The research described here examines the problems encountered by people when filling in forms. Subjects were required to complete forms on the basis of a situation sketch, while thinking aloud. From the completed forms, the observations, and the subjects' comments, conclusions could be drawn about the types of problems the subjects had encountered and about the strategies they had used. These conclusions, together with various suggestions found in the literature, provided a guideline for a thorough revision of seven forms. A test showed that, after revision, the number of forms completed unacceptably was reduced by about half.

In 1989, a remarkable report was published by Van Oorschot and Kolkhuis Tanke, two researchers from Tilburg University in The Netherlands [1]. This report described the findings of an international study into the so-called underconsumption of social security benefits. Having examined a large number of reports published by governments and universities in Western European countries, Van Oorschot and Kolkhuis Tanke conclude that many citizens who have a right to social security benefits receive little, if any, of what they are entitled to. To give some examples: in Great Britain, 49 percent of those entitled to family income supplements do not actually receive them; in France, 22 percent of the people over age sixty-five wait more than five years before claiming their old age pension; and
in the Netherlands, 55 percent of those entitled to a rent rebate do not receive the full amount [1]. Sociologists trying to explain these underconsumption figures have designed and tested some explanatory models [2, 3]. In these models, the process leading to a social benefit is described as an obstacle course. Thomassen, for instance, identifies five major obstacles that Dutch citizens have to surmount in acquiring entitlements [3].

1. **Lack of knowledge of rights** — people must be aware that they come under a particular scheme and therefore have a right to the benefits conferred by it;
2. **Mental resistance** — possible feelings of shame, fears of officialdom, and concern about being stigmatized must be overcome;
3. **Claiming rights** — an application has to be submitted; usually this means completing a form;
4. **The procedure for dealing with the claim** — the claimant must continue to respond adequately even when the claim is being processed; sometimes extra information must be supplied, and a degree of patience will be required in almost all cases; and
5. **Administrative law procedures for the protection of rights** — if the decision is unfavorable, the claimant may decide to exercise the right of appeal, in which case the same process may have to be followed all over again.

According to Thomassen, each of these obstacles can, and often does, cause things to go wrong [3]. In our thesis, we examine the various forms of official communication that are intended to make it easier for individuals to surmount the first and third obstacles [4]. In this article, we shall confine ourselves to the third obstacle: the completion of forms in order to apply for one's entitlements.

Just how much Dutch government forms are criticized is evident, for example, from a survey conducted by Huismans and Siegerist in the early seventies [5]. Of the 543 respondents (a representative sample of the Dutch population) 75 percent considered that the government used too many forms, 81 percent felt that some questions on forms were so unclear that it was impossible to know what to fill in, and 80 percent agreed with the statement that forms sometimes contain questions which make you wonder why on earth they would want to know the answers to them [5]. Some ten years later, major problems were reported with the application form for individual rent rebates. Only 18 percent of the claimants completed the form themselves; the other claimants had to have it done for them by a civil servant, their landlord, or a social or voluntary worker. Fifty percent of the respondents stated that they found it impossible to fill in the form themselves [6].

In a survey conducted among the clients of a local Dutch Social Services Department, over 30 percent of the respondents indicated that the department’s forms contained too many questions, almost 50 percent considered that the questions were unclear, and the same percentage thought that some of the questions were entirely unnecessary. Examples cited included questions about job applications, questions repeatedly asking for the same personal particulars, and questions about
the household and family situation which appeared irrelevant to the benefit in question [7].

**RESEARCH INTO THE PROBLEMS OF COMPLETING FORMS**

On the subject of designing and testing forms, only a limited amount of literature is available. Important research in this field has been done by Wright [8]. She carried out a series of laboratory tests in which the effect of different types of questions and of different completion formats was examined. The subjects of the experiments were required, for example, to answer a few dozen questions of a particular type. Wright then compared their performances with those of a group of subjects who had been given comparable questions in a different way. Among other things, it transpired that checking boxes was faster and led to fewer mistakes than underlining the correct answer or deleting the incorrect answers. Wright was also able to show in her experiments that ticking adjectives *(single/married)* took less time and produced fewer mistakes than replying to yes/no questions *(Are you married? yes/no)* [8, 9].

