MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAUS OF STANDARDS 1963-A
IMPROVING THE COMPREHENSIBILITY
OF A SIMULATED TECHNICAL MANUAL

David E. Kieras

University of Michigan

Technical Report No. 20 (TR-85/ONR-20)

June 20, 1985

This research was supported by the Personnel and Training Research Programs, Office of Naval Research, under Contract Number N00014-84-K-0729, Contract Authority Identification Number NR 667-513. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Approved for Public Release; Distribution Unlimited.
Improving the Comprehensibility of a Simulated Technical Manual

David E. Kieras

College of Engineering
University of Michigan
Ann Arbor, MI 48109

Office of Naval Research (Code 458)
Arlington, VA 22217

Reading comprehension, readability, documentation

It is clear that careful rewriting can improve the comprehensibility and usability of a technical document. But exactly how the document should be modified to accomplish this improvement has been unclear. In two experiments, a simulated technical manual for a simple piece of equipment was evaluated for comprehensibility using well-defined rules, and rewritten to eliminate the specific comprehension problems detected by the
rules. Unlike conventional readability formulas or writer's guidelines, these rules were based on results and theory from research on comprehension processes, and thus concern what aspects of text are actually difficult for the reader to process. Substantial improvements in performance were observed in tasks in which the subject had to learn how to operate the equipment using the manual. These results suggest that a computer-based system that could apply the comprehensibility rules would be useful in improving the quality of technical documents.
Improving the Comprehensibility of a Simulated Technical Manual

David E. Kieras

Abstract

It is clear that careful rewriting can improve the comprehensibility and usability of a technical document. But exactly how the document should be modified to accomplish this improvement has been unclear. In two experiments, a simulated technical manual for a simple piece of equipment was evaluated for comprehensibility using well-defined rules, and rewritten to eliminate the specific comprehension problems detected by the rules. Unlike conventional readability formulas or writer's guidelines, these rules were based on results and theory from research on comprehension processes, and thus concern what aspects of text are actually difficult for the reader to process. Substantial improvements in performance were observed in tasks in which the subject had to learn how to operate the equipment using the manual. These results suggest that a computer-based system that could apply the comprehensibility rules would be useful in improving the quality of technical documents.
Improving the Comprehensibility of a Simulated Technical Manual

David E. Kieras

The work presented in this report is part of a project to develop an advanced computerized system that will aid the writer of technical documents in preparing more comprehensible prose. The basic idea, described in Kieras (1985), is to use natural language processing techniques from artificial intelligence and results from psycholinguistics and cognitive psychology in the design of a system that would point out to writers where and why their prose is difficult for a reader to comprehend. Similar proposals at the theoretical level have been made by Miller and Kintsch (1980), and Kintsch and Vipond (1979), but in order to actually perform the complex analyses involved, a computer is necessary (see Kieras, 1985, for more discussion).

It is important to show that such a system would actually help the writer before major effort was put into developing it. It might seem obvious that a system that detected comprehensibility problems would be helpful in preparing more usable technical documents. However, the research on improving documentation yields no clear solutions to how documents can be improved, especially in terms of their usefulness in a task.

It is clear that documents can be improved by skilled writers. For example, work by Smith and Kincaid (1970) started with a document that originally scored at a very high (difficult) reading grade level according to a standard readability formula. The document was rewritten, resulting in a low reading grade level score. On-the-job performance using the rewritten document was substantially better, showing that the readability formula was a useful indicator of text quality. But, the problem is that the process of rewriting the document was poorly specified. Apparently, the researchers used their best writing skills to improve the document overall, and this improvement is reflected in the readability score. The problem is that it is not clear just how the document was been modified.

Since readability scores are widely assumed to reflect word familiarity and sentence complexity, Duffy and Kabance (1982) modified sample materials in a well-defined way intended to lower the readability score formula directly, by using higher frequency words and shorter, simpler, sentences. These changes, which produced very large improvements in the readability score, made little or no difference in people's ability either to recall the
information, or to answer questions of the sort found on standard comprehension tests.

Further work by Duffy, Curran, and Sass (1983) abandoned the readability formula approach and attempted to discover whether it was possible to make substantial differences in the usability of a technical document by any means at all. In this work, a sample of a technical manual was distributed to several technical writing firms who were invited to make whatever changes they felt were necessary in the material in order to improve its usability. In some cases, the changes were drastic, but again there was little or no improvement in the usability of the material, as measured both by recall and question answering tasks. In one case the "improved" material was actually worse than the original. Thus, apparent substantial modifications that improved the quality of the documents failed to improve their usability.

One possible problem with the Duffy, et al. work is that the improvements in the material were not addressing proper aspects of the material. Perhaps even technical writing firms do not really know how to improve such material. A more likely possibility is that the task used to assess comprehensibility may be crucial to whether problems with the material will show up. It could be argued that reading is so highly task-specific that unless the reading task is very carefully designed, even large changes in the apparent comprehensibility of the material may not produce performance effects. Thus, it is not clear whether the recall and comprehension measures used in Duffy's work have the appropriate relationship to the material and its intended uses.

Another possibility is that typical comprehension tests allow people to apply experiment-specific strategies that overwhelm otherwise important properties of the material. A good example of such effects appears in Johnson and Kieras (1983) who discovered that the effect of familiarity of the content of simple technical prose depended on the reading task. In an intentional recall task, there were substantial differences in the time subjects chose to study the material, but no difference at all in their recall performance. But, in an incidental recall task, where subjects did not know they would be tested later, there were substantial recall effects. Notice that Duffy's work, and most other work in this area, has relied heavily on various recall measures in an intentional recall situation. In other tasks used by Duffy, subjects were required to search the material looking for specific facts. It is possible, as Duffy has pointed out, that this task could be accomplished without having to comprehend the material to any depth. For example, the reader might simply scan for key words. Thus, subjects may adopt strategies that allow them to successfully respond to test questions without the comprehensibility of the material playing any significant role.
The experiments reported in this paper were designed to demonstrate worthwhile effects of improving the comprehensibility of technical material. The experiments had four important features: (1) A task situation was used that was a relatively realistic model of how actual technical documents are used to work with equipment. (2) Task performance measures and on-line reading time were used, rather than recall measures, to assess comprehension. (3) The good and bad versions of the material were realistically chosen. It is surely trivial to show that one can start with a good piece of text and destroy it to the point where subjects would not be able to perform a task using it. Rather, the goal was to start with a simulated technical document that is of realistically poor quality, and then show that improving it produces improvements in performance. (4) The changes made to improve the material were well defined. That is, the problem with demonstrations such as that by Smith and Kincaid (1970) is that the changes in the document are not clearly characterized, and so are neither experimentally reproducible, nor do they provide a basis for how a computerized comprehensibility evaluation system should work.

The approach in these experiments was based on the mental model paradigm employed by Kieras and Bovair (1984). The basic paradigm is that one group of subjects learns how to operate a device strictly by rote, while another group learns "how the device works" in terms of its internal components and interconnections. A simple control panel device, shown in Figure 1, was used. The subjects who learned a mental model for the device studied a block diagram (see Figure 2) and several pages of material in which it was explained that the control panel was for a "phaser bank" on the "Starship Enterprise". This "Star Trek" fantasy context apparently only serves to interest and motivate subjects to some extent; the work reported in Kieras and Bovair (1984) shows that the fantasy context does not in fact play an important role in determining people's ability to work with the equipment. Rather, what is important is the subject's understanding of the system topology, which is how the internal components are interconnected and how they relate to the controls and indicators on the front panel.

In these mental model studies, knowledge of how the system works produced substantial improvements in people's ability to learn and remember how to operate the device. Especially sharp improvements appear when subjects are asked to infer how to operate the device. That is, they are told what the goal state is, and then asked to set the controls in such a way as to produce this goal state. Since the device has internal components that can "malfunction", it is possible to test the subject's understanding of the system by simulating various mal-
Figure 1. The control panel device used in Kieras & Bovair (1984).
Figure 2. The block diagram supplied to the mental model groups in Kieras & Novair (1984).
functions of the system and asking subjects to attempt to achieve the goal state in spite of the malfunction.

