TIP-EXE: A software tool for studying the use and understanding of procedural documents
Franck Ganier, Ronan Querrec

To cite this version:
Franck Ganier, Ronan Querrec. TIP-EXE: A software tool for studying the use and understanding of procedural documents. IEEE Transactions on Professional Communication, 2012, 55 (2), pp.106-121. hal-00734005

HAL Id: hal-00734005
https://hal.univ-brest.fr/hal-00734005
Submitted on 20 Sep 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
TIP-EXE: A software tool for studying the use and understanding of procedural documents

Franck GANIER & Ronan QUERREC

Franck GANIER
Lab-STICC, UMR 6285 CNRS
Université Européenne de Bretagne,
Université de Brest,
Centre Européen de Réalité Virtuelle
25 rue Claude Chappe
29280 Plouzané – France
+33 (0)2.98.05.89.70
franck.ganier@univ-brest.fr

Ronan QUERREC
Lab-STICC, UMR 6285 CNRS
Université Européenne de Bretagne,
École Nationale d'Ingénieurs de Brest,
Centre Européen de Réalité Virtuelle
25 rue Claude Chappe
29280 Plouzané - FRANCE
+33 (0)2.98.05.89.50
ronan.querrec@enib.fr

Suggested running head:
TIP-EXE: Technical Instructions Processing – Evaluations and eXperiments Editor
TIP-EXE: A software tool
for studying the use and understanding
of procedural documents

Abstract – Research problem: When dealing with procedural documents, individuals sometimes encounter comprehension problems due to poor information design. Researchers studying the use and understanding of procedural documents, as well as technical writers charged with the design of these documents, or usability specialists evaluating their quality, would all benefit from tools allowing them to collect real-time data concerning user behavior in user-centered studies. With this in mind, the generic software TIP-EXE (Technical Instructions Processing – Evaluations and eXperiments Editor) was designed to facilitate the carrying out of such studies. Research questions: Does document design, and specifically the matching or mismatching of the terms employed in a user manual and on the corresponding device, affect the cognitive processes involved in the comprehension of procedural instructions? Can we use a software tool like TIP-EXE to assess the impact of document design on the use and understanding of a procedural document? Literature Review: A review of the methods employed to study either the use of procedural documents or their cognitive processing, and to evaluate the quality of these documents, revealed the lack of tools for collecting relevant data. Methodology: The TIP-EXE software was used to set up and run a laboratory experiment designed to collect data concerning the effect of document design on the performance of a task. The experiment was conducted with 36 participants carrying out tasks involving the programming of a digital timer under one of three conditions: “matching instructions”, “mismatching instructions”, “mismatching instructions + picture”. Based on a click-and-read method for blurred text, TIP-EXE was used to collect data on the time the users spent reading the instructions, as well as the time spent handling the timer. Results and Discussion: Results show that “matching instructions” (when the terms employed in the user manual match the terms on the device) enhance user performance. This instructional format results in less time spent consulting the instructions and handling the device, as well as fewer errors. This research shows that the TIP-EXE software can be used to study the way in which operating instructions are read, and the time spent consulting specific information contained therein, thereby revealing the effects of document design on user behavior.
Index Terms - procedural documents, real-time data, usability testing methods.

Franck Ganier is a professor of Cognitive Psychology and Ergonomics at the European University of Brittany, University of Brest, France. He is a member of the Lab-STICC Laboratory (UMR 6285 CNRS), European Centre for Virtual Reality (CERV), Brest, France. His research interests concern procedural learning, processes for comprehending instructions and methods for evaluating technical documents.

Ronan Querrec is an assistant professor in Computer Science at the European University of Brittany, École Nationale d'Ingénieurs de Brest (ENIB), France. He is a member of the Lab-STICC Laboratory (UMR 6285 CNRS), European Centre for Virtual Reality (CERV), Brest, France. His research interests concern virtual reality and knowledge explicitation in virtual environments by means of metamodeling. His work is applied to virtual environments for training purposes.
INTRODUCTION

Technical writers charged with the design of procedural documents have to choose which information content to give the user (for example, by determining a hierarchical order for presenting information, the level of detail when describing the device to be handled and the actions to be carried out, which terminology to use, etc.). They must also decide which format is best suited to convey this information (text, pictures, tables, graphics, etc.). Finally, they have to determine the layout of the document in order to help the user locate relevant information. These decisions rarely take into account the effects they have on the users’ understanding of the information [1].

Indeed, technical writers rarely receive feedback on how users process the information they have designed [1]. The underlying risk, therefore, is that poor information design can result in poor processing by the user. Given the difficulties in use or understanding that can arise from poor information design [2], the information content of procedural documents should be tested in order to assess how users are affected by it. Such an evaluation could subsequently lead to revising the organization, presentation and wording of this information [3], [4]. However, manufacturers who systematically use or provide this type of documentation with devices or appliances generally have no experience in the development of studies or techniques designed to improve the quality, i.e. the intelligibility and effectiveness, of procedural documents [1].

Moreover, previous studies have shown that the use of procedural documents (instructions for use, on-line help systems, recipes, etc.) frequently entails difficulties in the reading and carrying out of the instructions contained therein. Yet the role of the documentation provided with products aimed at the general public is to help users set up, operate or maintain the equipment or devices concerned. Some authors have put forward the hypothesis that, in some cases, these failures could be due either to the poor design of the information conveyed in these documents, or to difficulties encountered by users when processing the procedural information contained therein [5]-[8].