An advantage of laboratory research of this kind lies in the high degree of internal validity achieved by the often refined nature of the research structure. A disadvantage is the limited ecological validity, i.e., what the subjects of the experiment are required to do often differs quite markedly from what people need to do in practice.

In order to obtain a coherent picture of the way in which forms can be improved, it is therefore also necessary to investigate the problems that people experience in completing forms in real-life situations. To date, such research has been very limited.

Holland and Redish compared the behavior and products of people who completed the *Federal Government's Joint Application Form SF 171* with varying degrees of success [10]. They found, among other things, that applicants who had completed the form properly had paid much more attention to the rhetorical situation. In other words, they had endeavored to understand the function of the form as a whole and the functions of the individual questions. They had also tried to anticipate the possible reactions of the person processing the answers. Less successful applicants had concentrated mainly on the decoding strategies. They had limited themselves to understanding the specialized terms and the structure of the sentences [10].

Rose gathered data on the problems of completing American government forms by an indirect method: the critical incident approach [11]. He asked his respondents, who—in their professional capacities—often helped people to complete forms, to provide a detailed description of a problem they had recently encountered in practice. In addition to the expected problems regarding word choice and syntax, the respondents frequently mentioned cases in which the form-fillers had
been unable to answer the questions because they did not have the necessary information. In addition, the respondents mentioned the many awkward calculations that had to be made [11].

Frohlich arranged for eight English people to fill in an application form for Supplementary Benefit while thinking aloud [12]. He was particularly interested in the routes that the subjects of the experiment followed and in the mistakes they made. Frohlich found that the subjects immediately looked for the first question, and then answered the other questions one by one. Explanations and instructions were generally ignored. As a result, many irrelevant questions were answered and many relevant questions were answered incorrectly [12].

Design

In our research we used both a larger number of subjects and more forms than did the studies of Holland and Redish or Rose and Frohlich. Ninety-eight subjects, all of whom belonged to groups entitled to a specific benefit or subject to a specific (tax) regulation, were presented with one or two relevant government forms. They also received a situation sketch, which was meant to put them in the position of a Mr. John Doe and acquainted them with his particular financial situation. We asked the subjects to complete the form while thinking aloud and using this situation sketch. We used nine forms; three were from the tax authorities, five from the Ministry of Education and Science, and one from the council of the city of Hengelo. We collected four categories of information: the answers noted on the form, the remarks of the subjects during the experiment, observations of the behavior of the subjects, and replies to questions asked in a supplementary interview.

In analyzing the data, we were mainly trying to find the problems with which the subjects had been confronted. Four kinds of problematic events were distinguished. They are listed below in the order of importance we attach to them:

1. Ineffectiveness — participant incorrectly omitted a step in the process of completing the form, or incorrectly performed a step;
2. Inefficiency — participant took an unnecessary step;
3. Lack of understanding — participant performed a particular step for the wrong reasons and/or did not really understand what he or she was doing; and
4. Lack of autonomy — participant requested the help of third parties when performing a particular step.¹

¹ Of course, we do not object in any way to civil servants, friends, colleagues, etc. helping people to fill forms. But for the purposes of the research, we regarded it as a problem if a respondent could not succeed without outside help.
We also developed a model of the sub-tasks which appear necessary for the main task of correctly completing a form. The model is shown in Figure 1.²

Three levels of tasks are distinguished. In the first place, there are the sub-tasks that are directly connected with the main task, i.e., answering the questions on the form. Following the example set by Landa, we used the term *functional tasks* [13]. We distinguish between three categories:

- **Generating data** — producing an item of information such as a name, date, or amount. Sometimes data of this kind are known; at other times they have to be looked up.
- **Verifying conditions** — checking whether an item of information fits into a particular category.
- **Transforming** — converting an item of information by using it together with another item of information as an input for a calculation or by converting it into a code, i.e., into a verbal answer and/or a graphic sign (check in a box).

