>
<td>78</td>
</tr>
<tr>
<td>APPENDIX C</td>
<td>80</td>
</tr>
<tr>
<td>APPENDIX D</td>
<td>86</td>
</tr>
<tr>
<td>APPENDIX E</td>
<td>90</td>
</tr>
<tr>
<td>APPENDIX F</td>
<td>93</td>
</tr>
<tr>
<td>APPENDIX G</td>
<td>95</td>
</tr>
<tr>
<td>APPENDIX H</td>
<td>98</td>
</tr>
</tbody>
</table>


LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE 1</td>
<td>Differences between Coach and Critic help systems</td>
<td>16</td>
</tr>
<tr>
<td>TABLE 2</td>
<td>Demographic information</td>
<td>26</td>
</tr>
<tr>
<td>TABLE 3</td>
<td>Performance times for Tasks 1, 2, and 3 (sec)</td>
<td>28</td>
</tr>
<tr>
<td>TABLE 4</td>
<td>Efficiency scores for Tasks 1, 2, and 3 (100% maximum)</td>
<td>30</td>
</tr>
<tr>
<td>TABLE 5</td>
<td>Performance times for Tests 1 and 2 (sec)</td>
<td>32</td>
</tr>
<tr>
<td>TABLE 6</td>
<td>Efficiency scores for Tests 1 and 2 (100% maximum)</td>
<td>34</td>
</tr>
<tr>
<td>TABLE 7</td>
<td>Percentage of participants using help for Tasks 1, 2, and 3</td>
<td>37</td>
</tr>
<tr>
<td>TABLE 8</td>
<td>Percentage of participants using help for Tests 1 and 2</td>
<td>38</td>
</tr>
<tr>
<td>TABLE 9</td>
<td>Mean ratings behavior questions for first session (10 point scale)</td>
<td>39</td>
</tr>
<tr>
<td>TABLE 10</td>
<td>Mean ratings behavior questions for second session (10 point scale)</td>
<td>39</td>
</tr>
<tr>
<td>TABLE 11</td>
<td>Mean ratings behavior questions for third session (10 point scale)</td>
<td>40</td>
</tr>
<tr>
<td>TABLE 12</td>
<td>Mean ratings reaction questions for first session (10 point scale)</td>
<td>44</td>
</tr>
<tr>
<td>TABLE 13</td>
<td>Mean ratings reaction questions for second session (10 point scale)</td>
<td>45</td>
</tr>
<tr>
<td>TABLE 14</td>
<td>Overall ranking for performance measures</td>
<td>49</td>
</tr>
<tr>
<td>TABLE 15</td>
<td>Procedures for each group across all sessions</td>
<td>120</td>
</tr>
<tr>
<td>TABLE 16</td>
<td>Intercorrelations among major study variables</td>
<td>151</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIGURE 1</td>
<td>Basic model of an intelligent coach or critic help system</td>
<td>12</td>
</tr>
<tr>
<td>FIGURE 2</td>
<td>Intelligent help window</td>
<td>20</td>
</tr>
<tr>
<td>FIGURE 3</td>
<td>Training group by task interaction (performance time tasks 1-3)</td>
<td>28</td>
</tr>
<tr>
<td>FIGURE 4</td>
<td>Training group by task interaction (efficiency score tasks 1-3)</td>
<td>30</td>
</tr>
<tr>
<td>FIGURE 5</td>
<td>Training group by session interaction (performance time tests 1 and 2)</td>
<td>32</td>
</tr>
<tr>
<td>FIGURE 6</td>
<td>Training group by session interaction (efficiency score tests 1 and 2)</td>
<td>34</td>
</tr>
<tr>
<td>FIGURE 7</td>
<td>Percentage of participants performing knowledge test tasks correctly</td>
<td>36</td>
</tr>
<tr>
<td>FIGURE 8</td>
<td>Boxplot of group performance times for Task 1</td>
<td>122</td>
</tr>
<tr>
<td>FIGURE 9</td>
<td>Boxplot of group performance times for Task 2</td>
<td>123</td>
</tr>
<tr>
<td>FIGURE 10</td>
<td>Boxplot of group performance times for Task 3</td>
<td>124</td>
</tr>
<tr>
<td>FIGURE 11</td>
<td>Boxplot of group efficiency scores for Task 1</td>
<td>125</td>
</tr>
<tr>
<td>FIGURE 12</td>
<td>Boxplot of group efficiency scores for Task 2</td>
<td>126</td>
</tr>
<tr>
<td>FIGURE 13</td>
<td>Boxplot of group efficiency scores for Task 3</td>
<td>127</td>
</tr>
<tr>
<td>FIGURE 14</td>
<td>Boxplot of group performance times for Test 1</td>
<td>128</td>
</tr>
<tr>
<td>FIGURE 15</td>
<td>Boxplot of group performance times for Test 2</td>
<td>129</td>
</tr>
<tr>
<td>FIGURE 16</td>
<td>Boxplot of efficiency scores for Test 1</td>
<td>130</td>
</tr>
<tr>
<td>FIGURE 17</td>
<td>Boxplot of efficiency scores for Test 2</td>
<td>131</td>
</tr>
</tbody>
</table>
INTRODUCTION

In a perfect world computer software would be so easy to use that people could quickly learn new software and immediately begin utilizing it to accomplish their daily tasks. They could focus on each task itself rather than on the software tool. Most computer users wish that interaction with computer systems could be completely intuitive (Kearsley, 1988). However, anyone who has ever used software knows that such a perfect scenario, while incredibly desirable, does not currently exist. In fact, creating software that could immediately be efficiently used without any training is probably an unattainable goal. Computer software is becoming more powerful and complex with each new version and application release. Indeed, many users are overwhelmed by the plethora of features that are being added to their favorite applications, sometimes to the extent that they will revert to using an older version of the software (van Dam, 1997). Even infrequent users are now often faced with complicated and demanding tasks (Elkerton, 1988; Elkerton & Williges, 1985). Software functions and designs are changing so rapidly that current training methods cannot be expected to prepare users for long-term usage (Landauer, 1987; Newman & Riner, 1997). People will need to keep their skills up-to-date for the software programs that they utilize on a daily basis if they hope to use them most effectively. Most software designers attempt to aid users in this endeavor by providing better manuals, more usable on-line help systems, and more comprehensive tutorials. These aids are designed with the assumption that users who want to learn how to use the software will patiently read through the manuals, work through the tutorials, practice the skills, and be able to easily find answers in on-line help when necessary (Carroll & Rosson, 1987). Unfortunately, this assump-
tion is rarely true.

Most people dislike the feeling of "incompetence" that arises when learning how to use a new computer system (Carroll, 1987). They do not want to "suffer" through tutorials that are often completely unrelated to the actual tasks they encounter in their jobs (Newman & Riner, 1997). They prefer not to use the often overwhelming and confusing manuals and on-line help. Instead, users typically seek out a local expert who can help them (Ehrlich & Rohn, 1994; O’Malley, 1986). However, there are not enough expert users with the time available to help every user every time that a problem arises and this solution is expensive since it results in two people being "off-task" (Pilkington, 1992). Typically, users also prefer to be active rather than passive learners and attempt to learn software by performing their usual tasks and exploring the interface. The current GUI interfaces support this type of "on-the-job" learning, to an extent, by placing functions and commands in menus and on toolbars. Users can explore the interface and easily try out the available functions. Indeed, it has been discovered that users who adopt such an exploratory learning strategy may learn faster and perform better than those who learn by more traditional methods (Kamouri, Kamouri, & Smith, 1986; Robert, 1987; Wendel & Frese, 1987). However, people are often not completely successful in their learning endeavors. Ironically, the ease of use of many current software applications enables users to successfully accomplish their daily tasks without ever having read the user manuals or on-line help. But, if users do not take the time to explore this information, they often do not learn about all of the features that are available in a system. It is doubtful that all of these features will be discovered in the context of the user’s daily tasks either.

Discovering a system’s features and learning how to implement them are
both critical components of skill learning (Carroll, Mack, Lewis, Grischowsky, & Robertson, 1985). However, Fischer (1991) noted that as software continues to become more powerful and complex, mastering all of the available features becomes a task that exceeds the cognitive capabilities of most users. High-functionality systems have so many features that there is a “tool-mastery burden” (Fischer, 1993, p. 18). Users often are not aware of the existence of many functions. They do not know how to access them, how to use them, when they should be utilized, or how to plan a sequence of functions to achieve their goal (Breuker, 1992). They may suspect that a specific useful function exists, but they do not want to interrupt their task and search for help in large, confusing help manuals and systems. They do not feel that they should enter “poorly understood, error prone territory” in order to learn new functions just to complete a simple task (Ashworth, 1992, p. 446). In addition, many people feel pressure in their jobs to complete tasks quickly and successfully. They often resort to utilizing a few functions that they know well to accomplish their tasks, rather than searching for alternate, and possibly more effective, methods of unknown reliability. The cost of failing to complete a task on time is great enough to discourage “learning and acquisition of new and possibly better methods during the time of work” (Ericsson, Krampe, & Tesch-Romer, 1993, p. 368). This common situation results in users typically learning just enough about the application or system to accomplish their typical daily tasks. Carroll and Rosson (1987) referred to this situation as the production paradox. The production paradox is a term used to describe the finding that even experienced users rarely take the time to learn new methods or functions that would increase their efficiency. Instead, they tend to rely on a few well-learned and practiced procedures to accomplish most of their tasks, even if those procedures are not the most
efficient. Users do not seem willing to invest their limited time and energy in learning more about the systems they use. In fact, the majority of users only utilize a fraction of a system's features (Antsey, 1988; Bias & Mayhew, 1994; Sutcliffe & Old, 1987). A study of command usage in the Lotus 1-2-3 spreadsheet program by Napier, Batsell, Lane, and Guadagno (1992) found that even experienced users rarely used commands that are taught in intermediate and advanced classes. They suggested that many users may simply not be aware of the software's advanced features. In fact, only 3.6% of the available commands were used over 80% of the time and more than 50% of the 505 commands that were available were never used at all. This study is just one example that illustrates the fact that most users are not taking full advantage of the capabilities of the software that they use frequently. They are relying on a small subset of less-advanced commands, which may result in slow and inefficient usage of a very powerful software package.

Why is it important to be concerned that most people do not utilize the most efficient methods for the software applications that they use on a daily basis? There are several reasons why it is desirable to make software and help more usable and useful, thus enabling users to perform their daily tasks more efficiently. Kearsley (1988) noted that "improvements in the usability are translated into financial benefits in the form of reduced product support and marketing costs" (p. 71). Developers want people to buy their software. When people find software difficult to use and cannot easily find the answers they need in the manuals, tutorials, or on-line help, they become frustrated. If the option is available to them, they will quickly abandon such software for another product that they think will be more usable and help them accomplish their tasks more quickly and efficiently. Employers are also interested in efficiency since they invest a great deal in their software. The U.S.
Department of Defense alone spends approximately $30 billion annually on software (Rada & Moore, 1997). Considering this investment, it is not surprising that they want to get the most out of their software. Therefore, millions of dollars (e.g., for training materials, courses, travel expenses, and “lost” work time) and countless hours every year are invested in training people to use a variety of software packages. For example, Motorola spent approximately $120 million on training in 1992 (Henkoff, 1993). Providing advanced technical training can cost up to $50,000 per employee for just the initial year of employment alone (Caldwell, 1994). However, a great deal of what they are taught is never used in the workplace and the full capabilities of the software are not exploited. If people don’t have the immediate opportunity to use what they have learned when they return from training, they are likely to have forgotten the new methods by the time the opportunity presents itself. Also, most training sessions cover a fairly large amount of material. Many people simply can’t remember how and when to utilize the multitude of procedures that they were taught during a training session. For example, Napier, Lane, Batsell, and Guadagno (1989) found that only 36% of their subjects could successfully complete a spreadsheet problem that they had been taught less than 24 hours before. Consequently, many users often revert to using the inefficient methods with which they are familiar. Wasted training and slower, inefficient methods result in a loss of time and money. Finally, most people want to use their software more efficiently and will tolerate additional training if they believe that it will help them achieve this goal (Chignell, Hancock, & Loewenthal, 1989). Efficient software methods not only save time, but they usually also have the additional benefit of translating into fewer keystrokes and mouse movements, which may help reduce the incidence of repetitive stress injuries (Ehrlich & Rohn, 1994).
A great deal of training research in human-computer interaction (HCI) has focused either on the novice or the expert user. These studies have typically examined how to teach novices to use new software applications through a wide variety of training methods. However, this research has tended to ignore the more experienced, intermediate-level users who probably account for the largest percentage of total computer users (Lee & Barnard, 1993). Fischer (1993) believed that we need to focus on these domain workers, not the novices or computer experts:

Domain workers are not novice or naive users. They are people who have computational needs and want to make serious use of computers but are not interested in becoming professional programmers. They are skilled and knowledgeable in their respective domains; they use computers by choice, and over extended periods of time. (p. 21)

This study focused on this subset of users. These users can be described as people who utilize software to successfully accomplish their daily tasks and are fairly proficient with the methods that they know. However, they are usually not aware of more advanced methods that would help them accomplish their tasks more quickly and efficiently (Breuker, 1992; Fischer, Lemke, Mastaglio, & Morch, 1990; Paxton & Turner, 1984). They have special needs that require a different type of learning assistance than what would be effective for novices or experts (Santhanam & Wiedenbeck, 1993). They find traditional methods of computer assistance to be tedious or insufficient, since they are often directed at helping either novice users who know very little about the software or expert users who are quite familiar with it. For example, consider a scenario in which a user wants to select an entire line of text in a document that has been created in Microsoft Word™. The user carefully places the insertion point before the first word in the line and drags the
mouse across the entire line, character by character, to the end of the last word. This method is tedious and inefficient, but she continues to use it several times a day because she isn’t even aware of the much more efficient method of placing the cursor in the left margin and clicking once to select the entire line. She will never use this more efficient method because she doesn’t even know or may have forgotten that it exists. Intermediate users like this have different training and help needs than novices and experts. Since they are already familiar with the software they do not need introductory, basic help. Rather, they simply need to be made aware of the full capabilities of the software at points when a more efficient method could save them time and effort. Often, the savings are not immediately apparent to the user, but are greatly appreciated over time. Owen (1986) constructed a help facility for the UNIX operating system called “Did you know?” that attempted to address this issue. It provided users with information about features of the system so that they could better utilize the full range of its capabilities. In response to typing “DYK”, users were given useful tidbits of information and suggestions (e.g., Did you know that there is an easier way to clear the clipboard?). Even though this type of help system does not fully address the problems of intermediate users, it is a good example of an early attempt to provide help to this type of user. The question remains: How can users like this be provided with timely and effective training in more efficient methods?

Types of On-line Help

There are many types of help available to computer users ranging from hard-copy user manuals to active intelligent agents. This help varies on several dimensions: Active vs. passive, context-sensitive vs. static, integrated vs. separate, and immediate vs. delayed. Different types of help are more effective for different
types of users and there are individual differences in the type of information that users will prefer to seek out when they need help. For example, command reference manuals are generally better suited for expert users who simply need to refresh their memory or look up a specific command. They are typically too complex and difficult to search for novice and intermediate-level users. The focus of this study is on on-line computer help, which can be divided into the following general categories:

**Tutorial systems.** Tutorials have become a common addition to many software packages. Some tutorials are still found in hard-copy form, but many have moved on-line. They typically provide information and sequences of exercises that "walk" new users through the basics of using the application. Even operating systems like Apple Computer's System 7 ship with a tutorial that helps novices learn the basics of using the mouse and navigating the Macintosh™ desktop environment. While such static tutorials are fine for helping novices "learn the ropes", they are inadequate for users who are already familiar with the software. However, the development of intelligent tutoring systems (ITS) holds some promise for helping to train a wider range of users. Although ITS follow a set curriculum based on what the designer thinks the learners should know, they are capable of responding differently depending on the learner's answers. In addition, they build simple models of the learner's knowledge, which is used to branch to additional appropriate information and lessons. Early results with ITS have shown success in some domains. A study involving a LISP programming tutor at Carnegie-Mellon found that teaching a LISP course with an ITS was faster and students who used the ITS in addition to lectures produced grades that were 43% better than students who only attended the lectures (Anderson & Reiser, 1985). Some examples of other successful ITS are the
Recovery Boiler Tutor (Woolf, 1992), SOPHIE (Brown, Burton, & DeKleer, 1982), and Sherlock (Feifer, 1992).