The following situation is an example of problems encountered by individuals when dealing with instructions for use. Sometimes, especially when users come across new devices, the terms given in the user manual do not allow them to see exactly what has to be handled on the
device itself. When the terms used in the manual do not clearly match the labels on the device, the user attempts to replace the missing information with his or her own knowledge. To do so, however, the user has to make inferences. This may be a more or less costly activity in terms of cognitive resources, which thus creates a more or less heavy “cognitive load” [2], [10].

Indeed, given the purpose of procedural documents (i.e. to help users successfully carry out a set of tasks), two things become obvious. Firstly, the cognitive processes involved in the understanding of the instructions given in such documents need to be studied in order to shed light on possible sources of difficulties for users. Secondly, methods for evaluating the usability and effectiveness of these documents need to be developed. If we consider the problem described above, it is clear that researchers studying the reading, understanding and carrying out of procedural instructions, as well as technical writers charged with the design of these documents, or usability specialists evaluating the quality of procedural documents, would all benefit from tools allowing them to collect real-time data in user-centered studies, thereby revealing user behaviors – and the difficulties encountered – when dealing with procedural documents.

However, tools offering the double perspective of studying the cognitive processes involved in the understanding of procedural documents as well as evaluating the usability and effectiveness of such documents are lacking, both in the research and the professional fields [9]. The TIP-EXE (Technical Instructions Processing – Evaluations and eXperiments Editor) software was designed to fill this gap by facilitating the design and carrying out of experiments into the use and understanding of instructions displayed in complex documents (i.e. composed of several pages and combining different information presentation formats: text, pictures, tables, graphics, etc.).

These considerations raise the following research questions: (a) Does document design, and specifically the matching or mismatching of the terms employed in a user manual and on the corresponding device, affect the cognitive processes involved in the comprehension of procedural instructions? For example, is the cognitive load caused by the mismatch of the terms employed in the user manual and on the corresponding device higher than when the terms match? (b) Can we use a software tool like TIP-EXE to assess the impact of document design on the use and understanding of a procedural document?
To answer these questions, this paper proposes to (a) discuss the difficulty and complexity of conducting research on (and evaluation of) user reading and comprehension of procedural documents, (b) introduce the TIP-EXE software, and (c) report results of an experimental study using this software. To do so, the paper is organized into the following four sections: Literature review; Methodology for conducting a sample study with TIP-EXE; Results; Conclusions, limitations, and suggestions for future research.

**LITERATURE REVIEW**

The aim of this section is twofold. Firstly, it reviews the different user-centered research methods for studying the use and understanding of procedural instructions and evaluating the quality of procedural documents. Secondly, it presents the TIP-EXE software, designed to set up and carry out experiments (or evaluations), and to collect behavioral data on the use of procedural instructions. To do so, the literature review is organized as follows: Methods for Studying the Use of Procedural Documents; Methods for Studying the Understanding of Procedural Documents; Conclusions about the Methods for Studying the Use and Understanding of Procedural Documents; TIP-EXE: Evaluations and Experiments Editor for Procedural Documents; Ergonomic Evaluation of TIP-EXE.

The selected literature included in this review is based on a literature review by Ganier [1], who presented a panoramic view of the methods employed in testing the usability of procedural documents. This review has been regularly updated with new literature in the fields of the use and understanding of procedural documents (with research topics such as the comprehension of instructions, procedural text, technical manuals, usability testing, etc.).

Analysis of the literature in this field shows that the use of procedural documents can be studied at both global and local levels. At a global level, studies try to determine the general behaviors involved in the use of these documents through the observation and analysis of the behavioral activity of the user. At a local level, studies focus on the cognitive processing of the instructions given in these documents through rigorously controlled experiments. This latter approach examines the cognitive processes involved in the reading and understanding of procedural instructions, as well as in the carrying out of the corresponding actions.
While these two approaches (at both global and local levels) employ different means to collect and analyze data concerning behavioral patterns, they are both based on common theoretical foundations. Firstly, analysis of the user’s behavior helps determine the internal mental processes involved. Indicators such as the number of times the user consults the different sources (document and device), the errors made while carrying out the actions, or the successful accomplishment of the task, help determine the user’s informational requirements (or gaps in his or her knowledge). Such indicators therefore help reveal when and how the document is used, and whether it facilitates or hinders understanding [1].

Secondly, analysis of the time spent processing information is based on the premise that the greater or more complex the cognitive processes involved, the more time is required to carry out a task. Depending on the aims of the study, the on-line measurement of the time required to consult the different sources, to carry out the actions, or to complete the task, should reflect either the difficulties in understanding (or in planning the actions, etc.) encountered by the user, or the problems inherent in the design of the documents, which would be indicated by lower performances. Indeed, several authors suggest that inappropriate procedural instructions can induce an extraneous cognitive workload when the user has to process unnecessary information [2], [10], [11].

Thus, while poorly designed documents would tend to cause cognitive overload, characterized by an increase in the time spent processing information, well-designed documents would tend to facilitate understanding, proportionately reducing the time spent processing information. The following two sections describe the different methods for studying the use of procedural documents at a global level, and the understanding of procedural documents at a local level.

**Methods for Studying the Use of Procedural Documents.** Video recording is a commonly employed technique for observing how procedural documents are used. Analysis of the way in which individuals carry out a specific task is based on the behavioral patterns collected. This approach usually studies the number of times the document and the device are consulted, the length of time spent inspecting them, the number of actions carried out by the user and the time spent on each action, the organization of these different activities, and the procedures followed by the user compared with those predicted [12].
These data can be collected under controlled experimental conditions depending on the aims of the study and the hypotheses put forward by the researcher. In ecological settings, the user is required to carry out a task without receiving any prior information on how to use the instructions [13]. In more rigorously controlled experiments, the user is required to consult a booklet that contains instructions (with a different task to perform per page), and to scrupulously follow these instructions. In this case, the researcher provides no information with regard to the task. The participants in the study must discover the nature of the task by following the instructions [14].