This inference paradigm was used in the present studies as a way to obtain a measure of the comprehensibility of the information about how the system works. Two groups were used, both of which studied a simulated technical manual describing how the device works. One group studied the original bad version of the manual, while the other studied an improved good version. Both groups were then asked to infer how to operate the device in various situations. Their ability to successfully infer the operating procedures can then be taken as a measure of the availability of the information in the manual. Thus, this task provides a relatively realistic measure of how changes in the quality of the material affect the reader's ability to extract information from it.

EXPERIMENT 1

In the first experiment, face validity was given high priority, and a very simple data collection approach was used. The subjects had a paper manual in hand, and their behavior was recorded on videotape.

Method

Materials and Design. The design was a simple 2-group between-subjects design. One group used a good version of the manual; the other group used a bad version. The two versions of the simulated technical manual for the device are shown in The Appendix; the versions are illustrated by the short excerpts in Table 1. Neither version of the manual included a diagram of the system, unlike the work in Kieras and Bovair (1984), in order to force subjects to rely on the content of the text. The original, or bad, version was written by an engineer in what was felt to be an engineering style. This writer had experience in constructing complex systems and preparing documentation for them, and had considerable experience with engineering documentation. Like most engineering documents, the manual was organized in terms of the structure of the system, and focused on describing the system components and their relations. In this first experiment, the manual did not contain any procedural information on how to operate the system.

The good version of the manual was prepared by simulating the process of evaluation and revision that would be performed using a computerized evaluation system. The rules shown in Table 2 defined the problems to be detected and corrected. (See Kieras, 1985 for the background and justification for such rules). These
rules address only the presentation of the ideas, so the revision did not result in any substantial changes in the content. The document was rewritten in a criticize-rewrite cycle. In the criticize phase, the author of this report carefully went through each sentence of the document and attempted to apply the rules shown in the table, making notes where the rules were violated. This phase simulates the processing that would be done by the computerized comprehensibility evaluation system proposed in Kieras (1985). After completing this phase, the author proceeded to act the part of a technical writer, and attempted to correct the problems pointed out in the criticism by rewriting the material. However, note that no substantial changes in the content were made, other than eliminating redundant material. After five passes, no more problems were detected using these rules. Table 3 illustrates how the criticize-rewrite process was done.
Table 2
Comprehensibility Rules Used to
Revise the Bad Version

Reference

1. Referents must be referred to by an unambiguous and short (2-3 words) noun phrase that is used consistently throughout the document.

2. The identity of a referent must be trivially determinable from the referencing noun phrase; no inference should be required.

3. Pronouns (generally it) should refer only to the subject referent of the preceding sentence, unless grammatical gender allows an unambiguous specification of the referent.

4. Propositional pronouns should refer only to the main proposition of the preceding sentence.

Sentence structure

1. Parenthetical expressions that contain substantive content should be avoided.

2. Relative clauses must have a relative pronoun (which, that) unless the main proposition of the clause is based on a preposition.

3. Sentences should contain no more than two clauses and should be limited to a moderate amount of new information (roughly 5 propositions).

4. A noun phrase should contain no more than about 5 propositions.

5. Self-embedding constructions, which are quite rare, must be avoided.
Table 2 (continued)

Textual Coherence

1. Textually new referents and propositions should appear only in clause predicates.

2. Material should be grouped so that the following coherence rules can be followed: A new topic is introduced only by the first sentence of a new paragraph. The subject noun phrase of each sentence in a paragraph must refer either to the subject referent of the previous sentence, or to a textually new referent introduced in the predicate of the previous sentence (chained construction).

3. Although passive constructions should be avoided, especially in truncated form, a passive construction that is required to maintain coherence should be used rather than the active construction.

Textual Organization

1. Propositions should not be repeated, especially within noun phrases, except as required to avoid ambiguity.

2. In lists that introduce items to be described in detail later, the items should be referred to by their minimum referential form; no details should appear in the list.

3. Descriptions of partially parallel structures and processes should appear in parallel independent portions of text, rather than in constructions based on corresponding and respectively.

4. A series of items should appear as an explicit list.

5. Information relevant to operating procedures should appear in IF-THEN form, using the second person (you) where possible, and should correspond to a correctly specified production system.
Table 3
Illustration of Revision Process

Original

The phasing cylinder is an Aero-Labs Model 101 type A made of tungsten carbide and internally coated with a .001 mm thick coat of reflective silver.

[too many propositions, missing relative pronouns, passives, noun phrases too long]

The cylinder is filled with number 9 industrial grade zirconium beads.

[inconsistent reference: CYLINDER, passive] <didn't correct>

Power for the phasing cylinder is obtained from the energy accumulator bipolar dielectric plates.

[sentence content redundant, inconsistent reference, passive, noun phase too long, topic shift] <sentence deleted>

When phase-modulated power is applied to the cylinder the hyperons in the nuclei of the zirconium begin to oscillate.

[inconsistent references: PHASE-MODULATED POWER, CYLINDER, passive, redundant content]

Energy is emitted through a focussing aperture in one end of the cylinder as a high intensity beam of energy.

[topic shift: ENERGY, passive, too many propositions, inconsistent reference: CYLINDER]

Revision 1

The phasing cylinder is an Aero-Labs Model 101 type A. sentence is OK]

It is made of tungsten carbide and has an internal coat of reflective silver that is .001 mm thick.

[sentence is OK] <passive maintains topic>

The cylinder is filled with number 9 industrial grade zirconium beads.

[inconsistent reference: CYLINDER, passive]

Note. Comprehension problems shown in square brackets. Writer's comments shown in angle brackets.
When energy enters the cylinder, the hyperons in the zirconium nuclei begin to oscillate.

Revision 2

The beam passes through a focusing aperture in one end of the phasing cylinder.

Revision 3

Energy that enters the cylinder produces an oscillation of the hyperons in the zirconium nuclei.

Note. Comprehension problems shown in square brackets. Writer's comments shown in angle brackets. Unchanged sentences are not shown.
The two versions are almost of identical length despite the rewriting. The good version has less redundant material in it, but certain sections are less abbreviated. In particular, the bad version used constructions based on the word respectively to describe two similar parts of the system. In the good version, these two parts were each described in full, which as it happened, almost exactly compensated in length for the redundant material that was deleted. The two versions of the manual were prepared as mock-up paper-bound manuals, arranged so that corresponding sections in the two versions occupied the same pages.

Apparatus. The control panel device was driven by a laboratory computer, which controlled the indicator lights in response to the control settings in a way defined by the simulated internal structure of the device (see Figure 2), together with the specification of the simulated malfunction state of the device. The device and its behavior were identical to that described in Kieras and Bovair, except that the device labels were changed to be non-mnemonic, as shown in Figure 3. The device could be in one of five states. In the normal state, all components operate. There are four malfunction states, each corresponding to one of the four internal components (shown as boxes in Figure 2) being inoperative.

Procedure. Subjects were run individually. They were instructed that their task was to use the manual in order to figure out how to operate the device. The subjects were given a series of trials which started with the device fully operational, and then going on to various malfunction situations. In each trial they first were allowed to read the manual, and then had to set the manual aside before trying to operate the device. Thus, they could read as long as they wished, but could not operate the device while they were reading. Likewise, while they were operating the device, they would not refer back to the manual. This allowed the videotapes to be scored for reading activity versus operating activity. Each subject was tested in the five possible situations, with the first being the normal device state, followed by the four malfunction situations. The subjects were instructed to devise the most efficient procedure for operating the device in each situation. For the malfunction situations, the subjects were told that their task was either to devise a procedure that would succeed in operating the device in spite of the malfunction, or to determine that such an alternate procedure was not possible. They repeated the task in each situation until they produced the identical procedure three times in a row. Then they went on to the next situation. The entire session was videotaped, but the actions performed on the device were recorded by the laboratory computer.
Figure 3. The labels used on the control panel device for the comprehensibility experiments.
Subjects. There were ten subjects per group, assigned at random to the two groups. Three subjects were replaced for failure to follow the instructions. The subjects were of both sexes and were recruited through campus newspaper advertisements at the University of Arizona. They were paid $5.00 for their participation in the experiment. Prior to beginning the experiment, the subjects completed a brief questionnaire on their experience with technical manuals and equipment operation. This experience ranged from essentially no experience at all, to one female subject who had recently completed several years in the Army involving extensive use of equipment and technical manuals.