However, ITS are not well-suited to providing the type of training that intermediate users need. They are not typically geared toward providing training in procedures, but a larger problem is that they are not embedded within the application that the learners are using. As such, they cannot really provide context-sensitive help appropriate for the user's current task. Since the ITS resides outside of the application, the user is required to leave their task environment and engage in non-productive learning exercises. Therefore, even though ITS add artificial intelligence and flexibility to the tutorial help format, they cannot solve the production paradox that intermediate users face.

On-line help systems. On-line help has traditionally taken the information contained within an application's user and reference manuals and put it on-line so that it can be readily accessed while users are working with the application. This has often resulted in cumbersome help facilities that users have usually avoided. However, designers have recently attempted to make their on-line help systems more usable. They have provided hot links that provide users with explanations and additional information when a specially-marked word is clicked. They have created better indices of task-related help topics and search facilities. On-line help can be useful when users are unable to proceed with their current task by providing an immediately available source of information (Duffy, Palmer, & Mehlenbacher, 1992). It is most helpful for more advanced users who can formulate their problem and are familiar with the search terms. It also allows users to learn in a productive situation while working on their current tasks, rather than completing the artificial tasks of a tutorial.
However, on-line help is not the answer to all problems that users face. Often, the vocabulary used in help systems is restrictive and confusing (e.g., different applications use different terms for the same function) making it difficult for users to locate the information that they need (Gwei & Foxley, 1990). Users can quickly become frustrated because they know that the information is available, but they cannot find it. Indeed, many users will avoid using on-line help or manuals and seek help instead from a friend or local expert, which results in the undesirable situation of two people being “off task” (O’Malley, 1986). One of the problems with typical on-line help is that it is passive, forcing the user to suspend work while looking for help, and typically is not sensitive to the context of the user's current task. It also does not help intermediate users overcome the production paradox, since they have not reached an impasse that forces them to use the help. They may be accomplishing their tasks through inefficient methods and may not even realize that there are better methods that they could learn about through help (Breuker, 1988; Duffy et al., 1992; Fischer, 1993; Paxton & Turner, 1984). However, recent advances in on-line help systems (e.g., Wizards) suggest that users may be able to benefit from this more proactive help.

**Intelligent assistants.** Intelligent assistants are autonomous and adaptive computer programs that can operate in a variety of software environments (Roesler & Hawkins, 1994). They exhibit artificial intelligence and can perform a wide range of tasks for users to save time and frustration, based on the user's preferences. Although they are typically thought to be useful for such tasks as filtering electronic mail, scheduling meetings, and searching information databases, they could also perform tasks within an application for users when they perform inefficiently. Therefore, users would not be forced to learn how to use more efficient
methods. An assistant would simply perform an operation for a user if it detected that the user needed help. Obviously, assistants can be very useful tools.

AppleSearch, which was developed by Apple Computer, is an example of an intelligent agent that easily performs tasks which would otherwise be tedious, if not impossible, for a human user (Valauskas, 1994). An intelligent agent, called a reporter, searches and retrieves text from computers linked on an Appleshare network. It creates a ranked list of documents that fit the user’s search criteria and will retrieve a full copy of any requested document. Multiple reporters can be scheduled to automatically perform a variety of searches at specific times.

On the surface, this type of help sounds wonderful, but many people are uncomfortable with the thought of giving up such control (Norman, 1994). There are still unresolved issues of safety, trust, and expectations. Also, over time, assistants create a dependency. If users always rely on an assistant to perform certain tasks for them, they will never learn how to perform the methods themselves. The users cannot function without the assistant and they are at a loss if it is damaged in a disk crash or if they change jobs and lose the assistant. They are then suddenly performing again at the novice skill level without the benefit of the assistant. Obviously, intelligent assistants are not meant to be an effective training method.

**Coach systems.** A Coach system is similar to an intelligent assistant, but rather than helping users by automatically performing tasks for them, it teaches users how to do the tasks themselves. Coaches are advisory-style systems with the goal of educating the user, rather than just assisting. Unlike an assistant, they are intended to build the user’s skill level (Selker, 1994). Coaches are embedded intelligent help systems that proactively provide users with unsolicited help and training when they need it (Selker, 1989, 1994). A Coach records users’ performance with a
system in order to develop a user model and create personalized help to present to them (see Figure 1). It can provide them with tips and information to improve efficiency and productivity when it notices that they are making errors or performing inefficiently. It utilizes the user model to determine what the user needs help with and the coaching knowledge (i.e., a pedagogical model) to determine when and how to provide advice. A Coach provides immediate context-specific feedback during task performance each time that it determines that the user needs help, thus allowing immediate application of the newly-learned method to the task. Therefore, users are not forced to leave their tasks to search for help. Instead, the appropriate help is provided within the context of their current task. However, this type of frequent, interactive help is computationally intensive. Also, a Coach may misinterpret an individual user action as an error because it may be focusing on a small part

FIGURE 1. Basic model of an intelligent coach or critic help system
of the larger task without being aware of the user's overall goal.

Coaches have been found to be effective teaching tools. Selker (1989) compared a Coach system with a similar, but non-adaptive, system for teaching novice programmers the LISP language. He found that the Coach group wrote five times as many functions as the other group, liked LISP more, rated the Coach higher as a learning environment, and utilized all types of available help. The users who were helped with the adaptive Coach system showed greater satisfaction and productivity. This type of active assistance encourages users to learn new methods within the context of their current computing task, thus enabling learning by doing (Elkerton & Williges, 1989). In another study, Zellermayer, Salomon, Globerson, and Givon (1991) used a Coach called the Writing Partner to cue students with questions concerning planning and organizing while they were writing essays. Interestingly, they found that the students who used the Writing Partner took longer to write their papers and did not create better essays during the training period than students who worked without the Coach. However, two weeks later the Writing Partner students were writing better essays than the other students. Thus, the Coach in this study caused delays during training, but the long-term effects were beneficial. These results suggest that an intelligent Coach system may be one way to provide users with timely and effective training in more efficient methods within the context of their work tasks.

Critic systems. Critic systems are also advisory-style systems that attempt to educate the user. They are similar to Coaches in that they save users the trouble of searching for help by informing them of more efficient methods. They allow learning on demand by providing users with new knowledge within the context of their tasks (Fischer, 1993). The main difference between Critics and Coaches is that Crit-
ics provide "batch" feedback to the user at one specific point, rather than continual feedback during task performance. Typically, passive Critics provide this feedback when requested by the user or display it at the end of a task session. However, a Critic can be proactive and provide feedback when it determines that the user needs it. They are not as intrusive as Coaches, but the feedback that they provide is less context-specific and the user is obviously not able to take advantage of the help information at the time the original error occurred. Thus, Critics provide task-specific help at a more global level.

Critics have been successfully implemented in many different types of applications. One example is an architectural design system called JANUS that used the critic approach to help designers develop floor plan layouts and learn general construction principles. JANUS was found to be a useful tool for both design students and more experienced designers. Another example of a critic that was implemented within an entirely different domain is an active help system called ACTIVIST, which was designed for a screen-oriented text editor. Basically, ACTIVIST continually observed a user's keystroke sequences and suggested shorter, more efficient alternatives if they existed. Fischer (1991) has noted that the strength of such Critic systems is that they support users while they are working and integrate learning with their work environment. This also suggests that a more proactive Critic system may be another way to provide intermediate users with effective training in efficient methods within the context of their work tasks.

*Issues*

Intelligent help systems that proactively provide advice and training enable a more active learning environment than previous systems that have been developed to help train users, such as "training wheels" interfaces (Carroll & Carrithers,
They are similar to "command-selection aids" in that they help intermediate users learn new methods to improve performance (Elkerton & Williges, 1987). This type of help system also may effectively address the production paradox of such users because it volunteers the help that they need. It solves the "bottleneck problem that occurs when the user never asks for needed information because the user does not realize that the information is necessary" (Chin, 1991, p. 197). However, most research in this domain has typically focused on the technical aspects of intelligent help systems like Coaches and Critics. There has been very little behavioral research on how these systems should be designed (Elkerton & Williges, 1989). Issues involving feedback, practice, and intrusiveness need to be examined (see Table 1).

Coaches provide immediate, frequent feedback whereas the feedback from Critics is delayed. The research on the effect of frequency of feedback on learning has been mixed. Immediate feedback within the context of a problem has often been viewed as a necessary component of skill acquisition (Anderson, 1987). Indeed, without adequate feedback, effective learning and improvement of skills may be impossible (Ericsson et al., 1993). It has generally been believed that enhancing feedback by making it more immediate, detailed, and frequent will lead to better learning, which would suggest that a Coach system would be more effective than a Critic (Schmidt & Bjork, 1992). However, other researchers have noted that even though immediate and frequent feedback may result in better immediate task performance, it is not necessary and, in fact, may even interfere with long-term retention of skills (Gaynor, 1981; Schmidt & Bjork, 1992). Schmidt and Bjork (1992) proposed that frequent feedback may disrupt long-term learning because learners come to see it as part of the task and perform poorly when feedback is no
longer available. In addition, frequent feedback may disrupt information-processing activities during practice and hinder learning. This would suggest that a Critic system may be more effective than a Coach for promoting long-term learning, since it uses less-intrusive, infrequent feedback.

<table>
<thead>
<tr>
<th>Issues</th>
<th>Coach</th>
<th>Critic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feedback</td>
<td>• Immediate</td>
<td>• Delayed</td>
</tr>
<tr>
<td></td>
<td>• Continuous</td>
<td>• Infrequent</td>
</tr>
<tr>
<td>Context-sensitivity</td>
<td>• Within the context of an individual operation</td>
<td>• Within the context of the whole task</td>
</tr>
<tr>
<td>Intrusiveness</td>
<td>• Frequently interrupts user</td>
<td>• Interrupts user only once</td>
</tr>
<tr>
<td>Practice</td>
<td>• Distributed</td>
<td>• Massed</td>
</tr>
<tr>
<td></td>
<td>• Practice individual procedures</td>
<td>• Practice multiple procedures</td>
</tr>
<tr>
<td></td>
<td>• Immediate practice after &quot;errors&quot;</td>
<td>• Delayed practice</td>
</tr>
</tbody>
</table>

The feedback that Coaches provide is context-sensitive within the user’s immediate operations, whereas a Critic’s feedback relates more to the task as a whole, rather than to the individual sub-tasks. At first, it might seem best to provide feedback within the context of a user’s current operation so that the new method can be easily integrated and practiced. However, some researchers believe that the user should not be frequently interrupted in this manner since it may disrupt the coherence of the task and the long-term retention and transfer of skills (Carroll & Kay, 1985; Mobus, Pitschke, & Schroder, 1992; Schmidt & Bjork, 1992). Finally, the frequent feedback from Coaches permits frequent and distributed practice of the new methods whereas the single instance of feedback from Critics only permits massed practice of all of the new methods. Even though massed practice produces better performance during skill acquisition, distributed practice may prevent super-
ficial learning and lead to better long-term retention of skills. Thus, there are several differences between Coach and Critic help systems in terms of feedback, practice, and intrusiveness that are inherent in their designs, some which would suggest that Critics may be more effective for producing long-term retention of skills. However, previous research has not clearly shown how this combination of factors affects the acquisition and retention of skill, especially software skills. Therefore, the study of these issues within the context of intelligent help systems remains open to investigation.

This study investigated some of the issues involved in the design and application of proactive intelligent help systems. It examined the effects of timing and frequency of feedback, context-specificity of feedback, and distribution of practice on the learning, retention, and utilization of efficient software skills. These factors were studied within the framework of three training methods: Tutorial, Coach, and Critic. The participants received training in efficient methods (i.e., how to use styles) either before or during their performance of several tasks using Microsoft Word. The Tutorial method provided this training before the task performance, whereas both the Coach and Critic methods provided the training during the tasks. Word was selected for this study because it is an application that many people have used frequently, however, few people utilize many of its advanced features. Also, word processing provides a variety of tasks that can be successfully accomplished either through very efficient or less efficient methods.

The participants were either given a tutorial on styles before they began working on the word processing tasks (i.e., the Tutorial group) or they received training opportunities while they were working on the tasks (i.e., the Coach and Critic groups). They were tested on their use of efficient methods immediately after
they completed the training tasks and again when they returned two weeks later. In addition, the participants' use of efficient methods during their normal word processing were evaluated with a questionnaire administered approximately eight weeks after the second experimental session.

METHOD

Participants

Sixty-two undergraduate students from Rice University (38 women and 24 men) participated in this study in order to receive research credit for the psychology courses in which they were currently enrolled. The age of the participants ranged from 17 to 40 years ($M = 20.645, SD = 3.412$). All participants had been using Microsoft Word for at least six months ($M = 37.0, SD = 21.050$). They reported using Word about once a week, on average, for tasks such as writing short papers, lab reports, research papers, short essays, printing envelopes, creating tables, resumes, memos, signs, business letters, personal letters, and newspaper articles. They had no experience with creating and applying styles.

Design

A 3 x 3 mixed factorial design was used for this study. The first factor, training method, was between-subjects with three different types of training: Tutorial, Coach, and Critic. The second factor, session, was within-subjects, consisting of three sessions: the first session, a second session two weeks later, and a final session eight weeks after the second. Several measures were recorded during the training tasks and test performance: number of efficient methods (i.e., Styles) used, type of help that was accessed, time spent using help, and task performance time. Questionnaires were used during the first two sessions to acquire several measures such
as the participants’ confidence and impressions of their training method. In addition, the use of styles by the participants for their typical word processing tasks was assessed with a questionnaire administered during the last session. Scores on the Wonderlic Personnel Test, which is a measure of general intelligence, were used as a covariate for the analyses (Wonderlic, 1983).

Materials

**Hardware.** The experiment was conducted using two networked Macintosh® computers in separate rooms. The participant worked on a Macintosh LC II with a 12 inch color monitor in one room while the researcher conducted the study using a Macintosh IIgs with a 14 inch color monitor in a separate observation room. A VHS camera was used to record the participant’s actions and verbal comments. The camera was connected to a monitor in the observation room in order to monitor the participant’s progress. An L•TV video adapter manufactured by Focus Enhancements was used to connect the participant’s monitor to a separate television monitor in the observation room. This allowed the researcher to monitor and record the participant’s actions in the software application while completing each task.

**Software.** The participants completed a number of word processing tasks using Microsoft Word version 5.1a. During the tasks, they were presented with a simulated intelligent help system developed in FaceSpan™ version 2.0. The researcher simulated an intelligent help system by sending AppleScript™ commands from a HyperCard® stack to the help application running in the background on the participant’s computer. The commands opened the appropriate help window over the document window. The help informed the participants of a more efficient method for completing the current task, briefly explained why that method would
be more efficient, and referred them to a page number in the User manual to learn about the method (see Figure 2). There were four different panels of help information: Three for the methods of creating, applying, and modifying a style and one that referred to all three methods.

**Tutorial.** The participants in the Tutorial condition completed a short tutorial on the use of Styles before they began working on the training tasks. Basically, the tutorial taught them how to create, apply, and modify a style (see Appendix A).

**Tasks.** All participants performed a warm-up task that was used to determine their skill level (see Appendix B). The participants in each training condition were given different instructions prior to performing the training tasks. The instructions basically explained that the participants would be performing some word processing tasks and they were told to be efficient while working through the tasks, but not to spend too much time on each one (see Appendix C). There were three separate word processing tasks used as the training tasks. Each consisted of creating a docu-

![Help](image)

**There is a more efficient way to format this paragraph by creating a new Style.**

Creating a style for the paragraph now will make it easier to make changes to this paragraph later. It will also make it easy to create other paragraphs that have the same appearance.