**Methods for Studying the Understanding of Procedural Documents.** Some studies focus more specifically on the cognitive processes involved in the reading and carrying out of instructions given in procedural documents. In these studies, indicators (cues) that identify the processes involved per task can be collected either while the task is being carried out (on-line measures), or after a certain delay (off-line measures). These two approaches are described in the following paragraphs.

*On-line measures.* Different techniques are generally used in the on-line study of the processes of reading, understanding and carrying out procedural instructions. As previously mentioned, video recordings have often been used in experimental settings. They prove, however, to have certain limitations, notably where obtaining accurate on-line measures is concerned. To solve this problem, new techniques for collecting and more carefully analyzing behavioral data were developed.

In the mid-1970s, researchers in psycholinguistics developed techniques for studying the cognitive processes involved during reading. Amongst these, Self-Paced Reading techniques allowed the reader to display different segments of a text on a computer screen [15], [16]. The reader moved through these segments at his or her own pace by pressing a key on the keyboard. This allowed the computer to measure the Reading Exposure Time for each segment. Some authors, such as Dixon, adapted this technique to study the understanding of procedural instructions [17]. Here, the user read the instructions on the computer using the Self-Paced technique and carried out the task in a real-life environment (for example, by using a pencil and paper for a drawing task, or by handling a real appliance or other device designed for the purposes of the experiment).
Several other methods subsequently emerged that were derived from this technique. These have evolved towards a computer simulation of the device to be handled in addition to on-screen instructions [18]. In some cases, the on-screen instructions are encrypted or blurred; the user must click on them in order to read them. This allows for measuring the duration of display of the instructions, as with the click-and-read method proposed by Ummelen [19]. In other cases, the screen displays a working zone, a window that hides the instructions and in which the user can carry out the described actions. The computer alternately measures the duration of display of both the instructions and the working zone [20].

Moreover, the recording of ocular movements, a technique commonly used to study the reading of narrative or descriptive texts, has recently been employed in studies into the reading and understanding of procedural instructions [21], [22]. This technique involves measuring the ocular fixations and saccadic eye movements of users as they read instructions. Although it allows researchers to collect very precise data, eye-tracking systems are expensive, require operational expertise, and involve analyzing substantial quantities of data.

Whatever the technique, the same principle applies: determining the time the user requires to process (i.e. read and understand) all or part of the instructions, which are presented in different ways (text, pictures, graphics, etc.). This allows researchers to collect more detailed behavioral data and to determine precisely the nature, time and length of the processes involved. In addition to behavioral indicators related to reading procedural instructions, the actions carried out by the user also provide important evidence regarding how his or her cognitive processes operate. For example, based on the actions performed by the user immediately after reading the instructions, the researcher can make inferences regarding the user’s representation of the information processed, thereby enabling the researcher to follow the understanding processes step by step.

*Off-line measures.* Off-line measures of the reading and understanding of procedural instructions involve collecting data at the end of the reading-and-execution phase (immediately afterwards, or after a shorter or longer period of time). For example, data can be collected regarding how a task was carried out, in terms of success or failure, or the number of definite errors made by the user. The user can also be asked to verbally recall the instructions
he or she previously read, or to carry out the task again from memory. This allows researchers to obtain information on how procedures are memorized [23], [24].

**Conclusions about the Methods for Studying the Use and Understanding of Procedural Documents.** It is clear from the above that certain methods allow researchers to study how procedural documents are used and understood. However, these methods are sometimes difficult to implement (eye-tracking techniques) or require spending a substantial amount of time analyzing the data collected (video recordings). Moreover, the tools and equipment necessary for collecting more precise data, allowing more detailed analyses of such experiments, are generally developed as and when the need arises, in response to the specific requirements of researchers and usability specialists. A case in point is the programs used by Dixon [17], Duggan and Payne [18], Ummelen [19], and Ganier [20].

However, in these situations, the extent of the research is sometimes limited by the fact that the documents concerned often lack ecological validity. Indeed, in most cases, the instructions comprise only a few lines, presented in a text format as opposed to a format that uses both text and pictures. Moreover, the documents studied are limited to a single page. No generic tool exists for setting up experiments or observational studies into the use and understanding of complex documents, i.e. documents comprising both text and pictures, and composed of several pages. Moreover, and no doubt as a direct consequence, there is little, if any, transfer of these methods or developed tools to experiments in which the quality of procedural documents and the effectiveness of the instructions contained therein can be evaluated. Some generic software tools have been developed to evaluate documents, for example the DocuScope [25], but these are more text-centered than user-centered, and seem to address texts of a more narrative or expository nature. To date, they still need to be tested on procedural documents. The TIP-EXE generic software was therefore developed in order to facilitate user-centered evaluations.

**TIP-EXE: Evaluations and Experiments Editor for Procedural Documents.** The TIP-EXE (Technical Instructions Processing – Evaluations and eXperiments Editor) software is designed both for studies into how procedural documents are used and understood, and for the
ergonomic evaluation of their content. To this end, TIP-EXE is distributed free of charge\textsuperscript{1} to researchers studying the use and understanding of procedural documents, as well as to technical writers or usability specialists evaluating the quality of these documents.

TIP-EXE can be used to study how procedural instructions are read, and the time spent consulting information contained therein. To this end, TIP-EXE registers the reading path followed by the user, as well as the time he or she spends reading or checking the instructions. This allows researchers to compare either several sets of instructions that contain the same information presented in different material formats (e.g. one set of instructions given as text only and the other comprising both text and pictures), or the effect of information presentation on the user’s reading path (e.g. how the layout of a page influences the way in which it is read). The software displays an on-screen example of printed operating instructions, certain zones of which, specified \textit{a priori} by the researcher or the ‘evaluator’ (a usability specialist, for example), are blurred. By clicking on these zones in order to read them, the user activates a timer that registers his or her reading path and the time spent consulting each zone.