Results

For each trial, the time the subjects spent reading the manual, and the time spent operating the device, were measured from the videotape. The sum of the reading time and the operating time were also determined, and will be termed the total training time. The results presented in the first three lines of Table 4 are the totals obtained by summing over all trials. For example, bad version subjects spent an average of 204.5 seconds in reading during the entire experiment.

Table 4

Mean Time (Secs) and Quality Measures for Each Group

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Bad</th>
<th>Good</th>
<th>%Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Reading Time</td>
<td>204.5</td>
<td>126.0</td>
<td>38%</td>
</tr>
<tr>
<td>Total Trying Time</td>
<td>143.0</td>
<td>76.1</td>
<td>47%</td>
</tr>
<tr>
<td>Total Training Time</td>
<td>347.5</td>
<td>202.2</td>
<td>42%</td>
</tr>
<tr>
<td>Total Logical Actions</td>
<td>74%</td>
<td>89%</td>
<td>20%</td>
</tr>
<tr>
<td>Final Logical Actions</td>
<td>80%</td>
<td>94%</td>
<td>18%</td>
</tr>
</tbody>
</table>

There are substantial and significant improvements in the total reading, operating times and the time on task. The typical size of these improvements in time is about 40%. These results are significant at the .05 level by t-tests. The subjects...
reading the rewritten good version of the manual were able to complete the task of devising a procedure for operating the device substantially faster than the subjects who read the original bad version of the manual.

The last two lines in Table 4 present a measure of the quality of the procedures that subjects devised. Based on how the device works, certain actions, such as pushing the two buttons simultaneously, are not "logical". A high proportion of these illogical actions suggests a lack of understanding of how the device worked. The table reports the percentage of logical actions, both as the total percentage over all trials, and the final percentage based on the last three trials of the data where subjects have stabilized their procedures. Over all trials, the subjects reading the good version produced considerably more logical actions than the subjects who read the bad version. Thus, the good version subjects apparently had a better understanding of how the device worked than the bad version subjects. On the last three trials, the good version subjects are still substantially more logical in their actions. This means that the bad version subjects were not able to overcome their initial disadvantage by experimenting with the device. These effects are significant at $p<.05$.

For purposes of brevity, other details of the data will only be briefly summarized. The same basic patterns as presented in Table 4 appeared in the first three trials. Generally, there is a tendency for the amount of reading to decrease very rapidly, but this is especially true for the subjects reading the good version. However, the final trials show no important reading time differences at all. The data also suggested that there are substantial differences between the different malfunction situations. In some of the situations, the differences between the good and bad version groups were much more substantial, such as a 60% difference in total training time. However, in other situations, the differences were very slight. This is consistent with our earlier work (see Kieras & Bovair, 1983). Generally, procedures for normal situations are much easier for subjects to infer than those for malfunction situations. Certain malfunction situations produce very inefficient, essentially trial and error, behavior, when subjects do not understand how the equipment works. This was observed to some extent in these data, but the small sample size does not permit definite statistical conclusions.

Discussion

Although this experiment produced substantial effects, there are several serious problems with the results. First of all, there is extremely high variability in the data; the possibility of sampling error is quite high. Furthermore, the procedure provided no information on any of the details of reading. For
example, it is impossible to tell with any reliability what part of the manual the readers referred to, or how long they spent looking at each section of the manual.

A more serious problem is that the experiment as a whole seemed to be very difficult for the subjects. Many of the subjects seemed to be confused about the basic procedure of the experiment. For example, at least one subject read through the manual, and then did not know what they were supposed to do. This level of confusion and difficulty was quite surprising, given that several experiments using the same device and same general type of material had been conducted with very few such problems.

EXPERIMENT 2

The second experiment addressed the problems appearing in the first. In order to alleviate the problem that subjects had with following the procedure, the subject was guided through the experiment by a computer running on a terminal next to the device. The subjects were given ample prompts to help them keep track of where in the procedure they were.

More detailed reading data was collected by presenting the manual on the terminal. The manual was divided into sections, and a table of contents was presented in the form of a menu, allowing subjects to choose which section they would read. Once chosen, a section had to be read in its entirety. The sentences were read one at a time and the individual sentence reading times were recorded. Thus the experiment produced detailed data on (1) which section of the manual was read at what point in the subject's interaction with the device, (2) the individual sentence reading times, and (3) the total time spent reading. Finally, in situations in which the device was malfunctioning, subjects were asked to state what they thought was wrong. These statements would indicate how well subjects understood the internal structure of the device.

A section consisting of procedural text, which described how to operate the device in the normal situation, was added to both versions of the manual. A methodological reason for this change was to get the subjects started on operating the device, and thus relieve some of the confusion observed in the first experiment. Also, this change made the manual more realistic, because real equipment manuals often include at least the normal operating procedures.

A substantive reason for including the procedural section was to test a straightforward prediction about the difficulty of acquiring procedural information from text. As suggested in Kieras and Bovair (1985), procedural text has the function of conveying production rules to the reader. If the text conveys
the rules in a clear and unambiguous fashion, then the procedure should be easy to acquire from the text. If, however, the text presents the rules in an obscure and an indirect fashion, then the reader should have to work harder to acquire the production rules. The goal in this experiment was simply to demonstrate that the expected effect could be obtained. Future experiments could then investigate the details of the process.

In the good version, the normal operating procedure section was written in such a way that would map quite directly into a sequence of production rules. Table 5 shows the good version of this section, while Table 6, based on Kieras and Bovair (1985), shows some of the production rules making up the corresponding procedure. Notice the direct correspondence between the instructions and the rules, which means that the translation from sentences to production rules is simple. In the bad version, shown in Table 7, this translation would be difficult, not only because of the complex sentence syntax, but also because the order in which the information is presented is often the reverse of how it needs to be represented in the rules. Thus, to construct a correct set of production rules, the reader would have to analyze and reorder the information in various complex ways, leading to increased reading time, or less comprehension.

Table 5

<table>
<thead>
<tr>
<th>Good Version of Normal Operating Procedure Section</th>
</tr>
</thead>
</table>

5. NORMAL OPERATING PROCEDURE

If the system is functioning normally, and the command is to fire the phaser, then do the following:

Step 1. Turn switch S1 to the A position.
Step 2. Set selector switch S2 to the X setting.
Step 3. Press button B1, and then release it.
Step 4. If indicator I4 flashes, then the phaser has fired successfully.
Step 5. When indicator I4 stops flashing, set selector switch S2 to N.
Step 6. Turn switch S1 to the B position.
Table 6

Sample Production Rules for the Normal Operating Procedure

<table>
<thead>
<tr>
<th>Rule</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>IF ((SYSTEM: NORMAL) (COMMAND: FIRE PHASERS)) THEN (ADD GOAL DO STEP1))</td>
<td></td>
</tr>
<tr>
<td>IF ((GOAL DO STEP1)) THEN ((OPERATE S1 TO A POSITION) (DELETE GOAL DO STEP1) (ADD GOAL DO STEP2))</td>
<td></td>
</tr>
<tr>
<td>IF ((GOAL DO STEP2)) THEN ((OPERATE S2 TO X POSITION) (DELETE GOAL DO STEP2) (ADD GOAL DO STEP3))</td>
<td></td>
</tr>
<tr>
<td>IF ((GOAL DO STEP3)) THEN ((OPERATE B1 PRESS, RELEASE) (DELETE GOAL DO STEP3) (ADD GOAL DO STEP4))</td>
<td></td>
</tr>
<tr>
<td>IF ((GOAL DO STEP4) (DEVICE: I4 FLASHES)) THEN ((ADD NOTE PHASER FIRED SUCCESSFULLY) (DELETE GOAL DO STEP4) (ADD GOAL DO STEP5))</td>
<td></td>
</tr>
</tbody>
</table>

...
Firing the phaser when the system is operating under normal conditions is accomplished by the setting of J4 unit source switch S2 to the X setting, preceded by placing S1 on the J4 control panel in the ON (up) or A position. Discharge of the selected accumulator into the phasing cylinder to produce the desired energy output should then be caused by the depression of the appropriate trigger button (B1). Setting of the S2 selector to the N position should then be delayed until indicator light I4 on the J4 panel ceases to flash. Following the completed operation, setting S1 to the B position should be conducted.