**Please turn to page 15 in the User manual to learn how to create a new Style.**

**FIGURE 2.** Intelligent help window
ment with nine paragraphs (see Appendix D). Three different paragraph styles were used in the documents and each style was shared by three paragraphs. The tasks could be completed using inefficient procedures, but required the use of Word's styles feature for the most efficient performance.

**Tests.** Two tests were used to measure the participants' use of styles. One test was completed immediately after the training tasks during the first session and the second test was completed during the second session two weeks later. Both tests consisted of creating a document containing nine paragraphs with three different paragraph styles, each shared by three paragraphs (See Appendix E). Both documents contained the same number of characters but differed in surface similarity (i.e., different text and formats were used).

In addition, the participants were also given a short procedural knowledge test to explicitly test their knowledge of styles. It contained three sections that asked them to create, modify, and apply a style (see Appendix F).

**Questionnaires.** There were four questionnaires used for this study. A short background questionnaire was used to gather information about the participants and their experience with Word (see Appendix G). Two HyperCard-based questionnaires were used to acquire several measures such as the participants' confidence and their impressions of the intelligent help system, User manual, and tutorial (see Appendices H and I). The second questionnaire included additional questions regarding their use of styles. A final questionnaire was used to gather information about the participants' use of styles in their daily word processing tasks, their confidence, the number of opportunities they had to use styles, reasons for not using styles, and impressions of their training method (see Appendix J).

**User Manual.** The help information was presented in a hard copy form,
rather than on-line, due to the small size of the computer monitor. This allowed the participants to read the help information and still be able to see their document on the computer screen. A three ring binder was used to create a User manual with help sections separated by tabbed dividers. The manual contained pages describing how to create, apply, and modify Styles (see Appendix K). It also contained pages describing other Word functions for realism (e.g., creating tables). A separate application window was used to provide the intelligent help while the participants were performing the tasks. The help suggested a more efficient method (i.e., the use of Styles) and directed the participants to a page number in the User manual.

Procedure

The participants were randomly assigned to one of the three training conditions resulting in 20 participants in the Coach group and 21 participants each in the Critic and Tutorial groups (see Appendix L). They participated in three sessions. The first session lasted approximately one and a half hours, the second session lasted about one hour, and the last session also took approximately one hour. The first two sessions were spaced two weeks apart and the third session took place eight weeks after the second session. Both the participants and their computer screen were videotaped to aid the collection of the dependent measures. During the first session, the participants first completed a consent form and then a short background questionnaire about their experience and training with MS Word (see Appendix G). Then, the participants performed a short warm-up task to determine their level of skill using Word, specifically if they knew how to use styles. Participants without a minimum level of skill (i.e., novices) and participants who exhibited a high level of familiarity with the advanced features of the application (i.e.,
experts) were excused from the study.

Once the warm-up task was completed the participants were given three separate word processing tasks to complete. The Tutorial group received a tutorial on the use of Styles before they began working on the tasks. During the tasks, they had access only to the standard on-line help and the User manual. The Coach and Critic groups also had access to the standard on-line help, but in addition they received feedback during the tasks in the form of a simulated intelligent help system. After the second time the participants used an inefficient method while working on a task, the simulated intelligent help provided to the Coach group informed them each time they performed inefficiently that there was a more efficient method that would save them time. They were then referred to a page number in the User manual for a text-based description of the efficient method (e.g., how to create a Style). Appendix K contains the Styles help presented in the User manual. The participants in the Critic group received this feedback only once while working on a task when they had completed approximately 50% of the task. They were referred to page numbers in the User manual for descriptions of all of the more efficient methods (i.e., how to create, apply, and modify a style) that were related to the task. Thus, the Coach group received feedback at several points during task performance, whereas the Critic group received this feedback only once in the middle of the task. This design gave the Coach group distributed feedback and referred them to individual methods, whereas it gave the Critic group massed feedback and referred them to all of the methods as a whole.

Once all of the tasks were completed, all three groups were given an immediate post-test, which required the use of Styles for the most efficient task performance. The intelligent help system was not available during the test, but the
participants were able to use the on-line help or User manual if they wished. They then completed an online HyperCard-based questionnaire to rate the help, their confidence, and other items (see Appendix H). Upon completion of the questionnaire, they were excused and asked to return for the second experimental session in two weeks.

The second session began with a short warm-up task. Once the participants had completed the task, they were given a retention test that differed in surface similarity from the immediate test, but required the same methods (i.e., the use of Styles) for optimal task performance. They were then given a procedural knowledge test to determine if they knew how to use Styles, even if they had not actually used them during the retention test. When the participants finished this test, they completed an on-line questionnaire similar to the one completed during the first session with the addition of two questions regarding their use of styles during the previous two weeks (see Appendix I). They were then administered the Wonderlic Personnel Test. Upon completion of the test, the participants were reminded of the final session that was scheduled to take place in eight weeks.

The final session consisted of administering a questionnaire to the participants approximately eight weeks after the second experimental session (see Appendix J). The questionnaire asked them to provide information such as their use of styles in their daily word processing tasks, their confidence, and the number of opportunities they had to use styles (e.g., number of papers written). Once the questionnaire was completed, the participants were given a debriefing and thanked for their participation.
RESULTS

The results were analyzed with training method as the between-subjects factor and training tasks and session as within-subject factors. An alpha level of .05 was used for all statistical tests. The dependent measures were the number of participants in each group who accessed help, help use, task performance time, efficiency score, and the responses on the questionnaires. The help use was computed as the proportion of total task time spent using help, since time on task was different for each participant (Shute & Gawlick, 1995). Task performance time was measured as the amount of time it took to complete a task, not including the time spent using help. An efficiency score was computed as the proportion of efficient methods used (i.e., styles used correctly) out of the total number of opportunities to use them, which was nine for each task. Thus, if a participant used styles correctly for all nine paragraphs of text during a task, she would receive an efficiency score of 100%. All of the measures of help use refer only to the use of the User manual. Even though Word’s standard on-line help was available, the participants did not use it. Performance on the Wonderlic Personnel test was used as a covariate for the ANCOVAs for the task performance time measures.

The main effects of training method and session were examined, as well as the interaction of training method with test session for the dependent measures common to all of the groups. Planned comparisons were also performed to examine the differences between the Coach and Critic groups. In addition, the utilization of the help system was examined for the Coach and Critic groups.

Background Questionnaire

As can be seen in Table 2, the demographics of the three groups were fairly
similar. The majority of the participants were 3rd year students at Rice University (42%). Approximately 39% were Psychology majors, with the rest distributed among 18 other majors. About 68% of the participants used version 5.1a of MS Word, 23% used version 6, and the rest frequently used earlier versions of the software. They reported an average of 37 months of experience using MS Word ($SD = 21.050$). They reported using Word an average of about once a week. They used Word for a large variety of tasks such as writing short papers, lab reports, research papers, short essays, printing envelopes, creating tables, resumes, memos, signs, business letters, personal letters, and newspaper articles. The Coach and Tutorial groups scored slightly higher on the Wonderlic than did the Critic group, but the difference was not significant, $F (2, 59) = 1.190, p = .3113$.

<table>
<thead>
<tr>
<th>TABLE 2. Demographic information</th>
<th>Coach</th>
<th>Critic</th>
<th>Tutorial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Participants</td>
<td>20</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Mean Age</td>
<td>20.4 (2.5)</td>
<td>20.7 (2.9)</td>
<td>20.9 (4.6)</td>
</tr>
<tr>
<td>Men</td>
<td>45%</td>
<td>38%</td>
<td>33%</td>
</tr>
<tr>
<td>Women</td>
<td>55%</td>
<td>62%</td>
<td>67%</td>
</tr>
<tr>
<td>Experience using MS Word</td>
<td>40.7 (16.1)</td>
<td>31.9 (21.8)</td>
<td>38.6 (24.2)</td>
</tr>
<tr>
<td>(months)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often MS Word used</td>
<td>3.2 (1.0)</td>
<td>3.2 (1.2)</td>
<td>3.4 (1.1)</td>
</tr>
<tr>
<td>(scale of 1 to 5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wonderlic score</td>
<td>31.7 (4.8)</td>
<td>29.0 (7.4)</td>
<td>31.2 (5.2)</td>
</tr>
</tbody>
</table>

*Note.* Standard deviations are in parentheses

The participants were asked how they had learned to use MS Word. The overwhelming majority reported that they had taught themselves to use the application through trial and error (79%), 11% had learned from others, and the remaining participants had used other methods such as classes, reading manuals, and tutorials.
They were also asked to rate a number of methods for learning software in terms of how useful they thought it would be for them. Thirty-two percent of the participants rated “Learning from others” as the first most useful method and another 40% rated it as the second most useful method. The other two most popular methods were “Learning as you work, through trial and error”, which was rated number one by 29% and number two by 16%, and “Instructor-led training”, which was rated number one by 18% and number two by 19% of the participants. The other methods were only rated highly by a small number of the participants. Roughly in order, the remaining participants tended to prefer “On-line tutorials”, “Hard copy manuals”, and then “On-line help”. “Hard copy tutorials” and “Videotape training” were only selected by one or two participants.

Learning Measures

Performance for Training Tasks. The data for the performance times for the training tasks were skewed (see Figure 8, Figure 9, and Figure 10 in Appendix M), so they were transformed to satisfy normality and homogeneity of variance assumptions using a log(x+1) transformation. The suitability of the transformation was determined by the Shapiro-Wilk statistic test of normality and a test of homogeneity of variance. However, all of the figures and tables of means are based on the raw scores. Table 3 shows the mean and tri-mean for the task performance times for each group. The tri-mean and semi-interquartile range are better measures of central tendency and dispersion than the mean and standard deviation for skewed distributions since they are more resistant to extreme scores (Napier, Batsell, Lane, & Guadagno, 1992). Since the data for performance times were skewed, the tri-mean was used for all figures that show these measures. The tri-mean statistic is computed using the following formula (Q1, Q2, and Q3 are the first, second,
and third percentiles):

\[ Q1 + (2 \times Q2) + Q3 \]

\[ \frac{4}{4} \]

<p>| TABLE 3. Performance times for Tasks 1, 2, and 3 (sec) |
|-----------------------------------------------|--|--|------------------|</p>
<table>
<thead>
<tr>
<th>Tasks</th>
<th>Coaching</th>
<th>Critic</th>
<th>Tutorial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1</td>
<td>Mean</td>
<td>444 (140.7)</td>
<td>516 (210.7)</td>
</tr>
<tr>
<td></td>
<td>Tri-mean</td>
<td>424 (90.0)</td>
<td>477 (99.5)</td>
</tr>
<tr>
<td>Task 2</td>
<td>Mean</td>
<td>254 (116.0)</td>
<td>393 (175.2)</td>
</tr>
<tr>
<td></td>
<td>Tri-mean</td>
<td>245 (61.8)</td>
<td>367 (125.0)</td>
</tr>
<tr>
<td>Task 3</td>
<td>Mean</td>
<td>433 (93.7)</td>
<td>611 (251.4)</td>
</tr>
<tr>
<td></td>
<td>Tri-mean</td>
<td>429 (56.5)</td>
<td>578 (210.5)</td>
</tr>
</tbody>
</table>

*Note.* Standard deviations and semi-interquartile ranges are in parentheses.

The Tutorial group completed all of the tasks more quickly than the Critic and Coach groups (see Figure 3). However, the mean performance time of the

![FIGURE 3. Training group by task interaction (performance time tasks 1-3)](image-url)
Coach group was only longer than the Tutorial group by a matter of seconds. The Critic group's performance time was considerably longer than the other two groups for each task. The Wonderlic scores correlated significantly with the performance times for Task 1 \( (r = -0.5115, p < 0.001) \), Task 2 \( (r = -0.2550, p = 0.0455) \), and Task 3 \( (r = -0.2669, p = 0.0360) \). An analysis of covariance (with the Wonderlic scores as the covariate) revealed a significant main effect of training, \( F(2, 58) = 4.476, p = 0.0156 \), and task, \( F(2, 116) = 3.466, p = 0.0345 \). Since the tasks were quite different, the task effect is of little interest. There was no significant Training x Task interaction, \( F(4, 116) = 1.865, p = 0.1212 \). There was no evidence of a violation of homogeneity of slopes for Task 1 \( (F(2, 56) = 1.8095, p = 0.1732) \), Task 2 \( (F(2, 56) = 1.7056, p = 0.1910) \), or Task 3 \( (F(2, 56) = 1.3348, p = 0.2715) \).

A planned comparison of the Coach and Critic groups revealed a significant training effect, \( F(1, 38) = 8.802, p = 0.0052 \), task effect, \( F(2, 76) = 3.139, p = 0.049 \), and Training x Task interaction, \( F(2, 76) = 3.176, p = 0.0473 \). There was no significant difference between the two groups for performance time on Task 1, \( F(1, 38) = 0.488, p = 0.4891 \). However, the average performance time for the Coach group was significantly shorter than the Critics group's performance for Task 2, \( F(1, 38) = 8.154, p = 0.0069 \), and Task 3, \( F(1, 38) = 6.686, p = 0.0137 \).

An efficiency score was computed as the proportion of efficient methods used out of the total number of opportunities to use styles (i.e., nine for each task). Table 4 shows the mean and tri-mean for the efficiency scores for each group. The Tutorial group had the highest mean efficiency scores for the first two tasks, followed by the Coach group. The Critic group had the lowest efficiency scores for all tasks. The Coach group's efficiency scores improved across tasks such that their mean efficiency score for the last task was higher than the other two groups (see
TABLE 4. Efficiency scores for Tasks 1, 2, and 3 (100% maximum)

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Coach</th>
<th>Critic</th>
<th>Tutorial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1</td>
<td>Mean</td>
<td>47.22 (39.14)</td>
<td>35.45 (38.91)</td>
</tr>
<tr>
<td></td>
<td>Tri-mean</td>
<td>47.22 (38.89)</td>
<td>19.44 (38.89)</td>
</tr>
<tr>
<td>Task 2</td>
<td>Mean</td>
<td>60.00 (41.64)</td>
<td>26.98 (36.44)</td>
</tr>
<tr>
<td></td>
<td>Tri-mean</td>
<td>63.89 (44.44)</td>
<td>11.11 (22.22)</td>
</tr>
<tr>
<td>Task 3</td>
<td>Mean</td>
<td>83.89 (36.49)</td>
<td>43.39 (46.20)</td>
</tr>
<tr>
<td></td>
<td>Tri-mean</td>
<td>100.00 (0.00)</td>
<td>36.11 (50.00)</td>
</tr>
</tbody>
</table>

*Note.* Standard deviations and semi-interquartile ranges are in parentheses

Since the data for efficiency scores were skewed, transformations were attempted to satisfy normality and homogeneity of variance assumptions (see Figure 11, Figure 12, and Figure 13 in Appendix M). However, even after the transformations, the scores were still far from meeting the assumption of normality.

**FIGURE 4.** Training group by task interaction (efficiency score tasks 1-3)
Therefore, a nonparametric Kruskal-Wallis test was used to test the main effect of training (Howell, 1995). There was a significant difference among the groups for efficiency score for Task 1, $\chi^2 (2) = 11.022, p = .004$, Task 2, $\chi^2 (2) = 8.906, p = .012$, and Task 3, $\chi^2 (2) = 10.169, p = .006$. A comparison of the Coach and Tutorial groups for Task 1, using the Mann-Whitney test, revealed that the difference in efficiency scores was significant, $\chi^2 (1) = 5.113, p = .024$. A planned comparison of the Coach and Critic groups, using the Mann-Whitney test, did not reveal a significant difference in efficiency scores for Task 1, $\chi^2 (1) = 0.976, p = .323$. However, the difference between the two groups was significant for both Task 2, $\chi^2 (1) = 6.016, p = .014$, and Task 3, $\chi^2 (1) = 8.014, p = .005$, with the Coach group performing better than the Critic group.