The following stages are involved when using TIP-EXE: firstly, preparing the experiment or the evaluation of the document (‘New experiment’); secondly, carrying out the experiment and collecting the data (‘Run experiment’); and finally, exporting and analyzing this data (‘Export data’).

\textit{New experiment}. This stage consists of specifying the document(s) to be used, associating one or several pages to the document(s), deciding which zones in these pages will be visible, hidden, or blurred for the users, and specifying the task(s) to be carried out using the document(s). Creating a new experiment in TIP-EXE involves opening a parent window that contains three child windows: ‘Edit document’, ‘Edit task’, and ‘Edit page’. Figure 1 presents this display. The ‘Edit document’ window is used to create new documents, or to open previously created documents. The ‘Edit task’ window is used to specify the task(s) that the participants in the experiment will be required to carry out according to the instructions received. Finally, the ‘Edit page’ window is used to edit the page(s) associated with each document. This consists of specifying all of the pages that will make up a document and deciding which zones will be visible, hidden or blurred.

\textsuperscript{1} TIP-EXE for Windows can be downloaded at: \url{http://www.enib.fr/~querrec/Tip-exe.zip}
Deciding which zones in the document will be visible, hidden or blurred during the experiment represents the most important aspect of this stage. These zones can be a word, a line, a paragraph or a picture (only rectangular zones are allowed in this current version of TIP-EXE). A zone is created by clicking directly on the newly inserted page, then moving the mouse while holding down the button. An orange zone (blurred) will be created and the name ‘Zone 1’ will automatically appear. Figure 2 shows the creation of blurred zones in TIP-EXE. Each zone can be renamed, resized and moved.

Once a new document has been created, a set of associated tasks can be specified. These tasks are the instructions that will be given to the participants. The researcher or evaluator decides whether to display the instructions throughout the task or to make them disappear after a certain length of time. In the latter case, the participant can click on a button in order to display the instructions again for the same duration.

Run experiment. This stage involves, firstly, filling in all the information concerning the experimental conditions and the participants’ profiles and, secondly, the actual carrying out of the experiment. A window at the start of the experiment displays the experimental conditions.
and the participant’s profile (i.e. participant number, sex, age, mother tongue, education level, dominant hand, eyesight, miscellaneous information). Validating these data creates a *.LOG file with the name of the experiment followed by the participant’s identification number. The experiment consists of giving the participants one instruction per task, and providing them with the specific document that will help them perform each task. Figure 3 presents the windows seen by the participants in the experiment. Some zones in the document are visible, some are blurred, and others are hidden. The participants are asked to carry out one or several tasks consecutively, using one document or more, until a message appears to mark the end of the experiment.

**INSERT FIGURE 3**

*Figure 3. Windows seen by the participants in the experiment.*
On the left, the instruction describing the task to be performed.
On the right, a page of the document. Some zones are visible (the titles), some are blurred (the instructional paragraphs), and others are hidden (a paragraph, to the bottom-left of the document).

**Export data.** TIP-EXE allows for exporting the data collected from the experiments in order to analyze the results. The results file can be imported into a spreadsheet. This results file contains one ‘Experiment’ sheet, with the results obtained for all of the participants, and a number of other sheets, one per participant, with their individual results. Figure 4 shows the ‘Experiment’ sheet, that presents the results obtained for all of the participants (Part. 1, Part. 2, etc.) and for each task: the time spent consulting the document (Reading Exposure Time), the time spent carrying out the instructions by handling the device (Execution Time), the time taken to complete the task (Total Time), and the time spent consulting each of the zones (Zone 1, Zone 2, etc.).

**INSERT FIGURE 4**
Figure 4. The ‘Experiment’ sheet (results for all participants).

Figure 5 shows that the individual participant sheet is divided into 4 zones: ‘Information’, ‘Reading’, ‘Execution’ and ‘Statistics’.

**INSERT FIGURE 5**

Figure 5. Individual results sheet.

The ‘Information’ zone contains the participant’s personal information and identification number, as well as the conditions under which the experiment was carried out. The ‘Reading’ and ‘Execution’ zones are divided into three columns: the first indicates the time registered at the start of the event; the second, the time registered at the end of the event; and the third, the duration of the event. These times are measured in milliseconds. This measurement scale has been determined by the fact that the smallest zone that can be specified corresponds in size to one word. Depending on the aims of the experiment, times will probably have to be changed to either seconds or minutes before running statistical analysis. This can be easily done using the spreadsheet.

In the example shown in Figure 5, the participant began the experiment by consulting the “task to carry out” window and handling the device for 3,420 ms (‘Execution’: from Start time = 0 to End time = 3420), then consulted Zone 1 of the document for 3,430 ms (‘Reading’, Zone 1: from Start time = 3420 to End time = 6850). He then spent another 1,024 ms handling the device (‘Execution’: from Start time = 6850 to End time = 7874), before consulting Zone 1 again, etc. The ‘Statistics’ zone is divided into three columns: the first indicates the duration of the event; the second, the length of this event with regard to the total time spent carrying out the task; and the third, the number of occurrences for the consultation of each zone. In the example given in Figure 5, the participant spent 19,771 ms consulting the document (42.35 % of the total time to complete the task), and 26,912 ms
handling the device (57.65 % of the total time to complete the task). As regards the time spent consulting the document, the participant spent 6,038 ms on Zone 1 (30.54 % of the Reading Exposure Time), 7,878 ms on Zone 2 (39.85 % of the Reading Exposure Time), and 5,855 ms on Zone 3 (29.61 % of the Reading Exposure Time).