To ascertain what functional tasks must be performed, the user must consult the text of the form (and often the explanatory notes, too). The sub-tasks that must be performed in this connection are called *interpretation tasks*. We distinguish

---

² Although this model is not intended as a theory on sub-processes that play a role in the completion of forms, it could form a basis for such a theory. The model could undoubtedly be elaborated in more detail and assessed on its theoretical merits. This has not been done in this study; here the model had merely a heuristic function.
between two categories of sub-tasks, the first of which is again divided into two sub-categories:

- Understanding the verbal and visual aspects of the text at a local and at a general level; seeing the significance of visual signals and words, groups of words, sentences, and larger elements of text. This is called the **semantic aspect** of the interpretation.

- Inferring from the meaning of the text what functional acts must be performed in order to solve the problem—the **pragmatic aspect** of the interpretation.

The person completing the form must constantly make decisions on whether or not to perform a particular sub-task. The tasks involved in controlling this process are called **monitoring tasks**. Four types are distinguished among them:

- **Orientation tasks** — to obtain an overview of the document and its various parts, and of the overall task and the various steps that it comprises.

- **Selection tasks** — to decide which questions and explanatory passages need or need not be consulted and, if so, in what order this is to be done.

- **Checking tasks** — to ascertain whether the tasks are being performed correctly.

- **Switching tasks** — to interrupt the performance of interpretation tasks to perform functional tasks, and vice versa.

In looking for the **causes of the various kinds of problems**, we adopted three different approaches:

- We examined to what extent problems might be caused by inadequately performed sub-tasks other than the one giving rise to the problem. A wrong selection decision may, for example, be the result of inadequate orientation; an incorrect calculation may be the consequence of an incorrect plan of action which itself is due to an incorrect interpretation at the semantic level.

- We examined to what extent problems might have arisen because certain general preconditions for the successful performance of the tasks concerned were not met. Both orientation and interpretation of the text presuppose, for example, that the person completing the form has sufficient prior knowledge to be able to understand the context of the information presented. Also, where calculations have to be made, the respondent has to possess sufficient general mathematical skills.

- We tried to discover stereotype patterns in the behavior of the subject, i.e., to discover strategies of respondents which ultimately led to problems.
Results

The analysis of the data showed that problems occurred in the case of each of the separate sub-tasks. For example, it was evidently very difficult for the respondents to obtain and keep an overview of the form, and of the activities that they were required to undertake in this connection. They also had little understanding of textual and graphic selection instructions, and found it difficult to understand terms and syntactic structures. The calculations were difficult, and many of the replies were given in the wrong format, particularly if unconventional coding systems were used in a form.

Looking for the causes of the problems, we noted above all substantial discrepancies between the assumptions that are evidently made by the compilers of instructive documents regarding the attitudes, knowledge, and skills of the respondents on the one hand and the actual situation on the other. As far as attitude was concerned, it was striking that respondents did not, as they were evidently expected to do, read systematically through the information and follow the instructions step by step. They almost always made a beeline for their target, an approach that can be described as jumping the gun, or—to use a soccer term—as a "kick-and-rush strategy." The respondents wanted to answer the questions as quickly as possible on the dotted lines and in the appropriate boxes, and they hardly took the time to read the explanatory notes first, or even to read the questions carefully to see precisely what was being asked. They confined themselves to what they considered to be the minimum needed in order to perform their main task successfully.

A second conclusion was that the questions and explanatory notes often required greater reading proficiency than the majority of respondents possessed. With a few exceptions, questions that were phrased in a complicated sentence had to be read aloud a number of times before they were understood. The structure of the text of the explanatory notes often was not understood at all, resulting in, among other things, incorrect decisions about what to read or not to read. Sometimes the subjects also had an insufficient grasp of all kinds of graphic signs that were used in the forms, such as arrows, colors, and different typefaces. Also striking was the lack of arithmetic skills, which became apparent whenever even slightly complicated multiplications had to be done.