Friedman's Rank test was used to test the effect of task. There was a significant main effect of task since there was an obvious overall increase in the number of efficient methods used, $\chi^2 (2) = 9.710, p = .0078$. However, the effect of task was only significant for the Coach group, $\chi^2 (2) = 11.425, p = .0033$. The increase in the number of efficient methods used was not significant for the Critic group, $\chi^2 (2) = 1.881, p = .3904$, or the Tutorial group, $\chi^2 (2) = 1.238, p = .5385$. The Coach group's improvement from Task 1 to Task 2 was not significant, $\chi^2 (1) = 0.200, p = .6547$, but the improvement from Task 2 to Task 3 was significant, $\chi^2 (1) = 4.050, p = .0442$.

**Performance for Tests.** Performance on the immediate and delayed tests were used to measure the styles usage of the participants when the proactive help was no longer available and how well they would retain what they had learned two weeks later. Since the data for the performance times for the tests were skewed (see Figure 14 and Figure 15 in Appendix M), they were transformed to satisfy normal-
ity and homogeneity of variance assumptions using a log(x+1) transformation.

Table 5 shows the mean and tri-mean for the test performance times for each group.

<table>
<thead>
<tr>
<th>Tests</th>
<th>Coach</th>
<th>Critic</th>
<th>Tutorial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>Mean</td>
<td>419 (89.5)</td>
<td>475 (161.5)</td>
</tr>
<tr>
<td></td>
<td>Tri-mean</td>
<td>416 (78.0)</td>
<td>434 (109.5)</td>
</tr>
<tr>
<td>Test 2</td>
<td>Mean</td>
<td>392 (85.2)</td>
<td>478 (154.1)</td>
</tr>
<tr>
<td></td>
<td>Tri-mean</td>
<td>384 (55.0)</td>
<td>448 (107.5)</td>
</tr>
</tbody>
</table>

*Note.* Standard deviations and semi-interquartile ranges are in parentheses

The Coach group completed the tests more quickly than the Critic and Tutorial groups for both sessions (see Figure 5). The Tutorial group was the slowest group for the first session, whereas the Critic group performed slightly slower than the Tutorial group for the second session. In addition, the Coach group improved

![Graph showing performance times for Tests 1 and 2.](image)

**FIGURE 5.** Training group by session interaction (performance time tests 1 and 2)
their performance time between the first and second test, whereas both the Critic and Tutorial groups performed more slowly on the second test. The Wonderlic scores correlated significantly with the performance times for Test 1 \( r = -.4936, p < .0001 \) and Test 2 \( r = -.4039, p = .0011 \). However, an analysis of covariance (with the Wonderlic scores as the covariate) revealed that there was no significant main effect of training, \( F (2, 58) = 1.168, p = .3183 \), or session, \( F (1, 58) = 1.686, p = .1993 \). A test of the Training x Session interaction was not significant, \( F (2, 58) = 1.527, p = .2259 \). There was no evidence of a violation of homogeneity of slopes for Test 1 \( F (2, 56) = 0.5542, p = .5777 \) or Test 2 \( F (2, 56) = 1.2998, p = .2807 \).

A planned comparison of the Coach and Critic groups did not find a significant training effect, \( F (1, 38) = 1.774, p = .1908 \), session effect, \( F (1, 38) = 3.832, p = .0577 \), or interaction, \( F (1, 38) = 2.975, p = .0927 \).

Separate analyses of covariance were done for each session. There was no significant difference among the three groups for time spent performing the test during Session 1, \( F (2, 58) = 0.242, p = .7859 \), or Session 2, \( F (2, 58) = 2.215, p = .1183 \). A planned comparison of the Coach and Critic groups was not significant for the first session, \( F (1, 38) = 0.353, p = .5558 \). The difference between the Coach and Critic groups approached significance for the second session, \( F (1, 38) = 3.484, p = .0697 \).

An efficiency score was computed as the proportion of efficient methods used out of the total number of opportunities to use styles (i.e., nine for each test). Table 6 shows the efficiency scores (i.e., means and tri-means) for each group. The Tutorial group had the highest efficiency scores for both tests, followed by the Coach group. The Critic group had the lowest efficiency scores for both tests and the mean scores did not change across the two sessions. However, both the Coach
and Tutorial groups improved their mean efficiency scores slightly between the first and second tests (see Figure 6).

<table>
<thead>
<tr>
<th>Tests</th>
<th>Coach</th>
<th>Critic</th>
<th>Tutorial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>Mean</td>
<td>73.33 (39.39)</td>
<td>50.79 (45.48)</td>
</tr>
<tr>
<td></td>
<td>Tri-mean</td>
<td>88.89 (16.67)</td>
<td>58.34 (50.00)</td>
</tr>
<tr>
<td>Test 2</td>
<td>Mean</td>
<td>74.44 (39.67)</td>
<td>49.74 (45.76)</td>
</tr>
<tr>
<td></td>
<td>Tri-mean</td>
<td>91.67 (16.67)</td>
<td>58.34 (50.00)</td>
</tr>
</tbody>
</table>

*Note.* Standard deviations and semi-interquartile ranges are in parentheses.

The data for efficiency scores were skewed, therefore transformations were attempted to satisfy normality and homogeneity of variance assumptions (see Figure 16 and Figure 17 in Appendix M). However, after the transformations, the scores were still far from satisfying the assumption of normality. Therefore, a non-

![Graph showing efficiency scores](image)

**Figure 6.** Training group by session interaction (efficiency score tests 1 and 2)
parametric Kruskal-Wallis test was used to test the main effect of training. There was no significant difference among the groups for efficiency score during Session 1, $\chi^2 (2) = 3.549, p = .1696$. However, the difference was significant for Session 2, $\chi^2 (2) = 6.267, p = .0436$. Thus, there was a significant difference among the three training groups for the efficiency score on the second test. A planned comparison of the Coach and Critic groups, using a Mann-Whitney test, did not reveal a significant effect of training for the first test, $\chi^2 (1) = 1.879, p = .1705$, or the second test, $\chi^2 (1) = 3.164, p = .0753$.

Friedman’s Rank test was used to test the effect of session. No main effect of session was found, $\chi^2 (2) = 0.065, p = .7995$. The effect of session (i.e., change in efficiency scores) was not significant for any of the groups; $\chi^2 (1) < 0.000, p > 1.0000$ for both the Coach and Critic groups and $\chi^2 (1) = 0.191, p = .6625$ for the Tutorial group.

Knowledge Test. The participants were administered a Knowledge test at the end of the second session, which required the performance of three separate tasks: Creating, modifying, and applying a style. This was used to determine if the participants had learned how to use styles, but for some reason had not applied them during the earlier tasks. The Tutorial group performed the best for all tasks, followed by the Coach group, then the Critic group (see Figure 7). In the Tutorial group 18 of the 21 participants (86%) completed the task of creating a style correctly, whereas only 16 of the 20 participants (80%) completed the task correctly in the Coach group. In the Critic group, only 13 of the 21 participants (62%) completed this task correctly. A Chi-Square test of the difference in the proportions was not significant, $\chi^2 (2) = 3.529, p = .1713$. The second task involved modifying an existing style. In the Tutorial group 9 out of 21 (43%) completed the task of modifying a
style correctly and only 8 out of 20 (40%) completed the task correctly in the Coach group. In the Critic group, only 6 out of 21 (29%) completed this task correctly. This difference was not significant, $\chi^2 (2) = 1.0249, p = .5990$. The last task involved applying an existing style. In the Tutorial group 19 out of 21 (90%) completed the task of applying a style correctly and 17 out of 20 (85%) completed the task correctly in the Coach group. In the Critic group, only 13 out of 21 (62%) completed this task correctly. This difference in proportions approached significance, $\chi^2 (2) = 5.8070, p = .0548$.

Analysis of Help Use

The number of participants in each group that accessed help and the percentage of task time spent time using help were examined for Tasks 1, 2, and 3. Table 7 shows the percentage of participants in each group that accessed help. Since none of the participants used the standard on-line help, all help use refers to reading the

![Figure 7](image_url)

**FIGURE 7.** Percentage of participants performing knowledge test tasks correctly
User manual. Only one person in the Tutorial group looked at the manual, however several of the participants in the Coach and Critic groups referred to it during the tasks. A higher percentage of the participants in the Coach group used help during Task 1, however, a higher percentage of the Critic group used help during Tasks 2 and 3. A Chi-Square test of the difference in the proportions was not significant for Task 1, $\chi^2 (1) = 0.176, p = .6751$, Task 2, $\chi^2 (1) = 0.749, p = .3869$, or Task 3, $\chi^2 (1) = 0.344, p = .5578$.

The percentage of total task time spent using help was analyzed for those participants in the Coach and Critic groups that accessed help. The Coach group spent a slightly larger percentage of their time on Task 1 using help ($M = 26.028, SD = 12.451$) than the Critic group did ($M = 22.810, SD = 13.688$), but the difference was not significant, $t (33.7) = 0.738, p = .466$. The Coach group also spent a greater percentage of their time on Task 2 using help ($M = 34.704, SD = 24.254$) than the Critic group did ($M = 22.246, SD = 14.313$), however this difference was not significant, $t (21.8) = 1.751, p = .076$. For Task 3, the Coach group spent a considerably larger percentage of their time using help ($M = 35.478, SD = 16.767$) than the Critic group did ($M = 9.435, SD = 6.140$). This difference was significant, $t (4.8) = 3.318, p = .023$.

The number of participants in each group that accessed help was also exam-
ined for Tests 1 and 2. Table 8 shows the percentage of participants in each group that used help. However, very few participants actually accessed help during the tests. No one in the Tutorial group looked at the manual during the tests and only a

<table>
<thead>
<tr>
<th>Tests</th>
<th>Coach</th>
<th>Critic</th>
<th>Tutorial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>0%</td>
<td>14%</td>
<td>0%</td>
</tr>
<tr>
<td>Test 2</td>
<td>19%</td>
<td>5%</td>
<td>0%</td>
</tr>
</tbody>
</table>

small number of the participants in the Coach and Critic groups referred to it. A higher percentage of the participants in the Critic group used help during Test 1, whereas a higher percentage of the Coach group used help during Test 2. A Chi-Square test of the difference in the proportions was not significant for Test 1, $\chi^2 (1) = 3.083, p = .0791$, or Test 2, $\chi^2 (1) = 2.221, p = .1361$.

**Behavior**

The participants completed three different questionnaires, one during each session. Several questions were asked that measured their impressions of their training method, which will be discussed in the next session. The participants were also asked questions to measure how they were applying the methods that they had learned to their personal word processing tasks (i.e., did the training method influence their behavior?). Confidence ratings were included to provide a measure of self-efficacy (i.e., the belief in one's ability to perform a certain task), which has been shown to have a strong influence on actual behavior (Gist, Rosen, & Schwoerer, 1989).

**First Session.** Table 9 displays the mean ratings for the two confidence questions asked on the questionnaire administered during the first session. The differ-
ences among the groups were significant for “how confident” they were that they had learned how to use styles, \( F(2, 59) = 4.979, p = .0100 \); and “how confident” they were that they would actually use styles to format their own documents, \( F(2, 59) = 3.297, p = .0439 \). A planned comparison of the Coach and Critic groups revealed that only the difference in the learning confidence rating approached sig-

<table>
<thead>
<tr>
<th>Question</th>
<th>Coach(^a)</th>
<th>Critic(^b)</th>
<th>Tutorial(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>How confident are you that you have learned styles?</td>
<td>7.45</td>
<td>5.76(^c)</td>
<td>8.10(^b)</td>
</tr>
<tr>
<td>How confident are you that you will use styles?</td>
<td>6.85</td>
<td>5.91(^c)</td>
<td>7.91(^b)</td>
</tr>
</tbody>
</table>

*Note.* Scale of 1 = None/Not at all to 10 = Very. Significant differences indicated by superscript group letters.

nificance, \( F(1, 39) = 3.736, p = .0605 \), with the Coach group rating their confidence that they had learned styles higher than the Critic group.

**Second Session.** Table 10 displays the means for the four questions of interest for the questionnaire administered during the second session. In general, the

<table>
<thead>
<tr>
<th>Question</th>
<th>Coach(^a)</th>
<th>Critic(^b)</th>
<th>Tutorial(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>How confident are you that you have learned styles?</td>
<td>7.45</td>
<td>5.86(^c)</td>
<td>8.24(^b)</td>
</tr>
<tr>
<td>How confident are you that you will use styles?</td>
<td>6.45</td>
<td>5.48(^c)</td>
<td>8.00(^b)</td>
</tr>
<tr>
<td>How often have you used styles in the past 2 weeks?</td>
<td>2.00</td>
<td>1.52</td>
<td>2.48</td>
</tr>
<tr>
<td>How many opportunities have you had where you could have used styles in the past 2 weeks?</td>
<td>2.75</td>
<td>3.05</td>
<td>2.52</td>
</tr>
</tbody>
</table>

*Note.* Scale of 1 = None/Not at all to 10 = Very. Significant differences indicated by superscript group letters.
Tutorial group’s responses were again the highest, followed by the Coach group. The responses for the Critic group were generally the lowest. The differences among the groups were only significant for “how confident” they were that they had learned how to use styles, $F (2, 59) = 3.154, p = .0500$; and “how confident” they were that they would actually use styles to format their own documents, $F (2, 59) = 4.303, p = .0180$. The Coach group had slightly higher mean ratings than the Critic group for every question, with the exception of the “Opportunities” question. However, a planned comparison of two groups did not reveal any significant differences. A ratio was calculated for the rating of how often styles had been used during the past two weeks over the rating of opportunities to use styles. There was a significant difference among the groups for this ratio ($F (2, 59) = 3.8953, p = .0258$). A Tukey HSD post-hoc test revealed that the difference between the Critic and Tutorial groups was significant ($p = .0236$), with the Tutorial group using styles more often when the opportunity presented itself.

**Third Session.** Table 11 displays the means for five of the questions asked on the questionnaire administered during the third session. In general, the Tutorial

<table>
<thead>
<tr>
<th>Question</th>
<th>Coach^a</th>
<th>Critic^b</th>
<th>Tutorial^c</th>
</tr>
</thead>
<tbody>
<tr>
<td>How often have you been performing word processing tasks during the last 8 weeks?</td>
<td>5.25</td>
<td>6.24</td>
<td>5.52</td>
</tr>
<tr>
<td>How confident are you that you have learned styles?</td>
<td>6.85</td>
<td>5.62^c</td>
<td>7.67^b</td>
</tr>
<tr>
<td>How confident are you that you will use styles?</td>
<td>5.45</td>
<td>5.14</td>
<td>6.43</td>
</tr>
<tr>
<td>How often have you used styles in the past 2 weeks?</td>
<td>2.75</td>
<td>2.14</td>
<td>3.19</td>
</tr>
<tr>
<td>How many opportunities have you had where you could have used styles in the past 2 weeks?</td>
<td>3.00</td>
<td>3.52</td>
<td>3.33</td>
</tr>
</tbody>
</table>

*Note.* Scale of 1 = None/Not at all to 10 = Very. Significant differences indicated by superscript group letters.
group's responses were once again the highest, followed by the Coach group. The responses for the Critic group were still generally the lowest. The only difference among the groups that was significant was the rating of "how confident are you that you have learned styles", $F (2, 59) = 3.555, p = .0349$, with the Tutorial group still rating their confidence higher than the other two groups. The Coach group's mean learning confidence rating was slightly higher than the Critic group's mean rating, but this difference was not significant, $F (1, 39) = 1.956, p = .1699$. None of the other planned comparisons of the Coach and Critic groups revealed any significant differences. A ratio was again calculated for the rating of how often styles had been used over the opportunities to use styles. There were no significant differences among the groups for this ratio ($F (2, 59) = 0.0459, p = .9551$).

The effect of session and the interaction of training and session were examined for the two questions that were asked for all sessions regarding confidence in learning styles and confidence in actual use of styles. Both the Tutorial and Critic groups slightly increased their mean ratings of confidence in learning styles between the first and second sessions, whereas the Coach group's mean rating was unchanged. All three groups rated their confidence as slightly lower by the third session. However, neither the effect of session, $F (2, 118) = 2.365, p = .0984$, nor the interaction of training method and session was significant, $F (4, 118) = 0.206, p = .9347$. In general, all three groups' mean ratings of their confidence that they would actually use styles for their own word processing decreased across the three sessions. Thus, only the effect of session was significant, $F (2, 118) = 10.426, p = .0001$. A trend analysis revealed a significant linear component of the session effect (i.e., a negative linear effect), $F (1, 59) = 16.203, p = .0002$.