**Ergonomic Evaluation of TIP-EXE.** An ergonomic evaluation of TIP-EXE was carried out in order to assess the usability of the software. This evaluation involved the first operational version of TIP-EXE (version 0.9c) being tested by users likely to be interested in this software and who had a good working knowledge of such applications (four second-year students on the DEUST T-MIC\(^2\) course at the European University of Brittany, Brest, France).

The participants assumed the role of usability specialists asked to design an experiment using TIP-EXE. They had at their disposal the software user manual and a booklet containing screen captures of the different TIP-EXE windows. A method adapted from the Plus-Minus Method ([26], [27]) was used to evaluate the software. When the participants encountered a difficulty, they had to describe the nature of the problem (use or operation, window organization, terminology or symbols used, etc.) and to point it out on the corresponding screen capture. They could also recommend any information or other elements they considered lacking and which, in their opinion, would have enhanced the usability (i.e. ease of use and learnability) of TIP-EXE.

The feedback from this first evaluation was used to modify the software, the new version of which (version 0.9d) was then evaluated by other users (nine first-year students on the DEUST T-MIC course). The comparative evaluation of both versions highlighted an improvement in user performances, a decrease in the number of difficulties encountered, and an increase in user satisfaction. Since then, TIP-EXE has been used for the training of Technical Writers in Master degrees curricula at the Universities of Paris 7, Brest and Vannes (France), which has allowed for continuously improving the usability of this software. The TIP-EXE version presented in this article is version 0.9n.

**METHODOLOGY FOR CONDUCTING A SAMPLE STUDY WITH TIP-EXE**

\(^2\) 2-year “Technician in Interactive Media” scientific degree.
In order to test TIP-EXE in a real laboratory setting, an experimental study was set up using this software, from the preparation of the material to the analysis of the results. The purpose of this section is to present the aims of the study, the choice of research methodology, and how the study was conducted. This methodology could be extended to other studies involving the use of TIP-EXE.

**Aims of the Study.** A procedural document should allow the user to correctly handle the associated device, without any hesitation or difficulty on his or her part. However, instructional material presented in documents of this kind often fails to help the user (a) elaborate an appropriate representation of the actions to be carried out, or (b) precisely locate where these actions must be carried out on the device. This is notably the case when the device is unfamiliar (e.g. an innovative product) and when the terms used in the document do not allow for precisely identifying the elements to be handled on the device [6].

This type of situation recalls Glenberg and Robertson's indexical hypothesis [28], which we shall refer to as the *indexing process* in the present paper. These authors suggest that the interpretation of a text consists in *indexing* (i.e. putting into correspondence) three sources of information: the written text, the information coming from the situation (in this case, the device) and the user's prior knowledge. Thus, when the technical terms used in the instructions do not exactly correspond to the specific parts of the device to be handled, the user has to draw certain conclusions. To this end, he or she attempts to piece together the missing information from his or her interactions with the device, from information given in the document, and from his or her prior knowledge.

This inferential process can be difficult when part of this information is missing or incorrect, thereby influencing the mental efforts involved and the cognitive load to a greater or lesser extent (e.g. when the user is a novice). For example, Ganier & Pétillon [29] carried out a study using a programmable digital timer. Although marketed in France, the timer nevertheless had all its buttons labeled in English, while in the enclosed procedural document the terms designating the buttons were translated into French. Thus, the button *hour* on the timer was translated as *heure* in the document, *clock* became *horloge*, *week* was incorrectly translated as *jour* (*day*), and so on.
When confronted with such a situation, the user has to make a certain number of inferences in order to establish a correspondence between the terms he or she reads in the document and the elements he or she has to handle on the device. In such a case, facilitating the inferential process would ensure that the user develop coherent mental representations from the available documents [30]. Two different solutions facilitate the inferential process. The first, and easiest, solution would be to design the text and/or device in such a way as to facilitate the correspondences made by the user. Where this is not possible, for example when “mismatching” instructions for use have already been printed in a great number of copies, a second solution would be to include pictures of the device on a supplementary sheet of paper. Because they illustrate the content of the text and constitute an analogical representation of the real world, pictures help the user identify and interact with appropriate parts of the device and therefore facilitate the understanding of the information contained in the text [31].

An experiment was therefore designed using the TIP-EXE software to investigate how these two ways could be helpful to the user. This experiment aimed to address the following research questions: (a) Does document design, and specifically the matching or mismatching of the terms employed in a user manual and on the corresponding device, affect the cognitive processes involved in the comprehension of procedural instructions? (b) In the case of a poorly-designed document (with mismatching terms), how does a picture of the device help the user match terms used in the document with parts of the device? (c) Can we use a software tool like TIP-EXE to assess the impact of document design on the use and understanding of a procedural document? To these ends, using TIP-EXE, we tried to evaluate the cognitive load induced from the processing of instructions contained in a procedural document in two situations: (a) when the terms printed in the document did or did not correspond with the terms marked on the device to be handled (a programmable digital timer); (b) when a picture was added in the case of mismatching terms.

**Choice of Research Methodology.** To address the above research questions, we used an experimental methodology, with quantitative data collecting. Amongst other methods, this allows for highlighting causality effects between the different versions of the documents presented to the participants and the cognitive processes they induced. We also needed a tool that would measure these cognitive processes in a precise way. That is why, based on Ganier's model [32] identifying the processes implied in the use and comprehension of procedural documents, we used TIP-EXE. This software allows for separately and precisely measuring
different dependent variables corresponding to these cognitive processes, in a way that respects the natural behavior of users interacting with a procedural document and a device (i.e. interleaving between reading the instructions and carrying out the actions [18]).