Method

Materials and Design. As in Experiment 1, there was a separate group of subjects for each of two versions of the materials. The materials were essentially the same as Experiment 1 but with the following modifications. First, the new section on the normal operation procedure was included. This meant that subjects should have to do substantial inference of how to operate the device only in the malfunction situations. Second, the manual was divided into sections, each section numbered and designated with a heading. These headings and numberings were included in a menu, which is shown in Table 8.

Apparatus and Procedure. The subjects read the materials and all instructions on a standard video terminal placed next to the device. The subjects were run one at a time. The subjects were first required to read the manual all the way through, reading one sentence at a time. This gave an initial reading time on each sentence in both versions of the manual. Following this first reading, the menu shown in Table 8 was presented. The subjects could either proceed to choose a section of the manual to reread. If they chose a section, they read all sentences in that section at one time, and then returned to the manual. They could continue to choose and reread the menu. They could continue to choose and reread the sections of the manual until they were ready to operate the device. At any point during operating the device, they could abort the attempt and return to the menu. Subjects repeated this normal operation sequence until they had succeeded in getting the device to operate 3 times.
After the subjects completed the normal situation, they were briefly instructed about the experimental procedure for the malfunction situations. In each malfunction situation, they began with the menu and could either attempt to operate the device in that situation, or go back to read any section of the manual of their choice. They were instructed that after they completed operating the device, they had to type a code character on the terminal to indicate the final outcome of the trial. There were three possible codes: (1) a success code indicating that the device was operated successfully with no problem, as in the normal situation, (2) a code that the operation was successful, but there was a malfunction in the device, or (3) a code indicating that the operation was unsuccessful because, due to a malfunction, it is not possible to get the device to operate. In the second and third cases the subjects typed in a diagnosis statement, which was a one-line description of the malfunction.

Following the instructions, the subjects performed each of the four malfunction situations. In each situation, they were required to arrive at a correct decision regarding the outcome, and to type in the correct code. If they typed in an incorrect code, they were informed and returned to the menu, and had to repeat the situation until they produced the correct outcome code. Then they proceeded to the next situation.

After completing the last malfunction situation, they were given a sequence of 25 "practice" trials, arranged in 5
blocks, each block consisting of the normal situation and the four malfunction situations in a random order.

Subjects. Twenty subjects per group were run, recruited as in Experiment 1 and paid $5.00 for participating in the experiment.

Results

Total reading time. The total reading times are the sum of the reading times over all trials on each section; the means for each version and section are shown in Table 9. An analysis of variance showed that the main effect of version was significant only at $p=.063$, even though the bad version was read for 23% longer than the good version. There was a significant effect of manual section ($p<.05$), but the interaction between version and section was not significant ($p>.1$). Individual $t$-tests comparing good and bad versions of individual sections were nonsignificant ($p>.1$), except for the system overview section and the normal operating procedure sections ($p's<.01$). Although the good version is read faster, these statistically inconsistent effects on total reading times contrast sharply with the substantial effects observed in Experiment 1.

Table 9

Total Reading Times for Manual Sections

<table>
<thead>
<tr>
<th>Section</th>
<th>Bad</th>
<th>Good</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Overview</td>
<td>166.3</td>
<td>110.2**</td>
</tr>
<tr>
<td>Energy Booster</td>
<td>142.3</td>
<td>123.9</td>
</tr>
<tr>
<td>Accumulators</td>
<td>224.5</td>
<td>200.8</td>
</tr>
<tr>
<td>Phasing Cylinder</td>
<td>90.0</td>
<td>70.6</td>
</tr>
<tr>
<td>Normal Operating Procedure</td>
<td>110.1</td>
<td>63.8**</td>
</tr>
<tr>
<td>Malfunctions</td>
<td>98.2</td>
<td>73.6</td>
</tr>
<tr>
<td>Mean</td>
<td>138.6</td>
<td>107.1</td>
</tr>
</tbody>
</table>

**$t$-test significant at $p<.01$
First time reading. The first time reading times were obtained by totaling the reading times on the individual sentences in each section of the manual during the first reading only, before subjects had the opportunity to look back at any section. Table 10 shows the means. An analysis of variance showed that the version effect was not significant ($p=.13$), but there were strong effects of section, and a significant version by section interaction ($p's<0.1$). Individual $t$-tests comparing the sections between the two versions showed that the good version reading was faster than the bad on the system overview, normal operating procedure, and malfunction sections ($p's<.05$) and not significant otherwise. Note that the reading time for the accumulators section is actually reversed.

Thus, even the first time reading results show inconsistent effects on reading time, even though the variability in these data was considerably lower than the total time data.

Table 10

First Time Reading Times on Manual Sections

<table>
<thead>
<tr>
<th>Section</th>
<th>Bad</th>
<th>Good</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Overview</td>
<td>118.8</td>
<td>91.5*</td>
</tr>
<tr>
<td>Energy Booster</td>
<td>85.4</td>
<td>81.8</td>
</tr>
<tr>
<td>Accumulators</td>
<td>120.9</td>
<td>127.9</td>
</tr>
<tr>
<td>Phasing Cylinder</td>
<td>57.1</td>
<td>50.9</td>
</tr>
<tr>
<td>Normal Operating Procedure</td>
<td>64.9</td>
<td>37.3**</td>
</tr>
<tr>
<td>Malfunctions</td>
<td>52.5</td>
<td>40.0*</td>
</tr>
<tr>
<td></td>
<td>83.3</td>
<td>71.6</td>
</tr>
</tbody>
</table>

*$t$-test significant at $p<.05$; ** at $p<.01$

In considering the reading time results, it should be noted that the good version had more sentences of shorter length. The overhead involved in a sentence by sentence reading paradigm may thus have made the good version take a disproportionately long time to read.
Reading behavior. The choices made by subjects of which section to reread, and when in the experiment they chose to reread it, was examined. Figure 4 shows the number of subjects who reread any section of the manual as a function of the sequence of situations in the experiment. Overall, it appears that bad version subjects had to look back at the manual more often than good version subjects. Note that during the practice situations, re-reading occurred relatively rarely. Ninety percent of the bad version subjects had to refer back to the manual before trying the normal operating procedure, but only half of the good version subjects did so. This difference is significant by a chi-square test, p<.01.

In more detail, Figure 5 shows the number of subjects referring back to the section on normal operating procedure. Seventy-five percent of the bad version subjects looked back at this section before attempting the normal operating procedure, while only 40% of the good version subjects did so. The difference is significant by a chi-square test, p<.05.

Figure 6 shows the number of subjects who looked back at the section of the manual that described the "accumulators". This section contains most of the information required for understanding the malfunction situations. Many more bad version than good version subjects looked back at the accumulator section during the malfunction situation attempts. Only 20% of the good version subjects looked back, as contrasted with 70% of the bad version subjects. This difference is significant by a chi-square test, p<.01.

Thus the reading behavior results show that most bad version subjects needed to refer back to the manual in order to obtain the information required by the task. This suggests that the good version of the manual was indeed more intelligible to the subjects; most subjects understood the important information after only one reading.