The effect of session and the interaction of training and session were also
examined for the two questions that were asked for the last two sessions regarding
the use of styles and the opportunity to use them. The mean ratings for the self-
report of the use of styles during the two weeks prior to each session increased
slightly for all groups between the second and third sessions, with the effect of ses-
sion approaching significance, $F (1, 59) = 3.728, p = .0583$. The mean ratings for
the self-report of the number of opportunities to use styles also increased slightly
for all groups between the second and third sessions. Once again, the effect of ses-
sion approached significance, $F (1, 59) = 3.754, p = .0575$.

The participants were asked questions regarding their word processing
activity and styles use on the questionnaire during the last session. Eighteen partici-
pants in the Coach group reported writing an average of 8 papers using MS Word
($M = 8.056, SD = 7.248$) during the eight weeks prior to the last session. Nineteen
participants in the Critic group reported writing an average of 5 papers using MS
Word ($M = 5.263, SD = 5.516$) and 18 participants in the Tutorial group reported
writing an average of 6 papers using MS Word ($M = 6.056, SD = 4.399$). The differ-
eence among the groups in the number of papers written was not significant, $F (2,
52) = 1.116, p = .3353$. The participants were asked to estimate how many of the
total number of papers written had been formatted using styles. This was included
as a measure of transfer of training. In other words, did they transfer what they had
learned during the training session to their own word processing behavior in the
"real world"? Approximately 40% of the participants in the Coach group, 24% of
the participants in the Critic group, and 48% of the participants in the Tutorial
group reported that they had used styles while writing these papers. A Chi-Square
test of the difference in the group proportions was not significant, $\chi^2 (2) = 2.658, p
= .2648$. When asked how they had used styles, they reported using them for tasks
such as creating section headings, outlines, references, quotations, and modifying MS Word's "Normal" style.

The participants who had not used styles named a variety of reasons for why they did not use them. One of the main reasons given was that many of their papers were simple and used the same formats throughout. Some other reasons given were that they had not learned styles, they had reverted to their old methods, they wrote papers quickly and forgot to use styles, or that working with styles in other word processing applications was confusing. They were also asked to describe situations that they thought would persuade them to learn and/or use styles. They most commonly reported that they would use styles for papers that contained different fonts, headings, and sections (e.g., formal reports, research papers, resumes). Several participants also reported that they would learn and use styles if they had more time, if the software applications that they used easily supported them, or if they used a Macintosh computer because they thought styles were too difficult to use in the Windows version of the software.

The 23 participants who had used styles were asked what percentage of the total number of papers that they had written had been formatted, at least partially, using styles. The Coach group reported a range of approximately 7% to 100% of their papers had been written using styles ($M = 44.488$, $SD = 41.266$), the Critic group reported a range of 25% to 67% ($M = 45.060$, $SD = 16.362$), and the Tutorial group reported a range of 10% to 100% ($M = 55.750$, $SD = 35.234$). The differences among the groups were not significant, $F (2, 20) = 0.2856$, $p = .7546$. Interestingly, the ratings on the second session questionnaire for all participants regarding how confident they were that they would actually use styles had a significant positive correlation with the reported percentage of papers written using
styles, $r = .3616, p = .0039$. Thus, participants who had reported higher levels of self-efficacy eight weeks before the last session also tended to report that they had actually used styles for a larger percentage of their papers.

Reactions to Training

First Session. The participants completed three different questionnaires, one during each session, that measured their impressions of the intelligent help system, User manual, and tutorial, and their training method. Some of the comments from the participants regarding what they liked and disliked about the training methods and the User manual are listed in Appendix N. Table 12 displays the means for the four questions of interest for the questionnaire administered during the first session. In general, the Tutorial group’s responses were the highest, followed by the Coach group, and the responses for the Critic group were generally the lowest. The differences among the groups were only significant for “how helpful” the participants found their training method, $F (2, 59) = 8.920, p = .0004$. A planned comparison of the Coach and Critic groups revealed only a significant difference in the rating for

<table>
<thead>
<tr>
<th>Question</th>
<th>Coach$^a$</th>
<th>Critic$^b$</th>
<th>Tutorial$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>How helpful was the tutorial or help system?</td>
<td>6.95$^b$</td>
<td>4.57$^{ac}$</td>
<td>8.14$^b$</td>
</tr>
<tr>
<td>How much did you like the tutorial or help system?</td>
<td>6.35</td>
<td>5.14</td>
<td>7.14</td>
</tr>
<tr>
<td>How much did you like the User Manual?</td>
<td>6.45</td>
<td>5.71</td>
<td>5.62</td>
</tr>
<tr>
<td>How valuable do you think styles are?</td>
<td>8.05</td>
<td>7.48</td>
<td>8.67</td>
</tr>
</tbody>
</table>

*Note. *Scale of 1 = None/Not at all to 10 = Very. Significant differences indicated by superscript group letters.

“how helpful” they thought their training method was, $F (1, 39) = 5.794, p = .0209$, with the Coach group rating their method as more helpful.
Second Session. Table 13 displays the means for the four questions of interest for the questionnaire administered during the second session. In general, the Tutorial group's responses were again the highest, followed by the Coach group. The responses for the Critic group were generally the lowest. The differences among the groups were only significant for “how helpful” the participants found their training method, $F (2, 59) = 5.521, p = .0063$. The Coach group had slightly higher mean ratings than the Critic group for every question. However, a planned comparison of the two groups did not reveal any significant differences.

<table>
<thead>
<tr>
<th>Question</th>
<th>Coach$^a$</th>
<th>Critic$^b$</th>
<th>Tutorial$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>How helpful was the tutorial or help system?</td>
<td>6.30</td>
<td>5.52$^c$</td>
<td>8.00$^b$</td>
</tr>
<tr>
<td>How much did you like the tutorial or help system?</td>
<td>6.05</td>
<td>5.81</td>
<td>7.05</td>
</tr>
<tr>
<td>How much did you like the User Manual?</td>
<td>6.45</td>
<td>5.67</td>
<td>4.67</td>
</tr>
<tr>
<td>How valuable do you think styles are?</td>
<td>7.70</td>
<td>7.05</td>
<td>8.10</td>
</tr>
</tbody>
</table>

*Note. Scale of 1 = None/Not at all to 10 = Very. Significant differences indicated by superscript group letters.*

The effect of session and the interaction of training and session were examined for the questions of interest that were asked for both the first and second sessions (see Table 9 and Table 10). There was a significant interaction of training and session for the “how helpful” rating, $F (2, 59) = 3.778, p = .0286$. Both the Coach and Tutorial groups rated their training method as slightly less helpful between the first and second sessions, but this change was not significant, $F (1, 19) = 2.742, p = .1142$ and $F (1, 20) = 0.089, p = .7688$ respectively. However, the Critic group rated their training method as more helpful for the second session and this increase between sessions was significant, $F (1, 20) = 6.250, p = .0212$. For the question of “how much did you like the tutorial or help system”, both the Coach and Tutorial
groups had slightly lower mean ratings for the second session whereas the Critic group had a slightly higher mean rating, however this interaction was not significant, $F (2, 59) = 2.285, p = .1107$. For the question of “how much did you like the User Manual information”, the Tutorial group lowered their mean rating by the second session, whereas the Critic group’s mean rating only decreased slightly and the Coach group’s mean rating was unchanged. A test of the interaction was not significant, $F (2, 59) = 1.768, p = .1796$.

**Third Session.** The effect of session and the interaction of training and session were examined for the question regarding the value of styles across all three sessions. The mean ratings for all groups decreased significantly across sessions, $F (2, 118) = 13.888, p < .0001$. A trend analysis again revealed a significant linear component of the session effect (i.e., a negative linear effect), $F (1, 59) = 19.923, p < .0001$. Thus, all groups reduced the value rating across sessions. However, the participants who reported using styles for more of their papers also tended to rate styles as more valuable, $r = .4070, p = .0010$.

The participants were asked if they would have preferred a different training method than the one they received during the first session. Approximately 40% of the participants in the Coach group, 43% of the participants in the Critic group, and only 14% of the participants in the Tutorial group reported that they would have preferred a different training method. A Chi-Square test of the difference in the group proportions was not significant, $\chi^2 (2) = 4.732, p = .0938$. These participants were then asked what method they would have preferred. Three of these participants in the Coach group said that they would have preferred some type of tutorial whereas the rest would have preferred either a short introduction to styles and a chance to read about them before exposure to the Coach, a class, one-on-one
instruction, or a user-initiated help system. Most of the participants in the Critic group who would have preferred a different training method said that they would have preferred some type of one-on-one informal instruction and only one would have preferred a formal class. The few participants in the Tutorial group who would have preferred a different training method said that they would like one-on-one instruction, a chance to write down their own notes to remember commands, and one participant thought that the Coach training would have been very helpful.

The participants in the Coach group who were satisfied with their training method stated that they thought the Coach was “effective”, “helpful”, ”worked well”, and was good for learning and an “initial introduction” to the new methods. The satisfied participants in the Critic group stated that they thought the Critic was a “good way to learn”, a “good method”, “worked well”, “helpful, “useful”, and one participant stated that she “never would have known that styles existed without it”. Most of the participants in the Tutorial group were happy with the tutorial. They stated that it was “good”, “effective”, “helpful”, and “worked well”.

The participants who rated their training method as more helpful tended to have higher confidence that they had learned styles ($r = .5325. p < .0001$), confidence that they would use styles ($r = .3794. p = .0024$), and performed more efficiently during the retention test ($r = .5829. p < .0001$). Interestingly, there was a significant positive correlation between the rating of helpfulness and the number of papers that the participants had written in the eight weeks prior to the last session ($r = .2597. p = .0415$). The rating of how much they liked their training method followed a similar pattern of correlations. The participants with higher ratings of liking tended to have higher confidence that they had learned styles ($r = .5289. p < .0001$), confidence that they would use styles ($r = .3447. p = .0061$), and performed
more efficiently during the retention test ($r = .5179, p < .0001$). There was also an interesting relationship between the number of times that the intelligent help was presented to the participants in the Coach and Critic groups and the helpfulness rating ($r = -.3360, p < .0317$) and rating of liking ($r = -.4836, p = .0014$). The more times that the intelligent help was presented, the lower the participant rated the training method's helpfulness and the less they liked the training. This relationship becomes clearer when noting that the participants who were presented help more often also tended to have performed less efficiently on Test 1 ($r = -.4366, p = .0043$) and Test 2 ($r = -.3532, p = .0235$). Thus, these participants had been presented with help frequently during training since they were not performing efficiently. They did not learn how to use styles, as evidenced by the lower efficiency during the tests, and therefore did not rate the training method as very helpful or likeable.

DISCUSSION

This study compared three different types of training (i.e., Tutorial, Coach, and Critic) across three sessions, examining their performance on three training tasks and two tests. Table 12 shows the overall ranking of the three groups for each of the performance measures. Note that this table shows the pattern of the results only, not the significance of each measure. The participants in the Tutorial condition tended to perform the best on the training tasks and test. This is not surprising since they were explicitly trained in all of the required, efficient procedures. However, the participants in the Coach condition quickly learned the new methods and completed the training tasks almost as quickly as the Tutorial group. In fact, by the last training task they were performing at as high a level of efficiency as the Tutorial group. The Coach group consistently performed better than the Critic group
during the training tasks. The groups performed similarly for the first task, but the

<table>
<thead>
<tr>
<th>Measure</th>
<th>Coach&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Critic&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Tutorial&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance Time</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 1</td>
<td>2nd</td>
<td>3rd</td>
<td>1st</td>
</tr>
<tr>
<td>Task 2</td>
<td>2nd&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3rd&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1st</td>
</tr>
<tr>
<td>Task 3</td>
<td>2nd&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3rd&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1st</td>
</tr>
<tr>
<td>Test 1</td>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
</tr>
<tr>
<td>Test 2</td>
<td>1st</td>
<td>3rd</td>
<td>2nd</td>
</tr>
<tr>
<td>Efficiency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task 1</td>
<td>2nd&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3rd&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1st&lt;sup&gt;ab&lt;/sup&gt;</td>
</tr>
<tr>
<td>Task 2</td>
<td>2nd&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3rd&lt;sup&gt;ac&lt;/sup&gt;</td>
<td>1st&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Task 3</td>
<td>1st&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3rd&lt;sup&gt;ac&lt;/sup&gt;</td>
<td>2nd&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Test 1</td>
<td>2nd</td>
<td>3rd</td>
<td>1st</td>
</tr>
<tr>
<td>Test 2</td>
<td>2nd</td>
<td>3rd&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1st&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Knowledge Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Create Style</td>
<td>2nd</td>
<td>3rd</td>
<td>1st</td>
</tr>
<tr>
<td>Modify Style</td>
<td>2nd</td>
<td>3rd</td>
<td>1st</td>
</tr>
<tr>
<td>Apply Style</td>
<td>2nd</td>
<td>3rd</td>
<td>1st</td>
</tr>
</tbody>
</table>

*Note.* Significant differences indicated by superscript group letters.

performance of the Coach group noticeably improved as they moved on to the second and third tasks. This difference was most likely due to the fact that they had received more frequent feedback and opportunities to practice the individual procedures, whereas the Critic group was only presented with the feedback once during each training task. In addition, the Coach group completed the immediate test and retention test more quickly than the other two groups, although the differences were not significant. In fact, their performance time improved between tests, whereas the other two groups performed more slowly on the retention test. The
Coach group’s efficiency scores also improved between tests, with their mean efficiency score only slightly lower than the Tutorial group’s mean scores. The Critic group’s efficiency score was much lower than the Coach group’s and did not change between tests. Finally, even though the Tutorial group performed slightly better, a higher proportion of the participants in the Coach group were able to correctly perform the three-part knowledge test than the Critic group.

The rating scores for the various questions asked on the questionnaires at the end of each session followed a similar pattern. The Tutorial group tended to have the highest mean ratings, but of more interest is the comparison of the Coach and Critic groups since both groups of participants had been exposed to the alternative training method of an intelligent help system. Overall, the Coach group found their training method to be more helpful, they liked their method more, they rated their confidence in learning styles higher, their confidence in actually using styles was higher, and they rated styles as more valuable than the Critic group. During the final session, more participants in the Coach group than in the Critic group reported actually using styles for their own word processing tasks. Finally, a slightly higher percentage of participants in the Critic group than in the Coach group reported that they would have preferred a different training method.

It is important to point out that the Tutorial training method used in this study represented a “best case scenario” since all of the participants were explicitly trained in the most efficient methods required for the tasks that they were later asked to complete. In a real world scenario, users would not have access to a tutorial such as this that is tailored to their needs and gives them the exact training they need. Users would have to stop their current task, actively seek out tutorial training, and try to find a lesson that might address the problems that they are having. This
assumes that the users even know that they need help in the first place. Also, in
today’s busy work environments, users rarely have the time to stop working on
their daily tasks and try to train themselves with tutorials (Newman & Riner, 1997).
Tutorials simply cannot provide efficient, continuous learning on demand within a
user’s task context. The Coach and Critic training methods represent an attempt to
use a proactive intelligent help system to provide the participants with the neces-
sary training within the context of their “work” tasks. However, unlike the Tutorial
method used in this study, the participants were not required to learn how to use
styles. They were free to choose to use the advice and learn more about using styles
or ignore it and perform the tasks using their own methods. It was hoped that such
training would be successful in overcoming the production paradox, thus adding
more efficient methods to each participant’s command repertoire. The Coach and
Critic methods were compared to determine if users require immediate, context-
specific feedback and training on individual methods in order to acquire new soft-
ware skills, or if they can learn just as effectively from delayed, more general, task-
specific feedback and training on combined methods.