How the study was conducted

Participants. Thirty-six undergraduate students (32 women and 4 men) of the University of Brest volunteered to take part in this experiment. With a mean age of 21, all were French native speakers, and had no prior knowledge about the device. In France, where this type of study is not considered as biomedical research, there was no need to request and obtain approval from a Research Ethics committee. The study nevertheless strictly adhered to all ethical principles where the conducting of research with human participants is concerned (freedom to leave the experiment at any time, information about the nature of the study at the end of the experiment, confidentiality of the information obtained about the participants, etc.).

Procedure. The whole experiment was carried out using the TIP-EXE software. The participants were shown how to click on a blurred text to make it readable. They could then practice on a specific text created for this training stage. Once the participants were familiar with the use of TIP-EXE, they carried out the experiment, programming the digital timer with the help of the instructions presented in the TIP-EXE software. All participants were tested individually. They were randomly assigned to one of the three experimental conditions (see below).

During the experiment, the participants had to carry out three tasks with the programmable digital timer:

1. Adjust the day of the week to *Mercredi* (Wednesday) and the hour to 22:57.
2. Adjust the timer so that it started and stopped: a) every *Samedi* (Saturday) and *Dimanche* (Sunday), at 18:15 and 23:30; and b) every *Mardi* (Tuesday), at 20:00 and 22:00.
3. Delete the *Samedi* (Saturday) and *Dimanche* (Sunday) programming and reset the timer to start every *Vendredi* (Friday) at 20:45 and stop at 23:15.

In order to carry out these tasks, the participants had to follow the related sets of instructions extracted from the standard documentation provided with the device. The instructions were displayed blurred to the users, who had to click on the zones that they wished to consult in order to visualize them correctly.
Material. The device used for this experiment was the programmable digital timer shown in Figure 6. It is important to note that all the buttons on this timer were labeled in English.

INSERT FIGURE 6

Figure 6. The programmable digital timer used for the experiment.

Three experimental conditions were designed for the instructional material. In the first condition (“mismatching instructions”), all of the instructions were in French, including the terms designating the timer buttons. This condition corresponded to the procedural instructions distributed with the device by the manufacturer and was supposed to make the indexing process harder. In the second condition (“matching instructions”), the instructions were in French, but the technical terms designating the timer buttons were labeled in English in the text. This condition was hypothesized to allow the participants to easily index the technical terms with the corresponding parts of the timer. In the third condition (“mismatching instructions + picture”), while the instructions were in French, including the terms designating the timer buttons, they also included a picture of the timer in which the buttons were labeled in French. This condition is shown Figure 7, and was supposed to make the indexing process easier, but in a less efficient way than the “matching instructions” condition (due to the supplementary informational material to process). In each of these experimental conditions, the instructional material was displayed on a computer screen, using TIP-EXE.

INSERT FIGURE 7

Figure 7. The document used in the “mismatching instructions + picture” condition. In the picture of the digital timer, the terms designating the buttons are in French (but labeled in English on the real digital timer).
The data collected. The software registered different kinds of data corresponding to dependent variables. These dependent variables were the total time spent consulting the predefined zones corresponding to the instructions (Reading Exposure Time), the time spent carrying out the instructions by programming the timer (Execution Time), the time taken to carry out the whole procedure, including the three tasks (Total Time to Perform the Tasks), and the total number of errors (observed by the researcher at the end of each task).

Data analysis. To summarize the data collected from this experiment, we used descriptive statistics such as mean and standard deviation from the mean. To ensure the validity and generalizability of the data, the significance of the results was assessed using analysis of variance (ANOVA). As we had three experimental conditions with different participants, a one-way unrelated ANOVA was chosen. Inasmuch as ANOVA only allows for revealing a main effect, that is to say that there are some differences in means (and variances, equal to standard deviation squared) for the three conditions at a general level, this analysis was completed with analytic comparisons (contrasts), that involve comparing two means at a time or one mean with two cumulated means.

RESULTS

The purpose of this section is to address the research question concerning the degree to which matching instructions facilitate the indexing process and assist user comprehension and the completion of tasks given a procedural document. This section therefore provides the quantitative results and subsequent data analysis. It starts with a presentation of the descriptive statistics, presents the results of the ANOVA tests and the analytic comparisons, and then answers the research question.

The results of the descriptive statistics for all the dependant variables (Reading Exposure Time, Execution Time, Total Time to perform the tasks, and Number of errors) are shown in Table I.

Table I. Chronometric measurements (in seconds) and number of errors for each of the different ways in which
The instructions were presented (Standard Deviation between brackets).

INSERT TABLE I

The results of the ANOVA tests show that difference in types of instructions (“matching”, “mismatching”, “mismatching + picture”) had a significant effect on all of the dependent variables that were analyzed:

- Reading Exposure Time of the instructions: $F(2,33) = 5.37; p < .01$.
- Execution Time for handling the timer: $F(2,33) = 5.12; p < .03$.
- Total Time taken to perform the tasks: $F(2,33) = 6.61; p < .01$.
- Number of errors: $F(2,33) = 3.79; p < .05$.