Diagnosis statements. The diagnosis statements were printed out on individual slips of paper and sorted into categories by a judge who was blind to the original experimental conditions. The proportions of responses in the various categories are shown in Table 11. The responses are grouped into three major categories corresponding to the source of the information underlying the responses, and into subcategories that provide more detail. The manual-based responses referred directly to some component of the system that was described in the manual. In contrast, the superficial responses contain information that corresponds only to the "outside" of the device, such as a history of what actions were done, an observation of some event, such as an
Figure 4. Number of subjects referring back to any section of the manual, as a function of point in the experiment.
Figure 5. Number of subjects referring back to the normal operating procedure section, as a function of point in the experiment.
References to Accumulator Section

Number of Subjects

BN, DN - Before and During Normal Try
BM, DM - Before and During Malfunction Trials
BP, P1-5 - Before and During Practice Trials

Figure 6. Number of subjects referring back to the accumulators section, as a function of point in the experiment.
indicator failing to come on, or a surface malfunction, such as an indicator light being burnt out. Such responses seemed to suggest that the subject was assuming that the device had no internal components, even though the simulated manual described the internal components at considerable length.

The difference between versions in distribution of the response over the three major categories is significant by a chi-square test, \( p<.01 \). The good version subjects' diagnoses are more often based on the manual content, while the bad version group produces more superficial responses.

There are also important differences between the groups on the subcategories within the major categories. All three chi-square tests on the distributions within each major category are significant, \( p<.05 \). Within the manual-based responses, both groups are equally often correct in terms of proportion, but the good version group makes more such responses. It is not clear why the good version group makes more wrong responses than the bad version. Within the superficial responses, the bad version group makes more broken part responses. Within the other

Table 11
Proportion Diagnosis Responses in Each Major and Subcategory
(N=480 in each group)

<table>
<thead>
<tr>
<th>Major Category</th>
<th>Version</th>
<th>Subcategory</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual-Based</td>
<td>Bad</td>
<td>Good</td>
<td>Correct</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Incomplete</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wrong</td>
</tr>
<tr>
<td>Superficial</td>
<td>.37</td>
<td>.14</td>
<td>Observation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Broken Part</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>History of Actions</td>
</tr>
<tr>
<td>Other</td>
<td>.14</td>
<td>.13</td>
<td>Not trying</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Confused</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Frivolous</td>
</tr>
</tbody>
</table>
responses, the good version subjects were more often confused, while the bad version subjects more often made "not trying" and "frivolous" responses suggesting that they had given up trying to understand what was wrong with the system.

Discussion

The reading times on the bad version of the manual tended to be longer than on the good version, but the effect was not significant in a consistent and convincing way. Compared to the bad version subjects, the good version subjects did not have to look back at the manual as much, and apparently knew more about the device, as indicated by their diagnostic responses. Thus, despite the inconsistent reading time effects, the good version was clearly more usable than the bad version.

GENERAL DISCUSSION

Overall Conclusion

The overall result is that the systematic changes made in the original version of the simulated technical manual were effective in improving the ability of subjects to use the manual to operate the device. In the first experiment, these effects showed up strongly both in time measures and in other measures that reflected subjects' understanding of the device. In the second study, the time effects were inconsistent, but other clear effects appeared. With the bad version of the manual, subjects needed to refer to it more often, and were unable to be as definite and correct about the malfunctions.

The purpose of these studies was to determine whether well-defined changes in the original version of the simulated manual would result in definite improvements in subjects' ability to use the manual in a realistic task. If so, then a computerized system that was capable of indicating where such changes should be made would probably be effective in helping writers improve the usability of technical materials. Since on the whole, there were definite improvements due to the rewritten manual, these results suggest that such a computerized comprehensibility evaluation system would indeed be useful and effective.

Methodological conclusions

There is an important methodological problem of why the reading time effects that were so robust in the first experiment were weak and inconsistent in the second, and why researchers such as Duffy et al. have been unable to obtain clear effects of improving documents. This is an important problem to understand, because any efforts to improve the quality of technical materials must be able to reliably demonstrate real comprehension effects; reading time is clearly an important comprehension measure.
One explanation for the weak effects in Experiment 2 is that there was simply a procedural artifact, as discussed above. Several differences between the good and bad materials may have made simple overall reading time differences, based on single sentence reading times, quite misleading. More detailed analyses, and repeating the experiment without single sentence reading, would show whether this explanation was true.

Another explanation for the difference between the two experiments lies in the fact that there could have been gross differences in the overall information processing load imposed on subjects. In Experiment 1, subjects were not only required to comprehend the manual and apply it to the device, but also had to figure out just what they were supposed to do in the experiment as well. In contrast, the careful procedural prompting in Experiment 2 might have allowed subjects to devote more of their information processing capacity to understanding the manual. Likewise, since the manual used in Experiment 2 provided the subjects with a procedure, the subjects were in a position to get started in the experimental task with much less confusion. If these general processing load differences are responsible for determining the presence and absence of reading time effects, future research will have to be carefully planned with this in mind.

A statistical explanation is that these experiments combined both the reading of prose of natural complexity and a fairly complex problem solving task. In these situations, extremely high variability in the data would be expected. The lack of substantial significant effects in the second experiment could simply be due to these purely statistical factors.

A final set of explanations has to do with how subjects respond to the reading task itself. Subjects reading the bad version may optimize their performance by not attempting to read the manual carefully, and using their efforts on experimenting with the device. This is a good strategy because the device is in fact fairly easy to figure out by trial and error (see Kieras and Bovair, 1983, 1984). If subjects adapt their reading strategy to the situation, the overall reading time may be an ambiguous indicator of reading difficulty. Thus, the good version subjects in the second experiment may have read the material relatively longer because they thought that such effort would be rewarded, while the bad version subjects read it relatively quickly because they dismissed the manual as a useful source of information. This is consistent with reports on how actual incomprehensible manuals are dismissed by electronics maintenance workers (Bond & Towne, 1979).
In conclusion, applying standard comprehension research paradigms to realistic tasks involving technical manuals will have to address some important methodological problems.

References


APPENDIX

Bad and Good Versions
of the Simulated Technical Manual
OPERATOR'S MANUAL
for
SHIP-MOUNTED PHASER SYSTEM
MODEL 22-A

Prepared by
Rawlston Lewis Starship Equipment, Inc.
Founder's Port, Cygnus V.

STARSHIP ARMAMENTS DIVISION
STARFLEET SYSTEMS COMMAND
UNITED FEDERATION OF PLANETS
The Rawlston Lewis Model 22-A is a ship-mounted phaser system. It uses conventional technology to generate a beam of energy that has a strength of 30 billion joules. The phaser system operates by applying a phase-modulated current pulse to a phasing cylinder that is filled with zirconium. The phaser system is suitable for deterring most armed alien aggressors.

The phaser system consists of a control panel, an energy booster, a main energy accumulator, a secondary energy accumulator, and a phasing cylinder.

The control panel is type J-4. It has been modified by adding controls that allow using a secondary energy accumulator. The modification is in accordance with LSNA-Spec 45 A Rev. 4006.23.

The phaser system gets power from the ship's dilithium-controlled matter-antimatter conversion system. If you place the power switch S1 in the A position, the phaser system gets power. If the indicator light I1 is on, then the phaser system is getting power.
The purpose of the energy booster is to magnify the cmf potential of the power to a 60-gigavolt phase-modulated current pulse. The energy booster meets all Starfleet specifications (10A-A-21, 206B-3 rev 4, 301-J). The energy booster operates off the standard energon conversion cycle. In the energon conversion cycle, power that is generated by the ship's conversion system is converted to a current pulse. The current pulse is produced by the collapse and expansion of energon rings that are inside a power vessel. The power vessel is made by Wilkes-Farmer. It has a shield that consists of a 5mm thick coating of platinum-silver alloy. The shield limits radiation emission to .5 rad/hour. The energon rings rotate clockwise in a normally collapsed state inside the power vessel. They are G.E. size 00. The energon rings begin to expand when the power enters the power vessel. They continue to expand until they reach Blofeld's limit. Then they instantly collapse, releasing the current pulse. The current pulse goes to the main and secondary energy accumulators. If the indicator light 12 is on, then the energy booster is operating normally.
The main and secondary accumulators are General Electric bipolar plasmation dielectric energy accumulators. They operate by using the current pulse from the energy booster to create a plasma field between two bipolar dielectric plates. When the plates discharge into the phasing cylinder, the phaser system fires. Since the accumulators may malfunction, there is a secondary accumulator that acts as a backup to the main accumulator.