The performance of the Coach group is inconsistent with the concept that
immediate and frequent feedback may interfere with long-term retention of skills
(Gaynor, 1981; Schmidt & Bjork, 1992) because it disrupts information-processing
activities and the coherence of the task (Carroll & Kay, 1985; Mobus, Pitschke, &
Schroder, 1992). However, it is consistent with the belief that immediate, context-
specific feedback is required for effective learning and skill acquisition (Anderson,
1987; Ericsson et al., 1993). Thus, the disruption of the user’s task by the Coach
was not as unwelcome as might be expected. Any negative effect of the disruption
was outweighed by the advantages of both the immediacy and context-specificity
of the feedback. The Critic group received feedback within the context of the whole task, and as such, should have been better able to apply the appropriate procedures to the test task when they were required, but instead they actually performed more poorly than the Coach group. There are several possible reasons why the Coach method resulted in better performance than the Critic method. The Coach group received more frequent and distributed feedback than the Critic group did, which may have prevented superficial learning and led to better performance. However, the Tutorial group did not receive frequent, distributed feedback, yet they also performed well. This is most likely due to the fact that each participant had to go through the styles training (i.e., they had no choice) and they learned exactly when, where, and how to use styles. Similarly, the Coach group received context-specific feedback relevant to particular sub-tasks (e.g., apply a style). This helped the participants learn when and where to use styles, whereas the participants in the Critic condition had to learn on their own how to map the individual procedures of using styles to the parts of the overall word processing task. They did not have the benefit of the Coach method’s timing of feedback to match particular sub-tasks.

Increased frequency of feedback is a “double-edged sword” though. On one hand, the frequent feedback encourages and reminds users who want to learn about new methods. In fact, some of the participants’ comments in Appendix N can attest to this fact. They liked that it “bothered” them until they learned about the more efficient methods. They stated that it encouraged them to learn new techniques that they never would have discovered on their own. However, on the other hand, the frequent feedback annoyed users who felt that it interrupted their work and repeatedly bothered them with its persistent messages. In fact, one participant became visibly angry and appeared as if he might strike the computer. This issue is some-
thing that should be taken into consideration when designing a Coach help system. The help system used for this study was deliberately designed to be very obtrusive during the training session, since this was the only opportunity to attract the attention of the participants and provide them with the opportunity to learn how to use styles. Very few participants in the Critic group were annoyed by the help window appearing only once during each task, however, as the performance results demonstrate, feedback that is this infrequent will probably not be effective.

The general question that this study addressed was: How can software users of an intermediate skill level be provided with timely and effective training in more efficient methods? More specifically, can an intelligent help system provide an effective learning experience? How should this type of help system be designed in order to be most effective? The results of this study suggest that a Coach help system design is, in general, as effective as a “best case scenario” tutorial. Providing frequent, context-specific feedback was more effective than the delayed, more general task-specific feedback for learning the new software methods (i.e., creating and using styles to format a document). The results also demonstrate that disrupting the participants’ task flow was not detrimental to their performance. Interrupting users when they are performing inefficiently provides them with a timely opportunity to learn new, more efficient methods that they can add to their skill set, which ultimately leads to more efficient task completion. A Critic style of help system may be effective for users who are highly motivated to improve their skills and efficiency. But, they would have to be willing to take the time to review the Critic’s feedback and recommendations after they had already completed a task and apply what they learned to future tasks. Few users seem willing to do this in real-world contexts. However, it should be noted that many of the users in this study who were provide
with the Critic training method did learn how to use styles. They would not have
discovered styles without a proactive help system that introduced the new methods
to them. In addition, it should also be noted that Critic systems are technically easi-
er to implement since they monitor the user’s actions and typically only present
help at the point of task completion, which allows the system to more fully develop
the user and task model before providing feedback. Coach systems run the risk of
misinterpreting the user’s actions since they attempt to provide continual feedback
whenever errors or inefficient actions are detected.

Technical considerations aside, this study suggests that Coach systems are
more effective than Critics. The proactive help provided by a Coach is necessary
because users often don’t even know that they need help (Breuker, 1988; Chin,
1991; Duffy et al., 1992; Fischer, 1993; Paxton & Turner, 1984). It is highly
unlikely that the participants in this study would have discovered how to use styles
to accomplish their tasks more efficiently without some type of intervention.
McGraw (1997) would suggest that even this amount of help would not be suffi-
cient. She argues that designing an interface for a system that satisfies the individual
users should not be the goal, instead interfaces should be designed to support
performance in order to satisfy the user’s organization. The users should be able to
perform their jobs “better, faster, or with less training” (p. 19). McGraw describes a
performance support system that provides multiple types of integrated tools to sup-
port the users and improve their performance (e.g., on-line help system, coach, wizard, tutor). Building a system such as this that cleanly and effectively integrates all
of necessary types of assistance would certainly not be an easy task. For example,
Selker (1997) noted that he thought “making an adaptive help system could be
done in six months - it took me 14 years to put COACH into an actual product” (p.
89). However, Mallen (1996) described a methodology for designing an application such that it is completely integrated with its help system from the start. The actual "act of building the application produces a system that a computer based instructor can immediately use" (p. 349). This approach is promising. Since many help systems are added to software as an afterthought, they are not as effectively integrated as they could be and designing and developing the help system takes additional time and money that developers are reluctant to invest. A methodology that combines the two processes would benefit both the developer and the end users.

Based on the findings of this study and previous research, some guidelines for designing a proactive intelligent help system can be suggested. The first recommendation is to provide frequent feedback and learning opportunities within the context of individual sub-tasks. Even though the Critic approach is easier to implement, its method of providing feedback at the end of task completion introduces two problems. First, the user is presented with a large amount of information all at once, which could be somewhat overwhelming. The burden is completely on the user to find a way to "digest" this information and break it down into more manageable chunks. The second problem, is that it does not provide an opportunity to learn a method in the context of the smaller sub-tasks within the larger overall task. This allows the users to learn exactly when and where that method is applicable in the context of the user's current task, rather than trying to map the methods back to the parts of the task after the task has already been completed. Another recommendation is that notification of the feedback from the Coach should be made salient, but not annoyingly obtrusive. If the users regard interaction with the help system to be unpleasant, they will be less likely to use it or will even find a way to turn it off. However, if they find the experience enjoyable, they will likely be more motivated
to use the system and learn more about the software (Atlas, Cornett, Lane, & Napier, 1997; Ujita, Yokota, Tanikawa, & Mutoh, 1996). A well-designed Coach help system, which follows these guidelines, would probably be a very effective method for encouraging users to learn more about the software tools that they use and helping them learn how to perform their tasks more efficiently.

Intelligent help systems are simply one part of the overall approach to assisting software users in their learning endeavors. The ideal training solution is to initially provide users with formal and intensive training to introduce new software or advanced features of software that they already use. They should be taught the most efficient methods that can be applied to accomplish their daily work tasks. However, users cannot be introduced to every single feature of even the average software application and they cannot be expected to remember everything that they have learned in a training session. A Coach help system could then be effectively implemented to facilitate continuous learning activities by reminding users to apply efficient methods that they have already learned, reminding them of methods that they have forgotten, and introducing them to more efficient methods that they may not even know exist. The Coach should be linked to three different types of assistance to provide users with exactly the kind of help that they need: A quick reference help system that reminds users of a method and provides a brief description, an on-line tutorial that provides them with complete training in the use of the method, or a Guide or Wizard system that could “walk” the users through each step of using the method. An integrated training solution such as this is a very effective approach to increasing users' efficient use and knowledge of the software that they utilize every day, enabling them to accomplish their tasks more quickly and effectively. It also ensures that initial investments in software training are not lost and
supports the continual learning activities that are a necessary part of today’s workplace.
REFERENCES


Breuker, J. (1992). Generality watching: ITS caught between science and engineer-


Caldwell, B. (1994, August 8). The five percent solution: Companies can expect to pay big bucks to train staff on new technology. *InformationWeek*, 59.


Sutcliffe, A. G., & Old, A. C. (1987). Do users know they have user models? Some experiences in the practice of user modelling. In H. Bullinger & B. Shackel (Eds.), *Human-computer interaction - Interact '87* (pp. 35-41), New York: Elsevier Science Publishers B.V.


nois.


APPENDIX A
Tutorial

Tutorial Instructions

You will now complete a tutorial that will teach you how to create and use Styles while formatting a document. The tutorial has three sections:

1. How to create a style.
2. How to apply a style.
3. How to modify a style.

Work through the entire tutorial and let me know when you are finished by calling me on the intercom.
**Styles Tutorial**

A style is a collection of several paragraph formats, such as alignment, spacing, and fonts, that have been grouped together and given a name. This lets you apply all of the formats at one time. This style name can then be easily applied to all the paragraphs that you want to have the same appearance. Making changes to a style changes the formats quickly and easily for all paragraphs that have that style in a few easy steps. You can modify Word's standard styles or create and modify your own styles using the ruler or the Style command under the Format menu.

You will be using a document on the desktop named “Tutorial document” for the following exercises. The document should already be open, but if it is not, go ahead and double-click the icon to open it now.

1. **Creating a new style using the ruler**

This procedure creates a style based on the formats of a selected paragraph. This is the easiest way to create a style.

   1. Modify the formats of the first paragraph, which contains the words “Beginning Level”, so the text is bold and underlined.

      Note: You must use the ruler to indent and align text, not the tab key or spacebar, in order for those paragraph formats to be included in the style.

   2. Select one or more words in this paragraph. For example, the first paragraph is selected in the figure below so that the bold and underline formats are included in the style.
3. Move the mouse pointer to the Styles box and click the mouse button to highlight the current style name which is probably "Normal".

Note: The Styles box at the left end of the ruler displays the name of the current style of the paragraph. The "Normal" style is the default style that Word provides and you have probably used it most often. For example, the style is "Normal" in the figure below.

4. Type a new style name called "Heading" in the box to add it to the list of styles. Note: Don’t worry, this will not replace the style that was already listed in the box (usually the "Normal" style). More than one word can be used for style names.
5. Press the Return key on the keyboard and the following message dialog box will appear.

![Image of message dialog box]

6. Click the Define button in the message dialog box to define your new style based on the selected paragraph.

7. Now you can apply this style to any other paragraphs that you want formatted to have the same appearance.

   User tip: You can add your style to Word's default style sheet and make it available in all new documents that you create.

   1. Choose Style from the Format menu.

   2. Select your new style.

   3. Click the Use As Default button and the following message dialog box will appear:
4. Click the Yes button in the message dialog box.
5. Click the OK button in the Styles dialog box.
2. **Applying a style to a paragraph using the ruler or keyboard**

   This procedure applies all of a style's formats to a paragraph in a few easy steps. Applying a style to a paragraph gives it the same appearance as all of the other paragraphs that share that same style.

   1. To apply a style to a single paragraph, position the insertion point anywhere in the paragraph. If you want to apply a style to more than one paragraph, you will need to select each of the paragraphs that you want to change. Now, position the insertion point in the 4th paragraph, which contains the words “Intermediate Level”.

   2. Click the down arrow to the right of the Styles box at the left end of the ruler and select the style named “Heading” from the list that pops up. For example, the “Normal” style is selected in the list in the figure below. The paragraph will now be formatted with the formats that belong to the “Heading” style.

![Styles Dialog Box](image)

   User tip: For a keyboard shortcut, you can press Command+Shift+S to move to the style name area in the lower-left corner of the window. Then, type the style name or its abbreviation and press the Return key on the keyboard. For example, in the figures below the style name “Heading” was typed in over the word “Style”.
3. Repeat steps 1 and 2 to apply the "Heading" style to the 7th paragraph, which contains the words "Advanced Level".
3. **Redefining an existing style using the ruler**

When you redefine a style, Word automatically makes the changes to all paragraphs formatted with that style. Redefining a style also affects any other styles based on that style.

1. Select the first paragraph, “Beginning Level”, which has the style you want to redefine.

2. Modify the formats of the paragraph so that the text is italic, and not bold or underlined. Leave the selected paragraph highlighted.

3. Click the Styles box on the left end of the ruler to highlight the current style name, which should be “Heading”.

4. Press the Return key on the keyboard and the following message dialog box will appear.
5. Select the button next to “Redefine the style based on selection?” in the message dialog box to redefine the style based on the new formats of the selected paragraph.

Note: To cancel the changes you have made to the paragraph, select the button next to “Reapply the style to the selection?”.

6. Click the OK button in the message dialog box. Now, the other two paragraphs that share that same style will have the same appearance as the first paragraph.
APPENDIX B
Warmup Task

Building a World Wide Web Page on the IBM PC: Evening series
18a Basics Feb 7 Wed 7:00 pm - 9:00 pm

Building a World Wide Web Page on the IBM PC: Daytime series
19a Basics Feb 28 Wed 9:00 am - 12:00 noon
Instructions for Coach Condition

You will now be asked to complete 3 word processing tasks using Microsoft Word. Keep in mind that the tasks will ask you to create a first draft of a document. You will probably be asked to make changes to that document later, such as different fonts or formats, different spacing and alignment, etc. You can complete these tasks using the methods with which you are already familiar, but there are probably better methods that would help you perform these tasks more efficiently and often more quickly. We want you to be the most efficient that you can be while working through the tasks, even though the most efficient methods may not always seem to be the quickest. There may be “quick and dirty” ways to accomplish the tasks, but a document created in that way is often more time-consuming and difficult to revise and edit later. The most efficient methods make later revision much easier.

Remember, we want you to be the most efficient that you can be while working through the tasks, but don’t spend too much time on each task. Please feel free to use Word’s online help or the User manual at any time. In addition, we have added an experimental automated help system to Microsoft Word that will monitor your actions. Any time that it detects that there is a more efficient method of accomplishing the task on which you are currently working, it will open a window each time telling you about the method. It will also refer you to a page number in the User manual where you can learn how to use the method. Please feel free to use the help that this system provides. Sometimes the system is slow, so please be patient.

If you have any problems, call me on the intercom. I cannot offer you help in
completing the tasks, but I can help you if you accidentally close a window, the computer freezes or crashes, etc. Please call me on the intercom when you have completed a task and I will come in and set up the next task for you.

Do you have any questions?
Instructions for Critic Condition

You will now be asked to complete 3 word processing tasks using Microsoft Word. Keep in mind that the tasks will ask you to create a first draft of a document. You will probably be asked to make changes to that document later, such as different fonts or formats, different spacing and alignment, etc. You can complete these tasks using the methods with which you are already familiar, but there are probably better methods that would help you perform these tasks more efficiently and often more quickly. We want you to be the most efficient that you can be while working through the tasks, even though the most efficient methods may not always seem to be the quickest. There may be "quick and dirty" ways to accomplish the tasks, but a document created in that way is often more time-consuming and difficult to revise and edit later. The most efficient methods make later revision much easier.

Remember, we want you to be the most efficient that you can be while working through the tasks, but don’t spend too much time on each task. Please feel free to use Word’s online help or the User manual at any time. In addition, we have added an experimental automated help system to Microsoft Word that will monitor your actions. If it detects that there are more efficient methods of accomplishing the task on which you are working, it will open a window once telling you about all of the methods. It will also refer you to a page number in the User manual where you can learn how to use the methods. It will only provide you with this help information once. Please feel free to use the help that this system provides. Sometimes the system is slow, so please be patient.

If you have any problems, call me on the intercom. I cannot offer you help in
completing the tasks, but I can help you if you accidentally close a window, the computer freezes or crashes, etc. Please call me on the intercom when you have completed a task and I will come in and set up the next task for you.

Do you have any questions?
Instructions for Tutorial Condition

You will now be asked to complete 3 word processing tasks using Microsoft Word. Keep in mind that the tasks will ask you to create a first draft of a document. You will probably be asked to make changes to that document later, such as different fonts or formats, different spacing and alignment, etc. You can complete these tasks using the methods with which you are already familiar, but there are probably better methods that would help you perform these tasks more efficiently and often more quickly. We want you to be the most efficient that you can be while working through the tasks, even though the most efficient methods may not always seem to be the quickest. There may be “quick and dirty” ways to accomplish the tasks, but a document created in that way is often more time-consuming and difficult to revise and edit later. The most efficient methods make later revision much easier.

Remember, we want you to be the most efficient that you can be while working through the tasks, but don’t spend too much time on each task. Please feel free to use Word’s online help or the User manual at any time.