With regard to the Reading Exposure Time, analytic comparisons (contrasts) show a significant difference between the “matching instructions” condition and the two other conditions, either taken alone or cumulated. Thus, as shown in Table I, Reading Exposure Time is shorter for the “matching instructions” condition ($M = 141.23$ s; $SD = 44.91$), than for the “mismatching instructions” condition ($M = 204.16$ s; $SD = 84.71$): $F(1.33) = 3.95; p < .05$; or the “mismatching instructions + picture” condition ($M = 243.95$ s; $SD = 93.87$): $F(1.33) = 10.55; p < .01$. Results also indicate that Reading Exposure Time is shorter for the “matching instructions” condition than for the two other cumulated conditions: $F(1.33) = 9.14; p < .01$. Furthermore, Reading Exposure Time is longer for the “mismatching instructions + picture” condition than for the two other cumulated conditions: $F(1.33) = 6.77; p < .03$.

As regards the Execution Time, analytic comparisons show that it is shorter for the “matching instructions” condition ($M = 296.05$ s; $SD = 92.59$) than for the “mismatching instructions” condition ($M = 432.41$ s; $SD = 141.94$): $F(1.33) = 5.34; p < .03$; or for the “mismatching instructions + picture” condition ($M = 477.33$ s; $SD = 184.10$): $F(1.33) = 9.45; p < .01$. Analytic comparisons also indicate that Execution Time is shorter for the “matching instructions” condition than for the two other conditions when cumulated: $F(1.33) = 9.67; p < .01$. Furthermore, Execution Time is longer for the “mismatching instructions + picture” condition than for the two other cumulated conditions: $F(1.33) = 4.90; p < .03$.  

21
Total Time to perform the tasks is shorter for the “matching instructions” condition ($M = 437.28$ s; $SD = 126.30$), than for the “mismatching instructions” condition ($M = 636.57$ s; $SD = 201.25$): $F(1.33) = 6.18; p < .03$; or for the “mismatching instructions + picture” condition ($M = 721.29$ s; $SD = 243.43$): $F(1.33) = 12.55; p < .01$. Total Time to perform the tasks is also shorter for the “matching instructions” condition than for the two other cumulated conditions: $F(1.33) = 12.11; p < .01$. Furthermore, Total Time to perform the tasks is longer for the “mismatching instructions + picture” condition than for the two other cumulated conditions: $F(1.33) = 7.05; p < .03$.

Turning to the Number of errors, analytic comparisons show a lower number of errors for the “matching instruction” condition ($M = 0.08; SD = 0.29$) than for the “mismatching instructions + picture” condition ($M = 0.58; SD = 0.67$): $F(1.33) = 6.60; p < .03$. The number of errors is also lower for the “mismatching instructions” condition ($M = 0.16; SD = 0.39$) than for the “mismatching instructions + picture” condition: $F(1.33) = 4.58; p < .05$. These analytic comparisons also indicate that the number of errors is higher for the “mismatching instructions + picture” condition than for the two other cumulated conditions: $F(1.33) = 7.39; p < .01$.

Together, these results show a superiority effect for the “matching instructions” condition. The total time required to perform the different tasks (including Reading Exposure Time and Execution Time) was shortest, and the number of errors was lowest for the “matching instructions” condition, i.e. when the terms used in the instructions corresponded to the terms marked on the timer (clock in the document corresponded to clock on the timer). Thus, in this condition, the Reading Exposure Time of the instructions, as well as the time required to manipulate the timer are shorter than for the “mismatching” and “mismatching instructions + picture” conditions (Table 1).

Besides the superiority effect for the “matching instructions” condition, the results indicate that, firstly, in the particular case of adding a picture when terms mismatch, the presence of a picture in the procedural document does not seem to facilitate the indexing process inasmuch as it leads to a higher number of consultations and a greater processing time than for the “matching instructions” condition. These results could be explained by the fact that consulting the picture increased the time required to read the document as well as the total activity induced from interleaving the consultation of the instructions, the consultation of the picture,
and the handling of the timer. Secondly, as regards the “mismatching instructions” and “mismatching instructions + picture” conditions, some of the users probably abandoned the consultation of the instructions to the benefit of programming the timer by trial-and-error. However, the real device used in this experiment did not allow for precisely measuring the interactions of the user with the programmable digital timer. We will consider some possible solutions intended to enhance the studying of the user-document-device interaction in the following part of this paper.

**CONCLUSIONS, LIMITATIONS, AND SUGGESTIONS FOR FUTURE RESEARCH**

**Conclusions.** Based on the observation that individuals sometimes encounter difficulties using and understanding procedural documents, the purpose of this paper was to present TIP-EXE, a software tool designed to help researchers, technical writers and usability specialists set up user-centered experiments and evaluations by measuring reading and execution behaviors. A sample study was run with TIP-EXE in order to investigate how document design, and specifically the matching or mismatching of the terms employed in a user manual and on the corresponding device, affect the cognitive processes involved in the comprehension of procedural instructions. This experiment showed that matching instructions improve user performance by facilitating what we have referred to in this paper as the *indexing process*. Moreover, it showed that TIP-EXE could be a useful tool to run such studies, corresponding to both researchers’ as well as practitioners’ needs.

The design of TIP-EXE therefore attempted to address many different requirements. One requirement was for generic application software that could be used for research purposes as well as for the evaluation of procedural documents. This software had to allow for studying the use of documents composed of several pages, and sometimes combining both text and pictures. It had to be portable, fully autonomous and capable of being used in both occupational and laboratory settings. It had to allow fast and easy exploitation of collected data, exported in spreadsheet format. Finally, it had to be cheap and easy to use. The TIP-EXE software meets all of these requirements. Furthermore, the role and application of TIP-EXE can be extended beyond that of evaluating the use and understanding of procedural documents: it can, for example, be used to study reading paths when dealing with narrative texts, pictures, or web sites.
**Limitations.** The main limitation of the TIP-EXE software is that it does not allow for as natural a reading activity as that registered using eye-tracking systems, since the user must click on the zone that he or she wishes to consult. More research therefore needs to be carried out into the legitimacy of using TIP-EXE as opposed to eye-tracking systems. Moreover, the current version of TIP-EXE does not allow for accurately measuring the user's interactions with the device to be handled. Indeed, the time spent handling the device is obtained by subtracting the Reading Exposure Time from the total time required to carry out the task, both of which are measured using TIP-EXE. While this can be acceptable when evaluating the quality or effectiveness of documents provided with real devices, it proves to be somewhat limited when evaluating either the cognitive processes involved in the understanding of procedural instructions, or the user's sequence of actions following the reading of these instructions.