The energy source selector switch S2 on the control panel controls which accumulator powers the phasing cylinder. The selector switch has three settings, as follows:

1. If you select the Neutral (N) setting, then neither the main nor the secondary accumulator can transfer energy to the phasing cylinder. If you do not anticipate immediate operation of the phaser system, then you must maintain the selector switch in the Neutral setting. This is required by Starfleet OPSREG 100.35.7.

2. If you select the A setting, then the main accumulator can send energy to the phasing cylinder. The main accumulator fire control button B1 controls the flow of energy from the main accumulator to the selector switch. If you press the B1 button, and the selector switch is on the A setting, then the main accumulator discharges into the phasing cylinder.

3. If you select the B setting, then the secondary accumulator can send energy to the phasing cylinder. The secondary accumulator fire control button B2 controls the flow of energy from the secondary accumulator to the selector switch. If you press the B2 button, and the selector switch is on the B setting, then the secondary accumulator discharges into the phasing cylinder.

If the indicator light I3 is on, then the main accumulator is operating properly. The secondary accumulator does not have an indicator light that signals normal operation.
The phasing cylinder is an Aero-Labs Model 101 type A. It is made of tungsten carbide and has an internal coat of reflective silver that is 0.001 mm thick. The phasing cylinder is filled with number 9 industrial grade zirconium beads.

Energy that enters the cylinder produces an oscillation of the hyperons in the zirconium nuclei. The oscillation then produces a high intensity beam of energy. The beam passes through a focusing aperture in one end of the phasing cylinder. The radiation from the initial beam causes the oscillation to continue. The oscillation then produces another beam of energy. The oscillation depletes the zirconium nuclei after four beams of energy. This causes the oscillation to stop. If indicator light I4 flashes four times, then the phaser system has emitted the beams of energy.

If the phaser system malfunctions, then the malfunction is probably in the energy booster, the accumulators, or the phasing cylinder. The malfunction is probably not in the wiring, indicator lights, or control switches. This is because the control panel is triply redundant and has not failed in 1000 M man-hours of operation. If a malfunction occurs, then you should use standard troubleshooting procedures to correct it. You should develop procedures for emergency situations.
OPERATOR'S MANUAL
for
SHIP-MOUNTED PHASER SYSTEM
MODEL 22-A

Prepared by
Rawlston Lewis Starship Equipment, Inc.
Founder's Port, Cygnus V.

STARSHIP ARMAMENTS DIVISION
STARFLEET SYSTEMS COMMAND
UNITED FEDERATION OF PLANETS
The Rawlston Lewis Model 22-A ship-mounted phaser system uses conventional technology to generate a thirty billion joule (30,000 M joule) beam of energy, by the application of a phase-modulated energy field to a zirconium filled phasing cylinder, suitable for the deterring of most armed alien aggressors. The Model 22-A consists of a type J-4 control panel (modified by the addition of controls allowing the use of a secondary power accumulator as per LSNA-Spec 45 A Rev. 4006.23), a 60 gigavolt phase-modulated energon cycle current pulse energy booster, two (2) heavy duty General Electric bipolar plasmation dielectric energy accumulators and an Aero-Labs model 101 tungsten carbide phasing cylinder filled with number 9 industrial grade zirconium beads.

All power to the Rawlston-Lewis Model 22-A is obtained from the ship's on-board dilithium-controlled matter-antimatter conversion system. The secondary General Electric bipolar plasmation dielectric energy accumulator does not have a Normal Operation Indicator on the J-4 when modified according to LSNA-Spec 45 A Rev. 4006.23. Input of power to the system is indicated by illumination of indicator light I1.

Power for the main and secondary General Electric bipolar plasmation dielectric energy accumulators originates in the 60-gigavolt phase-modulated current pulse energy booster. This booster meets all Starfleet specifications (10A-A-21, 206B-3 rev 4, 301-J). Access to the ship's main power converters is obtained by placing the switch S1 in the A position.
The purpose of the phase-modulated energy booster is to magnify the cmf potential of the power obtained from the shipboard power converters to a 60-gigavolt phase-modulated current. The booster operates off the standard energon conversion cycle, whereby power generated by the ship's matter-antimatter conversion system is converted to a 60-gigavolt phase-modulated current pulse through the collapse and expansion of rotating energon rings within a shielded Wilkes-Farmer power vessel. The Wilkes-Farmer power vessel has been shielded with a 5mm thick coating of platinum-silver alloy limiting radiation emission to .5 rad/hour. Internal to the power vessel, the energon rings rotate (clockwise rotation) in a normally collapsed state. These rings are G.E. size 00. When power, obtained from the ship's matter-antimatter converter by placing switch S1 in the A position, is inserted into the power vessel, the normally collapsed rings begin to expand. Expansion continues until they reach Blofeld's limit, at which time they instantly collapse, releasing a phase-modulated current pulse to the main and secondary bipolar power accumulators. Normal operation of the power booster is indicated by the light (labeled I2) on the modified J-4 control panel being on.
The main and secondary General Electric bipolar plasmation dielectric accumulators operate by using the phase-modulated 60-gigavolt current obtained from the power booster to create a plasma field between two bipolar dielectric plates. When the phaser is operated, the plates discharge into the phasing cylinder. Since the accumulators are subject to the appearance of malfunction, provision has been made to provide a secondary General Electric plasmation accumulator acting as a backup to the main storage system. Which accumulator operative is controlled through the energy source selector switch S2 on the J-4 panel.

Selector S2 on the J-4 control panel has three settings. Energy can not be transferred from either the main accumulator or the secondary dielectric plates to the cylinder if the Neutral (N) setting is selected. Starfleet OPSREG 100.35.7 requires that the J-4 S2 be maintained in the Neutral (N) setting when immediate operation of the Model 22-A phaser system is not anticipated. The A and B settings of the J-4 S2 switch connect the corresponding dielectric energy accumulators, the main accumulator and the secondary energy storage system, respectively, to the phasing cylinder, through the fire control buttons B1 and B2 corresponding to the main and secondary systems, which control the flow of energy from the accumulators \(S2\). When the corresponding button (B1 or B2) is pressed the selected accumulator discharges into the phasing cylinder. Proper operation of the main accumulator is indicated on the J-4 control panel when the indicator light (I3) is on.
The phasing cylinder is an Aero-Labs Model 101 type A made of tungsten carbide and internally coated with a .001mm thick coat of reflective silver. The cylinder is filled with number 9 industrial grade zirconium beads. Power for the phasing cylinder is obtained from the energy accumulator bipolar dielectric plates. When phase-modulated power is applied to the cylinder the hyperons in the nuclei of the zirconium begin to oscillate. Energy is emitted through a focussing aperture in one end of the cylinder as a high intensity beam of energy. The radiation from the initial beam sustains the oscillation of the hyperons leading to the release of another beam of energy. The release of four beams of energy depletes the zirconium nuclei and thus causes a cessation of oscillation. Emission of the pulsed phaser beam is indicated by the flashing four (4) times of indicator I4 on the J-4 control panel.