If you have any problems, call me on the intercom. I cannot offer you help in completing the tasks, but I can help you if you accidentally close a window, the computer freezes or crashes, etc. Please call me on the intercom when you have completed a task and I will come in and set up the next task for you.

Do you have any questions?
APPENDIX D
Training Document for Task One

Psychology 101

Intro to Psychology
Mon (9-10am)  Wed (10-11am)  Fri (1-2pm)

English 203

Composition
Tue (1-2pm)  Thu (3-4pm)

History 110

Western Civilization
Mon (8-10am)  Tue (1-3pm)  Wed (1-3pm)
Training Document for Task Two

**Psychology 101**

*Intro to Psychology*

Mon (9-10am)  Wed (10-11am)  Fri (1-2pm)

**English 203**

*Composition*

Tue (1-2pm)  Thu (3-4pm)

**History 110**

*Western Civilization*

Mon (8-10am)  Tue (1-3pm)  Wed (1-3pm)
Training Document for Task Three

RESUME

Objectives

• A position utilizing my market research skills and knowledge
  - An opportunity to receive real-world training

Education

• University of Houston - Houston, TX
  - August 1994 to Present

Honors

• Dean’s List
  - 3 semesters
APPENDIX E
Document for Test One

Introduction

Study 1

The study was conducted to compare different methods of list memorization. It also examined the effects of distraction tasks on memory scores. It was believed that a counting distraction task would reduce the number of digits recalled.

Method

Procedure

The subjects were given 2 minutes to memorize a list of fifty digits. They then performed a counting distraction task for thirty seconds. Finally, they attempted the recall task.

Conclusion

Study 1

The results for this study were not surprising. The distraction task did have a negative effect on the digit recall task. The subjects recalled only 10 percent of the digits.
Selling Your Home

Preparing the House

The house should be professionally cleaned before contacting your real estate agent. You should consider repainting the exterior of the house and any of the interior rooms that have paint which is faded, stained, or chipped.

Buying a New Home

Contacting the Agent

Call an agent to set up an appointment to inspect the new house. You should take the time to walk through the house with the agent and locate any trouble spots.

Obtaining a Loan

Qualifying for a Loan

You will need to arrange an appointment with a lending institution officer. Check over your credit report prior to the meeting for any mistakes.
APPENDIX F
Knowledge Test

1. Modify the first line of the Knowledge test document (which contains the words “Going to Court”) so that the text is:

   - Times 14, Bold, Italic, and Centered

   - Create a new style named “Title” for the first line (based on the formats of its text)

2. Now, modify the style for the second line (which contains the words “Choosing a Lawyer”) so that the text is:

   - Times 12, Underline, and Flush left (not centered)

3. Type the words “Legal Representation” in a new paragraph at the end of the document. Apply the style that you just modified for step #2 above to this new paragraph.
Background Questionnaire

Age ___________ Gender (circle one)  M       F

Year in school ______________

Major ______________________

Which version of Microsoft Word do you normally use (4.0, 5.1, 6.0, etc.)?
__________________________

How often do you use Microsoft Word? (circle one)

1. Every day          4. A few times a month
2. 2 to 4 times a week 5. Rarely
3. About once a week

Describe the word processing tasks that you typically perform using Microsoft Word:

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
How long have you been using Microsoft Word? ______ years ______ months

How did you learn how to use Microsoft Word?


Rank how useful you think each of the following methods of learning the software might be for you by placing a "1" in the space next to the method that you think is most useful, a "2" next to the second most useful method and so on. Also, if you have actually used that method in the past to learn Microsoft Word, then estimate how much time (in hours) you have spent using it.

<table>
<thead>
<tr>
<th>Method</th>
<th>How useful</th>
<th>Time spent (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard copy manuals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hard copy tutorials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>On-line tutorials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>On-line help</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor-led training</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Videotape training</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learned from others</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learned as you worked</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(through trial and error)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX H
Questionnaire for Session One

The following questions ask you to respond on a scale of 1 to 10. Please circle one of the numbers for each question.

1. How often did you use the on-line help while working through the word processing tasks?

1  2  3  4  5  6  7  8  9  10
Never                                           Always

2. How often did you use the User manual while working through the word processing tasks?

1  2  3  4  5  6  7  8  9  10
Never                                           Always

3. How helpful was the tutorial/intelligent help system?

1  2  3  4  5  6  7  8  9  10
Not at all                                       Very helpful
4. How much did you like the tutorial/intelligent help system?

1  2  3  4  5  6  7  8  9  10
Not at all

5. What did you like or dislike about the tutorial/intelligent help system?

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

6. How much did you like the Styles help information in the User Manual?

1  2  3  4  5  6  7  8  9  10
Not at all

7. What did you like or dislike about the information in the User Manual?

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
8. How confident are you that you have learned how to use Styles to format documents?

1 2 3 4 5 6 7 8 9 10
Not at all Very confident

9. How confident are you that you will use Styles to format documents for your personal word processing tasks (e.g., papers, reports, etc.)?

1 2 3 4 5 6 7 8 9 10
Not at all Very confident

10. How valuable do you think Styles are?

1 2 3 4 5 6 7 8 9 10
Not at all Very valuable
APPENDIX I
Questionnaire for Session Two

The following questions ask you to respond on a scale of 1 to 10. Please circle one of the numbers for each question.

1. How often did you use the on-line help while working through the word processing tasks?

   1  2  3  4  5  6  7  8  9  10
   Never                      Always

2. How often did you use the User manual while working through the word processing tasks?

   1  2  3  4  5  6  7  8  9  10
   Never                      Always

3. How helpful was the tutorial/intelligent help system?

   1  2  3  4  5  6  7  8  9  10
   Not at all                  Very helpful
4. How much did you like the tutorial/intelligent help system?

1  2  3  4  5  6  7  8  9  10
Not at all                     Very much

5. What did you like or dislike about the tutorial/intelligent help system?

__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________

6. How much did you like the Styles help information in the User Manual?

1  2  3  4  5  6  7  8  9  10
Not at all                     Very much

7. What did you like or dislike about the information in the User Manual?

__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
8. How confident are you that you have learned how to use Styles to format documents?

1 2 3 4 5 6 7 8 9 10
Not at all Very confident

9. How confident are you that you will use Styles to format documents for your personal word processing tasks (e.g., papers, reports, etc.)?

1 2 3 4 5 6 7 8 9 10
Not at all Very confident

10. How valuable do you think Styles are?

1 2 3 4 5 6 7 8 9 10
Not at all Very valuable

11. How often have you used Styles during your word processing in the past two weeks?

1 2 3 4 5 6 7 8 9 10
Not at all Always
12. How many opportunities do you think you have had where you could have used Styles during your word processing in the past two weeks?

1  2  3  4  5  6  7  8  9  10

None  Very many
Questionnaire for Session Three

1. How often have you been performing word processing tasks during the last 8 weeks (i.e., since the 2nd session of this experiment)?

   1  2  3  4  5  6  7  8  9  10
   Never  Several times a day

2. Have you been using Microsoft Word for your word processing? If not, what do you use?

   

3. How many papers have you written using Microsoft Word during the last 8 weeks?

   

4. Did you use Styles while writing any of those papers?  Yes  or  No

5. If “Yes”, what percentage of the total number of papers were formatted, at least partially, using Styles (e.g., 0%, 50%, 100%, etc.)?

   

6. If “Yes”, how did you use Styles for those papers (e.g., to create headings and subheadings, etc.)?

7. How confident are you that you have learned how to use Styles to format documents?

1 2 3 4 5 6 7 8 9 10
Not at all  Very confident

8. How confident are you that you will use Styles to format documents for your personal word processing tasks (e.g., papers, reports, etc.)?

1 2 3 4 5 6 7 8 9 10
Not at all  Very confident

9. How valuable do you think Styles are?

1 2 3 4 5 6 7 8 9 10
Not at all  Very valuable

10. How often have you used Styles during your word processing in the past two weeks?

1 2 3 4 5 6 7 8 9 10
Not at all  Always
11. How many opportunities do you think you have had where you could have used Styles during your word processing in the past two weeks?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>Very many</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

12. If you do not use Styles frequently to format your documents (or didn’t use them at all during the experiment), please explain why you think you do not use them more often.

13. What situations would persuade you to learn or encourage you to use Styles for your word processing?

14. Would you have preferred a different training method than the one you received during the first experimental session? Explain your answer.
APPENDIX K
User Manual

Styles

A style is a collection of several paragraph formats, such as alignment, spacing, and fonts, that have been grouped together and given a name. This lets you apply all of the formats at one time. This style name can then be easily applied to all the paragraphs that you want to have the same appearance. Making changes to a style changes the formats quickly and easily for all paragraphs that have that style in a few easy steps. You can modify Word's standard styles or create and modify your own styles using the ruler or the Style command under the Format menu.

- Creating a new style using the ruler

This procedure creates a style based on the formats of a selected paragraph. This is the easiest way to create a style.

1. Modify the formats of a paragraph until it has the appearance you desire.

2. Select one or more words in this paragraph. For example, the first line of the first paragraph is selected in the figure below so that the bold and underline formats are included in the style.

![Image](Beginning_Leve.png)
3. Move the mouse pointer to the Styles box and click the mouse button to highlight the current style name.

Note: The Styles box at the left end of the ruler displays the name of the current style of the paragraph. The "Normal" style is the default style that Word provides and you will probably use it most often. For example, the style is "Normal" in the figure below.

4. Type your new style name in the box to add it to the list of styles. Note: Don’t worry, this will not replace the style that was already listed in the box (usually the “Normal” style).
5. Press the Return key on the keyboard and the following message dialog box will appear.

![Define style "Heading" based on selection?](Define)

![Cancel](Cancel)

6. Click the Define button in the message dialog box to define your new style based on the selected paragraph.

7. Now you can apply this style to any other paragraphs that you want formatted to have the same appearance.

User tip: You can add your style to Word's default style sheet and make it available in all new documents that you create.

1. Choose Style from the Format menu.
2. Select your new style.
3. Click the Use As Default button and the following message dialog box will appear:

![OK to record style in default style sheet?](Yes)

![No](No)

4. Click the Yes button in the message dialog box.
• Applying a style to a paragraph using the ruler or keyboard

This procedure applies all of a style's formats to a paragraph in a few easy steps. Applying a style to a paragraph gives it the same appearance as all of the other paragraphs that share that same style.

1. To apply a style to a single paragraph, position the insertion point anywhere in the paragraph. If you want to apply a style to more than one paragraph, select all of the paragraphs that you want to change.

2. Click the down arrow to the right of the Styles box at the left end of the ruler and select the style you want from the list that pops up. For example, the "Normal" style is selected in the list in the figure below. The paragraph will now be formatted with that style's formats.
User tip: For a keyboard shortcut, you can press Command+Shift+S to move to the style name area in the lower-left corner of the window. Then, type the style name or its abbreviation and press the Return key on the keyboard. For example, in the figures below the style name "Heading" was typed in over the word "Style".

![Image of style name area]

![Image of style name area]
• Redefining an existing style using the ruler

When you redefine a style, Word automatically makes the changes to all paragraphs formatted with that style. Redefining a style also affects any other styles based on that style.

1. Select the paragraph that has the style you want to redefine.

2. Modify the formats of the paragraph until it has the appearance you desire and leave the selected paragraph highlighted.

3. Click the Styles box on the left end of the ruler to highlight the current style name.

4. Press the Return key on the keyboard and the following message dialog box will appear.
5. Select the button next to “Redefine the style based on selection?” in the message dialog box to redefine the style based on the new formats of the selected paragraph.

Note: To cancel the changes you have made to the paragraph, select the button next to “Reapply the style to the selection?”.

6. Click the OK button in the message dialog box.
APPENDIX L
### Procedures

**TABLE 15. Procedures for each group across all sessions**

<table>
<thead>
<tr>
<th>Session</th>
<th>Coach</th>
<th>Critic</th>
<th>Tutorial</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First Session</strong></td>
<td>• Background and experience questionnaire</td>
<td>• Background and experience questionnaire</td>
<td>• Background and experience questionnaire</td>
</tr>
<tr>
<td></td>
<td>• Warmup task</td>
<td>• Warmup task</td>
<td>• Warmup task</td>
</tr>
<tr>
<td></td>
<td>• Training tasks</td>
<td>• Training tasks</td>
<td>• Tutorial</td>
</tr>
<tr>
<td></td>
<td>• On-line help, manual, and Coach feedback</td>
<td>• On-line help, manual, and Critic feedback</td>
<td>• On-line help and manual</td>
</tr>
<tr>
<td></td>
<td>• Immediate test</td>
<td>• Immediate test</td>
<td>• Immediate test</td>
</tr>
<tr>
<td></td>
<td>• Rating questions</td>
<td>• Rating questions</td>
<td>• Rating questions</td>
</tr>
<tr>
<td><strong>Second Session</strong></td>
<td>• Warmup task</td>
<td>• Warmup task</td>
<td>• Warmup task</td>
</tr>
<tr>
<td>(2 weeks later)</td>
<td>• Retention test</td>
<td>• Retention test</td>
<td>• Retention test</td>
</tr>
<tr>
<td></td>
<td>• Wonderlic test</td>
<td>• Wonderlic test</td>
<td>• Wonderlic test</td>
</tr>
<tr>
<td></td>
<td>• Rating questions</td>
<td>• Rating questions</td>
<td>• Rating questions</td>
</tr>
<tr>
<td><strong>Third Session</strong></td>
<td>• Use questions</td>
<td>• Use questions</td>
<td>• Use questions</td>
</tr>
<tr>
<td>(8 weeks later)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX M
Boxplot Figures

Boxplots of Performance Time for All Training Tasks (1-3)

FIGURE 8. Boxplot of group performance times for Task 1
FIGURE 9. Boxplot of group performance times for Task 2
FIGURE 10. Boxplot of group performance times for Task 3
Boxplots of Efficiency Scores for All Training Tasks (1-3)

FIGURE 11. Boxplot of group efficiency scores for Task 1
FIGURE 12. Boxplot of group efficiency scores for Task 2
Boxplots of Performance Time for Both Tests (1 and 2)

FIGURE 14. Boxplot of group performance times for Test 1
FIGURE 15. Boxplot of group performance times for Test 2
Boxplots of Efficiency Scores for Both Tests (1 and 2)

FIGURE 16. Boxplot of efficiency scores for Test 1
FIGURE 17. Boxplot of efficiency scores for Test 2
APPENDIX N
What the participants reported liking about the Coach method:

I had never used the style feature before and never would unless someone mentioned it to me or an intelligent help system like this pointed it out. I like its detection of inefficiency.

I liked that it kept bothering me to use the “style format” until I finally did. That way, I was forced to use it.

It was an effective way to learn about something I would never even think of going to a manual to learn (since I didn’t know that function existed).

I liked that it kept coming up, which forced me to learn how to use styles. If it had not been so persistent, I probably would have ignored it.

After my method became tedious, the help system reminded me that there was a better was to go about things.

I liked the fact that it would come up each time that I made a mistake that it could catch.

It identified errors as I made them, cutting the time I would have to spend redoing the poorly formatted portion of my document. It showed me a new technique I’d never even considered and wouldn’t consider w/o the help.

I liked that it would interrupt me while I was typing to tell me about a more efficient way to format my paragraphs.

The good thing was that it did refer me to the manual when a faster way could help decrease the time it took for me to complete a task

It was easy to follow and gave very helpful tips that I will be able to use in the future.

The help system brought to my attention a more efficient way of solving the
problem which helped my efforts for the rest of the tasks presented.

It seems to make things more difficult at first, but ends up helping in the end. Advice given can be useful.

It helped in that it referred to the page number of the User Manual, and also in that it was correct about an easier way to perform the task. However, it would be easier if the information in the User Manual was in the on-line help system.

It knew exactly what I needed to do to make my job easier. However, it could be distracting when it just pops up on the screen.

A little slow, but effective.
What the participants reported disliking about the Coach method:

I didn’t like that it interrupted my work, but I did like that it suggested options for me to try that I would never have considered trying.