**Suggestions for future research.** Future evolutions of the TIP-EXE software may be expected to allow for collecting more accurate data concerning the moments at which the device is handled and their duration, thanks to two types of applications still under development. The first of these aims to allow TIP-EXE to collect all user actions on a real electronic device (e.g. a programmable digital timer) through an electronic interface including a microprocessor for identifying which buttons the user pressed, at what point and for how long. The second involves merging TIP-EXE with a computer simulation of the device (a 3D object that can be handled, [33]), using virtual reality. This latter solution offers multiple advantages and should interest both basic research and industrial applications.

Where basic research is concerned, beyond being able to measure both cognitive and behavioral processes, the use of a virtual 3D object together with TIP-EXE would allow for carrying out studies into document and product design. As regards industrial applications and practitioner research, it would make it possible to evaluate both procedural documents and the corresponding devices or equipment, in the course of their design process. Indeed, this solution would solve the availability problems of different prototypes when the product is still in its design stage. Consequently, it would make the evaluation of the product and its documentation much easier, thereby reducing costs relative to product design evolutions.
ACKNOWLEDGMENTS

The authors would like to thank Charlotte Hoareau, for carrying out the data collecting for the study, as well as Kathryn Smyth and Morgane Roger, for their generous and friendly help in the translation and the improvement of an earlier version of this manuscript. The authors would also like to thank the editor of the journal for his helpful guidance in the revision of this article.

REFERENCES


Figure 1. Creating an experiment: TIP-EXE parent window and three child windows.
Figure 2. Creation of blurred zones in TIP-EXE.
Figure 3. Windows seen by the participants in the experiment. On the left, the instruction describing the task to be performed. On the right, a page of the document. Some zones are visible (the titles), some are blurred (the instructional paragraphs), and others are hidden (a paragraph, to the bottom-left of the document).
Figure 4. The ‘Experiment’ sheet (results for all participants).
<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Age</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Dominant hand</td>
<td>Right</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Education level</td>
<td>Master 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Experimental condition</td>
<td>Trial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Eyesight</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Misc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Mother tongue</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Participant number</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Sex</td>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reading</td>
<td>Execution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Task</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Zone</td>
<td>Start time</td>
<td>End time</td>
<td>Duration</td>
<td>Start time</td>
<td>End time</td>
<td>Duration</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Zone1</td>
<td>3429</td>
<td>6968</td>
<td>3439</td>
<td>0</td>
<td>3429</td>
<td>3439</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Zone2</td>
<td>15180</td>
<td>16833</td>
<td>1643</td>
<td>10240</td>
<td>15186</td>
<td>4704</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Zone2</td>
<td>17611</td>
<td>20737</td>
<td>3126</td>
<td>16833</td>
<td>17616</td>
<td>796</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Zone2</td>
<td>20737</td>
<td>30169</td>
<td>3114</td>
<td>20737</td>
<td>20757</td>
<td>632</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Zone3</td>
<td>23141</td>
<td>26108</td>
<td>4667</td>
<td>23159</td>
<td>31411</td>
<td>1244</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Zone3</td>
<td>43912</td>
<td>43930</td>
<td>18</td>
<td>43914</td>
<td>43912</td>
<td>780</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>45090</td>
<td>46683</td>
<td>1693</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Reading exposure time</td>
<td>19771</td>
<td>4735%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Execution time</td>
<td>26912</td>
<td>57.65%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Total time</td>
<td>45883</td>
<td>100.00%</td>
<td>Occurrences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Time for Zone1</td>
<td>6938</td>
<td>30.54%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Time for Zone2</td>
<td>7975</td>
<td>33.65%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Time for Zone3</td>
<td>5885</td>
<td>29.81%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 5. Individual results sheet.*
Figure 6. The programmable digital timer used for the experiment.
Figure 7. The document used in the "mismatching instructions + picture" condition. In the picture of the digital timer, the terms designating the buttons are in French (but labeled in English on the real digital timer).
Table I. Chronometric measurements (in seconds) and number of errors for each of the different ways in which the instructions were presented (Standard Deviation between brackets).

<table>
<thead>
<tr>
<th></th>
<th>Matching instructions</th>
<th>Mismatching instructions</th>
<th>Mismatching instructions + picture</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reading Exposure Time</strong></td>
<td>141.23 s (44.91)</td>
<td>204.16 s (84.71)</td>
<td>243.95 s (93.87)</td>
</tr>
<tr>
<td><strong>Execution Time</strong></td>
<td>296.05 s (92.59)</td>
<td>432.41 s (141.94)</td>
<td>477.33 s (184.10)</td>
</tr>
<tr>
<td><strong>Total Time to perform the tasks</strong></td>
<td>437.28 s (126.30)</td>
<td>636.57 s (201.25)</td>
<td>721.29 s (243.43)</td>
</tr>
<tr>
<td><strong>Number of errors</strong></td>
<td>0.08 (0.29)</td>
<td>0.16 (0.39)</td>
<td>0.58 (0.67)</td>
</tr>
</tbody>
</table>