The J-4 control assembly has been modified to accept a secondary accumulator and its corresponding controls. Because the J-4 unit is triply redundant and has not failed in 1000 M man hours of operation, malfunctions of the 22-A ship-mounted phaser are not likely to be due to failures in the wiring, indicator lights, or control switches, but generally will occur in the energon cycle power booster, the dielectric accumulators, or the phasing cylinder subsystem. Standard troubleshooting procedures should be used to correct malfunctions as they occur, and emergency compensation procedures should be evolved for use in critical situations.
Robert Ahlers  
Code N711  
Human Factors Laboratory  
NAVRADADCEN  
Orlando, FL 32815

Dr. Ed Aiken  
Navy Personnel R&D Center  
San Diego, CA 92152

Dr. Meryl S. Baker  
Navy Personnel R&D Center  
San Diego, CA 92152

CDR Robert J. Biersner, USN  
Naval Biodynamics Laboratory  
P.O. Box 39407  
New Orleans, LA 70189

Dr. Alvan Bittner  
Naval Biodynamics Laboratory  
New Orleans, LA 70189

Code N711  
Attn: Arthur S. Bliwes  
Naval Training Equipment Center  
Orlando, FL 32813

Dr. Robert Blanchard  
Navy Personnel R&D Center  
San Diego, CA 92152

Dr. Nick Bond  
Office of Naval Research  
Liaison Office, Far East  
APO San Francisco, CA 96502

Lt. Alexander Bory  
Applied Psychology  
Measurement Division  
NAVRADADCEN  
NAS Pensacola, FL 32508

Dr. Richard Braby  
TAES  
Naval Training Equipment Center  
Orlando, FL 32815

Dr. Robert Breaux  
NAVRADADCEN  
Code N-095R  
Orlando, FL 32815

Dr. Richard Cantone  
Navy Research Laboratory  
Code 7210  
Washington, DC 20375

CDR Robert Carroll  
NAVOD 115  
Washington, DC 20370

Dr. Fred Chang  
Navy Personnel R&D Center  
San Diego, CA 92152

Dr. Susan Chipman  
Code 442PT  
Office of Naval Research  
800 N. Quincy St.  
Arlington, VA 22217

Chief of Naval Education and Training  
Liaison Office  
Air Force Human Resource Laboratory  
Operations Training Division  
WILLIAMS AFB, AZ 85534

Dr. Stanley Collier  
Office of Naval Technology  
800 N. Quincy Street  
Arlington, VA 22217

CDR Mike Curran  
Office of Naval Research  
800 N. Quincy St.  
Code 270  
Arlington, VA 22217

Dr. Charles E. Davis  
Personnel and Training Research  
Office of Naval Research (Code 442PT)  
800 North Quincy Street  
Arlington, VA 22217

Edward E. Edwards  
NAVTTEC  
Naval Air Station  
Corpus Christi, TX 78419

Dr. John Ellis  
Navy Personnel R&D Center  
San Diego, CA 92152
Dr. Richard E. Elster  
Department of Administrative Sciences  
Naval Postgraduate School  
Monterey, CA 93940

Dr. Carl E. England  
Naval Health Research Center  
Code 8660 Environmental Physiology Dept  
P.O. Box 85121  
San Diego, CA 92158

Dr. Marshall J. Farr  
2520 North Vernon Street  
Arlington, VA 22207

Dr. Pat Faderico  
Code P13  
NPRDC  
San Diego, CA 92152

Dr. Cathy Fernandez  
Navy Personnel R&D Center  
San Diego, CA 92152

Mr. Paul Foley  
Navy Personnel R&D Center  
San Diego, CA 92152

Dr. Jude Franklin  
Code 7510  
Naval Research Laboratory  
Washington, DC 20375

LT Steven D. Harris, MSC, USN  
R&D 1, Box 243  
Arlington, VA 22414

Ms. Rebecca Hetter  
Navy Personnel R&D Center (Code 62)  
San Diego, CA 92152

Dr. Jim Holman  
Code 51  
Navy Personnel R & D Center  
San Diego, CA 92152

Mr. Dick Hoshaw  
NAVOP-152  
Arlington Annex  
Room 207A  
Washington, DC 20370

Dr. Edward E. Huthwaite  
Navy Personnel R&D Center  
San Diego, CA 92152

Dr. Kenan J. Karr  
Chief of Naval Education and Training  
Code 0542  
Naval Air Station  
Pensacola, FL 32509

Dr. Peter L. Kinsaid  
Training Analysts & Evaluation Group  
Dept. of the Navy  
Orlando, FL 32815

CRAWAL D. Lambert  
CSCW Systems & Technology Dept.  
Naval Ocean Systems Center  
San Diego, CA 92152

Dr. Samui Lang  
Navy Personnel R&D Center  
San Diego, CA 92152

Dr. William L. Moly (62)  
Chief of Naval Education and Training  
Naval Air Station  
Pensacola, FL 32509

Dr. Kneale Marshall  
Chairman, Operations Research Dept.  
Naval Post Graduate School  
Monterey, CA 93940

Dr. James F. McBride  
Navy Personnel R&D Center  
San Diego, CA 92152

Dr. Barbara McDonald  
Navy Personnel R&D Center  
San Diego, CA 92152

Dr. Joe McLachlan  
Navy Personnel R&D Center  
San Diego, CA 92152

Dr. James McMichael  
Navy Personnel R&D Center  
San Diego, CA 92152
Navy

* Dr. Al Hewett
Office of Naval Research
Code 471
800 N. Quincy
Arlington, VA 22217

* Dr. William Montague
NPRDC Code 13
San Diego, CA 92152

* Ms. Kathleen Moreno
Navv Personnel R&D Center (Code 62)
San Diego, CA 92152

* Library, Code P201L
Navv Personnel R&D Center
San Diego, CA 92152

* Technical Director
Navv Personnel R&D Center
San Diego, CA 92152

* Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

* Office of Naval Research
Code 433
800 N. Quincy Street
Arlington, VA 22217

* Office of Naval Research
Code 441NP
800 N. Quincy Street
Arlington, VA 22217

* Psychological Sciences Division
Code 442
Office of Naval Research
Arlington, VA 22217

* Director
Engineering Psychology Program
Code 442EP
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

* Organizational Effectiveness
Research Program, Code 442OE
Office of Naval Research
Arlington, VA 22217

Navy

6 Personnel & Training Research Program
Code 442PT
Office of Naval Research
Arlington, VA 22217

1 Psychologist
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101

1 Office of the Chief of Naval Operations
Research Development & Studies Branch
DP 115
Washington, DC 20350

1 Nair Paulson
Code 52 - Training Systems
Navy Personnel R&D Center
San Diego, CA 92152

1 LT Frank L. Patno, MSC, USN (Ph.3)
CNET (N-422)
NAS
Pensacola, FL 32508

1 Dr. Gary Pocock
Operations Research Department
Code 55PK
Naval Postgraduate School
Monterey, CA 93940

1 Dr. Gil Ricard
Code N711
NTED
Orlando, FL 32817

1 Dr. Bernard Rialland
Navy Personnel R&D Center
San Diego, CA 92152

1 William Rizzo
Code 712 NTED
Orlando, FL 32817

1 Dr. Carl Ross
CNET-PED
Building 90
Great Lakes NTC, IL 60068
1) Dr. Donald Woodward   
Office of Naval Research (Code 441)  
300 North Quincy Street  
Arlington, VA 22217

2) Dr. Wallace Wulff, III  
Naval Personnel R&D Center  
San Diego, CA 92152

3) Capt. Bruce Young   
COMNAVCRUITCOM  
Code 21  
4015 Wilson Blvd  
Arlington, VA 22206

4) Cdr. Joe Young  
HS, MRPDC  
ATTN: MRPDC-P  
2500 Green Bay Road  
North Chicago, IL 60064

5) Dr. Steven Zornetzer   
Associate Director for Life Sciences  
Office of Naval Research  
300 N. Quincy St.  
Arlington, VA 22217

6) Marine Corps  
1) Capt. Rick Butler   
CAT Project Office  
HQ. Marine Corps  
Washington, DC 20380

2) Mr. Paul Dicenzo   
Commandant of the Marine Corps  
Code LDC-4  
Washington, DC 20380

3) H. William Greenup   
Education Advisor (COL)  
Education Center. MCDEC  
Quantico, VA 22134

4) Maj. John Keene   
ADP Systems Branch  
CC Development Center (2104)  
MCDEC  
Quantico, VA 22134