It would be good if it gave you a link to access the corresponding on-line help section instead of giving you the manual’s page number. Hard copy manuals are good when you are learning to use the program step by step, but in cases like this one in which you want to know something specific on-line help is better. The problem with traditional on-line helps is that sometimes you don’t know where to find what you are looking for, but in the case of the smart help this could be solved by adding the hot-link to the suggested section.

Made the task easier. But interrupted my concentration when it popped up.

Though I regret not having utilized the intelligent help system, I feel nonetheless that it is too persistent. If a user neglects to follow the intelligent help system after its second or third appearance it can be a nuisance.

It was very helpful except it could become annoying for experienced users if it popped up too often unnecessarily.

It was annoying when I was trying to make corrections my way.

It kept repeating itself. I understood it the first time.

It might have been better if the information had been given directly on the screen as an option. I could see where people would simply neglect to stop and use the manual.

I disliked the interruptions.

If I didn’t want the help, it still kept repeating the same suggestion.

The fact that it was activated so often bothered the hell out if me. If denied
once it should be turned off in the future for similar tasks.

It kept telling me I was making the same errors over and over again and since I could not figure out how to correct those errors, it just got very annoying.

It kept interrupting and slowing me down and when I finally did look up what it told me to look up, it made little sense.

Good suggestions but the speed of the system was too slow (it took a long time for the window to pop up and disappear) and there was not way to disable the suggestion from keep popping up.

I liked that it gave me good and applicable advice that was easy to understand and implement, but I disliked being interrupted in the middle of the task because I thought it would tell me something I didn't need to know.
What the participants reported liking about the Critic method:

I did not even know about the style option in MW and would probably never have known if the help system hadn't told me. It is fairly effective with things like this, but I would like to use it more to be sure.

I liked the help system because after the initial frustration period I was able to think about the benefits and felt that it would help me if I learned to use the Microsoft Word as efficiently as possible. Any type of system that will alert you when you are being inefficient is good to have even if you choose not to use the information every time.

I like this system because it tells you about options on the Mac that you would otherwise never now about unless someone else tells you about it. I would never have known about the styles option unless this help system told me.

I liked how it made me aware of certain features I otherwise would not have known about. However, I didn't like being nagged by the help (i.e., Keep on telling me the same thing).

It reminded me that I shouldn't be so stubborn by trying to figure things out on my own instead of utilizing the system and tools provided.

I liked very much the fact that it only came up once to alert the user about a specific function of MS Word and then it disappeared. It would have gotten annoying to have it keep coming up, even though I might not have chosen to follow some of its advice.

I think that the intelligent help system was very useful. It gives you the advice while you are working on the task and lets you know where to look for the info. that you need to do the task more efficiently.
I liked how it let me know where in the manual to look.

It showed me the “style” option which I did not know about before. I am not sure how useful overall that option is, but I guess it is a good thing to know about.

I liked the fact that it helped me be more efficient.

I liked when it gave page numbers to be more specific about where to look for tips to be more efficient.
What the participants reported disliking about the Critic method:

It looked good at first but I had a difficult time understanding the manual. It would perhaps be more efficient to have the help system actually HELP instead of refer you to a manual.

Honestly, I didn't pay much attention to what was actually said. It was helpful, however, to put the page number in the manual that the helpful information would be on in the intelligent help system box.

It did not give enough information as to exactly how I needed to be helped. It's not worth it to me just to be told that there is "something" I'm doing that is inefficient. If I'm doing something inefficient then just tell me what I'm doing wrong and walk me through it. If the help system is smart enough to pick out when I'm being inefficient then it should be able to tell me exactly how to correct it step by step without sending me to the manual. The help system needs to be more interactive. If you can put the information in a manual then you should be able to put even more information in the semi-intelligent help system.

I enjoy figuring things out for myself without windows popping up at random. I think in such a way that doing things in a more efficient manner takes me more time because I get used to formatting, etc. by hand instead of having it done automatically.

I liked the fact that it offered an alternative way to perform a task, but it was not very helpful to have to refer to the manual because the manual was not clear for me at first. It would be good if the help system could demonstrate the task the alternative way rather than the manual.

For such a short application, though, it really was more of a nuisance. Trying
to learn all aspects of the “Heading” application, for example was actually hindering performance.

It should have given the suggestion right on the screen rather than referring you to the book.

Eventually, it would be helpful to eliminate the hard copy of the manual; the help system would identify the problem and then provide the info. on-line.

In the long-run I know it will cut-down on time, but when you are limited in time in present it is difficult to stop and look up helpful hints that will allow you to decrease time spent formatting.

It really didn't bother me, except that it took a little while to close.
What the participants reported liking about the Tutorial:

It wasn't complicated. The information of style redefining was useful.

The tutorial explained every step very thoroughly, but sometimes I was confused about which part of the document I was supposed to be modifying. It also presented information about Microsoft Word that was new to me, so that was interesting.

Now that I applied what it showed me, I can say that it really worked. I like how it takes you step by step and shows you examples. I myself learn faster by seeing what I am doing than by just reading directions.

I liked the fact that it showed me how little I know about all the capabilities of Microsoft Word--which is perceived to be a very simple piece of software--I should be able to do more than I actually can.

It wasn't too long; didn't belabor a point. It was instructional and informative without being condescending.

It was good that it gave examples and showed everything that would come up on the screen.

I was able to work through some problems and learn by using the skills that I had learned and apply them to become familiar with them.

Definitely gave me practice in learning what I was doing

It was concise, easy to follow, hands on.

I liked the way it provided a lot of information relative to the task I was doing. It was clear, also.

The tutorial was fine. I liked the way it involved me. That is, I liked the way it had me perform the computer operations rather than having me just read some-
thing.

The tutorial showed me how little I know about the technical side and the short-cuts of WORD.

I liked that it told me exactly what to do. I liked that it told me what would happen (what boxes would pop up) so I knew that I was doing it right.
What the participants reported disliking about the Tutorial:

The tutorial was dull; however, it was also very helpful.

Some of the wording in the tutorial was difficult to decipher. I had to go back to the beginning once or twice to understand what was being said.

It was not easy if you had a problem to specifically look up its answer. I prefer on-line tutorials that work you through the areas in which you are having problems. It is easier to see the things on the computer than having to transfer them from text to the computer.

It was informative, but not really fun.

The page setup was pretty cramped. I did not think that the information was presented very clearly. A step by step example of creating one setting from start to finish without the extra comments might have been more helpful. Perhaps put the extra comments at the end of the orig. information.

The tutorial was a bit ambiguous in some places, but after figuring out what it was saying, I found the tutorial to be very helpful.

It seemed too remedial at times, but I guess that is unavoidable when you aren't sure of the background of your audience. Otherwise, it was very helpful!
What the Coach participants reported liking about the User manual:

I liked the diagrams that it had. Without the pictures of where to click, I don't know if I would have been able to figure out how to change styles.

It was easily accessed - dividers, labels etc. It was also clear, concise and helpful as well as illustrative.

Very easy step by step instruction. Good illustrations (what comes up on the screen).

I liked the pictures so that I did not have to read the whole thing to understand what to do.

I liked that it was illustrated--I didn't have to read all the words in the book. The dialogue boxes that showed what to expect assured me that I was on the right track or let me know when I wasn't. The index tabs were also nice.

I liked that it was simple to read (a few short sentences and an example).

I liked that it was easy to follow and very helpful.

What the Coach participants reported disliking about the User manual:

Some of the phrasing in the manual were not all that clear.

Some of the wording of the manual didn't "flow" easily. Computer manuals have a reputation for being "difficult" anyway and perhaps this biased my opinion.

I thought some of the wording was a little misleading and not to the point. Although the User Manual was better than others I have seen.

It was clear and straightforward. However, sometimes when you are typing a document it seems bothersome to have to consult a manual. You'd rather do it the
way you already know that might be a little bit slower.

It was just hard to understand and took a while to read. If it had been a quick step by step explanation to how to use it, I may have paid more attention to it.
What the Critic participants reported liking about the User manual:

I liked it because it was presented in a concise, easy-to-read format. I could follow the directions easily.

It told me what I needed to know to get the job done. I guess you really can't learn how to do something entirely unless you actually go through the procedure yourself.

The user manual was fairly clear in what it wanted you to do. This section explained in detail what it wanted from you and was not very ambiguous as these things tend to be.

The info. really made the task go a lot faster and it was info. that I had never encountered before, so it was very useful. It also gives you step by step instruction as to how to work the styles section.

I liked how the user manual was very concise and had a lot of graphics to facilitate understanding.

Like the fact that it has dialog boxes identical to what you see on the monitor.

The pictures they gave as examples were helpful.

It was clear and straight forward, I liked the manual because you could skip some steps if you had already figured them out, rather than having to scroll through or wait.

I like the manual for my questions about how to do things (i.e., how to set-up tabs).
What the Critic participants reported disliking about the User manual:

When I tried to use the info, I was confused and set back in my task. The style is easy to read and follow, but I could not figure out how to make it help my specific 'problem' (which wasn't really a problem in my view until the helper came on the screen.)

I guess I did not take the time to read every sentence, but what I did read could have been a little more clear. I tried to implement the Styles function 2/3 times and did not figure it out until I learned from my own trial and error. It could be that I didn't take the time to read through it thoroughly, but the time it would take me to read every sentence would exceed the way I was already doing it which makes it more inefficient considering the short length of the tasks.

It was a typical user manual, and I will admit that I am afraid of user manuals because of my experiences with them in the past. Sometimes it's better to learn on your own, except in cases where the manual and other tools are actually helpful.

The information was way too ambiguous. I tried to do what it said and I ended up doing more work than I would have if I had just done it myself. I am not as interested in trying to make rewrites easier as I am interested in getting the assignment at hand finished. If the manual can help me accomplish that then fine, but it only seemed to confuse me. Maybe because I was only interested in finishing this experiment and getting out of here. Possibly if this were my own computer I might really try and get more out of the manual. As it stands now it took too much time to really get involved with trying to figure things out. I probably will be using the manual more often in the future.

Similar to the help system though, sometimes if you have time limitations, it
is difficult to stop and look-up ways to decrease on formatting time. In the long-run knowing the information is extremely beneficial
What the Tutorial participants reported liking about the User manual:

I like how it gave very detailed instructions and what we should anticipate next.

It was very clear in terms of directions and it broke the steps down really well. Also, the pictures matched what was on the screen exactly.

The information was concise and straightforward, but people unfamiliar with computer terminology might have trouble understanding it.

What the Tutorial participants reported disliking about the User manual:

I didn't use the user manual so I have no opinion.

I did not use the user manual at all.

I didn't open the manual...I'm a male and don't typically ask for directions.

I did not use it --- I'm too stubborn to look something up in a book...

or too lazy.
APPENDIX O
**Correlation Matrix**

**TABLE 16. Intercorrelations among major study variables**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Wonderlic score</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. How often use MS Word</td>
<td>0.0991</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. How long used MS Word</td>
<td>0.1540</td>
<td>0.0459</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. How helpful was training</td>
<td>0.0516</td>
<td>-0.1930</td>
<td>-0.1140</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>5. Liked training</td>
<td>-0.0118</td>
<td>-0.1402</td>
<td>-0.0391</td>
<td>0.8829*</td>
<td>-</td>
</tr>
<tr>
<td>6. Liked manual</td>
<td>-0.0001</td>
<td>-0.1224</td>
<td>-0.0213</td>
<td>0.5519*</td>
<td>0.7107*</td>
</tr>
<tr>
<td>7. Final confidence had learned styles</td>
<td>0.1327</td>
<td>0.0095</td>
<td>0.2085</td>
<td>0.5325*</td>
<td>0.5289*</td>
</tr>
<tr>
<td>8. Final confidence will use styles</td>
<td>-0.1764</td>
<td>-0.0801</td>
<td>-0.0345</td>
<td>0.3794*</td>
<td>0.3447*</td>
</tr>
<tr>
<td>9. Final styles value rating</td>
<td>-0.3072*</td>
<td>0.0475</td>
<td>-0.2002</td>
<td>0.1184</td>
<td>0.0185</td>
</tr>
<tr>
<td>10. How often used styles past 2 weeks (3rd session)</td>
<td>-0.1252</td>
<td>0.0997</td>
<td>-0.0148</td>
<td>0.0855</td>
<td>0.0416</td>
</tr>
<tr>
<td>11. Opportunities in past 2 weeks (3rd)</td>
<td>-0.3094*</td>
<td>-0.0905</td>
<td>-0.2035</td>
<td>0.0356</td>
<td>0.0418</td>
</tr>
<tr>
<td>12. Number of papers written in past 8 weeks</td>
<td>-0.2071</td>
<td>-0.4073*</td>
<td>0.1380</td>
<td>0.2597*</td>
<td>0.2055</td>
</tr>
<tr>
<td>13. % papers formatted using styles</td>
<td>-0.0993</td>
<td>-0.0333</td>
<td>-0.0936</td>
<td>0.1210</td>
<td>0.0540</td>
</tr>
<tr>
<td>14. Number of times intelligent help was presented</td>
<td>-0.0547</td>
<td>-0.0717</td>
<td>-0.0063</td>
<td>-0.4396*</td>
<td>-0.4885*</td>
</tr>
<tr>
<td>15. Total time spent using help</td>
<td>-0.1917</td>
<td>-0.1067</td>
<td>-0.0149</td>
<td>0.0115</td>
<td>0.1311</td>
</tr>
<tr>
<td>16. Knowledge test</td>
<td>0.1488</td>
<td>0.1816</td>
<td>-0.0046</td>
<td>0.0552</td>
<td>0.0386</td>
</tr>
</tbody>
</table>

*Note. * indicates significance at p < .05
# TABLE 16. Intercorrelations among major study variables

<table>
<thead>
<tr>
<th>6. Liked manual</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Final confidence had learned styles</td>
<td>0.4796*</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Final confidence will use styles</td>
<td>0.4407*</td>
<td>0.6138*</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>9. Final styles value rating</td>
<td>0.0340</td>
<td>0.1156</td>
<td>0.4812*</td>
<td>-</td>
</tr>
<tr>
<td>10. How often used styles past 2 weeks (3rd session)</td>
<td>0.1440</td>
<td>0.3962*</td>
<td>0.6754*</td>
<td>0.3780*</td>
</tr>
<tr>
<td>11. Opportunities in past 2 weeks (3rd)</td>
<td>0.0349</td>
<td>0.0639</td>
<td>0.4256*</td>
<td>0.4121*</td>
</tr>
<tr>
<td>12. Number of papers written in past 8 weeks</td>
<td>0.2183</td>
<td>0.0694</td>
<td>0.2357</td>
<td>0.0634</td>
</tr>
<tr>
<td>13. % papers formatted using styles</td>
<td>0.1203</td>
<td>0.3269*</td>
<td>0.6571*</td>
<td>0.4070*</td>
</tr>
<tr>
<td>14. Number of times intelligent help was presented</td>
<td>-0.1432</td>
<td>-0.3397*</td>
<td>-0.1906</td>
<td>0.0741</td>
</tr>
<tr>
<td>15. Total time spent using help</td>
<td>0.4308*</td>
<td>-0.0252</td>
<td>-0.0799</td>
<td>-0.0645</td>
</tr>
<tr>
<td>16. Knowledge test</td>
<td>-0.0523</td>
<td>0.0662</td>
<td>-0.0211</td>
<td>-0.1394</td>
</tr>
</tbody>
</table>

*Note. * indicates significance at p < .05
<table>
<thead>
<tr>
<th></th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>11. Opportunities in</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>past 2 weeks (3rd)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Number of papers written</td>
<td>0.1580</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in past 8 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. % papers formatted using</td>
<td>0.6859*</td>
<td>0.0142</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>styles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Number of times intelligent</td>
<td>0.1395</td>
<td>0.0818</td>
<td>0.0967</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>help was presented</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Total time spent using</td>
<td>-0.0974</td>
<td>0.1686</td>
<td>-0.2252</td>
<td>0.1757</td>
<td></td>
<td></td>
</tr>
<tr>
<td>help</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Knowledge test</td>
<td>0.0287</td>
<td>-0.0540</td>
<td>0.0755</td>
<td>-0.2185</td>
<td>-0.2751*</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* *indicates significance at p < .05