5) Col. Ray Leidich   
Headquarters. Marine Corps  
MPI  
Washington, DC 20380

6) Headquarters. U. S. Marine Corps  
Code MPI-20  
Washington, DC 20380

7) Lt. Col. Jim Murphy   
HQ. Marine Corps  
Code MRRP  
Washington, DC 20380

8) Special Assistant for Marine Corps Matters  
Code 100M  
Office of Naval Research  
800 N. Quincy St.  
Arlington, VA 22217

9) Dr. A.L. Slafsosky   
Scientific Advisor (Code RD-1)  
HQ, U.S. Marine Corps  
Washington, DC 20380

10) Dr. Sydell Weiss   
Headquarters. USMC  
Code TDE 22  
Rm 2701  
Navy Annex  
Washington, DC 20380
Marine Corps

Major Frank Schanen, USMC
Headquarters, Marine Corps
Code "PI-20"
Washington, DC 20330

Army

1 Technical Director
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22330

1 Mr. J. Barber
HSS, Department of the Army
DAPE-IER
Washington, DC 20310

1 Dr. Kent Eaton
Army Research Institute
5001 Eisenhower Blvd.
Alexandria, VA 22330

1 Lt. Col Rich Enlich
H3, Dept. of the Army
CCSA(DPCS-DFM)
Washington, DC 20310

1 Dr. Beatrice J. Farr
U. S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22330

1 Dr. Myron Fischl
U. S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22330

1 Donald Haggard
Ft. Knox Field Unit
Army Research Institute
Steele Hall
Ft. Knox, KY 40121

1 Lt. Col. Ron Hanaer
USARCHD-AR
Ft. Sheridan, IL 60067

1 Chief, ARI Field Unit
P. O. Box 5787
Presidio of Monterey
Monterey, CA 93944

1 Dr. Milton S. Katz
U. S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22330
Dr. Joseph Martin
Army Research Institute
5001 Eisenhower Blvd.
Alexandria, VA 22333

Dr. Karen Mitchell
Army Research Institute
5001 Eisenhower Blvd.
Alexandria, VA 22333

Dr. William E. Nordbrock
FMC-4000 Box 25
APO, NY 96710

Dr. Harold F. O'Neill, Jr.
Director, Training Research Lab
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Commander, U.S. Army Research Institute
for the Behavioral and Social Sciences
ATTN: PERI-DR (Dr. Judith Orasanu)
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Ray Perez
5001 Eisenhower Avenue
PERI-IC
Alexandria, VA 22333

Joseph Psotka, Ph.D.
ATTN: PERI-IC
Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

Mr. Robert Ross
U.S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Mr. Lou Ruberton
DAPE-MPA-C3
Department of the Army
Washington, DC 20310

Dr. Robert Sasaor
U.S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333
Air Force

1. Air Force Human Resources Lab
   AFHRL, APS
   Brooks AFB, TX 76235

   Life Sciences Directorate, FL
   Bolling Air Force Base
   Washington, DC 20332

3. Dr. Earl A. Allworth
   AFHRL/AFD
   Brooks AFB, TX 76235

4. Col. Roger Campbell
   AFHRL/LRT
   Pentagon, Room 4E190
   Washington, DC 20330

5. Dr. Raymond E. Christianson
   AFHRL/MOE
   Brooks AFB, TX 76235

6. Brian Dollsah
   AFHRL/AFD
   Lowry AFB, CO 80237

7. Dr. Robert Denton
   AFHRL/PMXO
   Bolling AFB, TX 78148

8. Dr. Alfred R. Fregly
   AFHRL/AFD
   Bolling AFB, DC 20332

9. Dr. Thomas Killion
   vDRI
   P. O. Box 44
   Higley, AZ 85236

10. Dr. Patrick Killion
    AFHRL/MOE
    Brooks AFB, TX 78235

11. Dr. T. M. Longridge
    AFHRL/OTE
    Williams AFB, AZ 85224

12. Dr. Sylvia R. Mauer (TAMU)
    HQ Electronic Systems Division
    Hanscom AFB
    Bedford, MA 01731
Department of Defense

1: Mr. Bob Brandwein
Defense dipping Data Center
500 Cameron El Estero, #200
Monterey, CA 93940

2: Mr. J. Burgener
MEPCOM
1850 Green Bay Road
North Chicago, IL 60064

3: Dr. Dennis Blease
Computer Education Coordinator
200 Dependent Schools
2401 Eisenhower Avenue
Room 172
Alexandria, VA 22331

4: Cmnr. T. Dean
MEPCOM
1850 Green Bay Road
North Chicago, IL 60064

5: Defense Technical Information Center
Camber Station, Bldg S
Alexandria, VA 22331

6: Dr. Craig H. Fields
Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

7: Dr. Frank Kramer, Ph.D.
Department of Clinical Investigation
Walter Reed Army Medical Center
6300 Georgia Ave. N.W.
Washington, DC 20017

8: Dr. Anita Lancaster
Accession Policy
DASD/HUMCAP/MAP
Pentagon, Room 2B771
Washington, DC 20301

9: Dr. Jerry Lennus
DASD/HUMCAP
Washington, DC 20301

Department of Defense

1: Dr. Clarence McDonald
HQ. MEPCOM
1850 Green Bay Road
North Chicago, IL 60064

2: Military Assistant for Training and Personnel Technology
Office of the Under Secretary of Defense
for Research & Engineering
Room 3E129, The Pentagon
Washington, DC 20301

3: Cpl. Van Fazner
HQ. MEPCOM
ATTN: Director MEPCOM
1850 Green Bay Road
North Chicago, IL 60064

4: Dr. W. Steve Seidman
Office of the Assistant Secretary of Defense (MEA & RF)
23269 The Pentagon
Washington, DC 20301

5: Mr. John Sweeney
MEPCOM
1850 Green Bay Road
North Chicago, IL 60064

6: Major Jack Thorne
DASD/HUMCAP
1400 Wilson Blvd.
Arlington, VA 22209

7: Dr. Robert A. Wiser
U.S. Army Institute for the Behavioral and Social Sciences
3001 Eisenhower Avenue
Alexandria, VA 22331
Private Sector

1. Dr. anders Ericsson
Department of Psychology
University of Colorado
Boulder, CO 80309

2. Dr. Wallace Feurzeig
Department of Educational Technology
99 Moulton St.
Cambridge, MA 02278

3. Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

4. Dr. Debre Gentner
University of Illinois
Department of Psychology
Urbana, IL

5. Dr. Robert Glaser
Learning Research & Development Center
University of Pittsburgh
3939 O'Hara Street
PITTSBURGH, PA 15260

6. Dr. Marvin D. Slocik
217 Stone Hall
Cornell University
Ithaca, NY 14853

7. Dr. Josiah Soguen
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

8. Dr. Henry M. Half
Half Resources
4919 33rd Road, North
Arlington, VA 22207

9. Dr. Reid Hastie
Department of Psychology
Northwestern University
Evanston, IL 60201

10. Dr. Joan L. Heffer
Graduate Group in Science and Mathematics Education
c/o School of Education
University of California
Berkeley, CA 94720

Private Sector

11. Melissa Holland
American Institutes for Research
1033 Thomas Jefferson St. NW
Washington, DC 20536

12. Dr. Marcel Just
Department of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

13. Dr. David Kieras
Program in Technical Communication
College of Engineering
1225 E. Engineering Building
University of Michigan
Ann Arbor, MI 48109

14. Dr. Walter Kintsch
Department of Psychology
University of Colorado
Boulder, CO 80302

15. Dr. Alan Leaad
Learning & Development Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

16. Dr. Jay McClelland
Department of Psychology
MIT
Cambridge, MA 02139

17. Dr. Allen Munro
Behavioral Technology Laboratories
1845 Elena Ave., Fourth Floor
Redondo Beach, CA 90277

18. Dr. Donald A Norman
Cognitive Science, E-015
Univ. of California, San Diego
La Jolla, CA 92039

19. Dr. Nancy Pennington
University of Chicago
Graduate School of Business
1101 E. 58th St.
Chicago, IL 60637