The third conclusion was that the subjects had too little background knowledge about the procedures to enable them to complete the form successfully. Many of the professional terms used were unfamiliar to the subjects, and they often completely failed to understand the purpose for which the information would be used. Nor was it, in many cases, clear to them what importance particular information had.

On the basis of these results, we formulated three criteria for the drafting of forms. The first criterion is that the questions and the various kinds of explanatory notes should be drafted from the perspective of the required actions or, to use a
term introduced by Flower et al., according to the *scenario principle* [14]. The information should be presented as much as possible in terms of specific situations, followed by instructions for the actions that have to be taken in each situation. The text should, therefore, not be intended primarily to explain the benefit or regulation, but to help solve the problem facing the respondent in his or her individual situation.

Strict control of the behavior of the respondent is the second criterion. It should be made clear in detail to the reader precisely what is expected of him or her. Clear instructions should be given, both regarding the way in which the information must be generated, verified, or transformed, and on the way in which the answers must be coded. The so-called routing instructions also seem to be of great importance, i.e., instructions for skipping irrelevant questions and references to the explanatory notes (and from there back to the next question).

The third criterion is adequate background information. This is needed in order to compensate for the observed lack of prior knowledge. General background information should be restricted to a concise summary of the system of the scheme, and an indication of the purpose of the form. Such information should be marked in such a way that the person completing the form can easily recognize its function and is encouraged to read it through before starting to fill in the rest of the form. In view of the *kick-and-rush strategy*, detailed information should, wherever possible, be given at a local level, i.e., alongside the questions to which the information is relevant.

**RESEARCH INTO POSSIBLE IMPROVEMENTS**

The next step in our research consisted of revising a number of the forms that had been examined. We will illustrate our approach with some examples of the measures we took.3

**Revision**

One major modification is related to the order of the various questions on the form. We calculated which of all the possible sequences in which the questions could be answered would, on average, cost the people completing them the least time. We also expressly sought a thematic connection between the questions; questions which related to the same themes were then grouped together.4 The effect of altering the order of the questions is evident in Figures 2 and 3.

---

3 A more complete picture of the measures taken is offered in our thesis [4]. Together with four colleagues, we have written a handbook in which these measures are presented as guidelines for form designers [15]. Both publications are written in Dutch.

4 In our thesis, one chapter is devoted to a detailed treatment of this subject. An adaptation in English is forthcoming [16].
The original version of the form (Figure 2) does not show what route the person completing the form should follow. Indeed, there is even some uncertainty about how many questions should be answered (for example whether only one or more of the first six boxes should be checked). Nor is it clear whether the presumed date of starting/leaving the national service should be completed by all respondents, or only by those who have passed the medical examination. In the revised version (Figure 3) the route is clarified.

A major change in the design of the form was the marking of the various kinds of information. Often a so-called three-column approach was chosen; the form was divided into a left-hand column containing the notes explaining the questions, a central column containing the questions themselves, and a right-hand column in which the replies had to be entered. By placing the information in accordance with the natural reading direction, i.e., from left to right, we expected that the respondents would be encouraged to consult the relevant explanatory notes first, then to read the question, and only afterwards to reply to it. An example of a three-column structure is shown in Figure 4.

In formulating the questions, we tried to be direct and simple. Where possible, we opted for closed questions, preferably yes/no questions. Where we came across questions involving a denial, we replaced it with positive versions. To give some examples:

*Was there any period in this quarter during which you had to no work?* was rephrased as:

*Have you worked continuously throughout this quarter?*

Composite questions, such as:

*Are you married and are you not judicially separated?* were always split to form combinations, such as:

1. *Are you married?*

and

2. *Do you live with your spouse?*
### Military service

<table>
<thead>
<tr>
<th>Question</th>
<th>Options</th>
<th>Go To</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 Are you performing your national service at present?</td>
<td>yes no</td>
<td>16</td>
</tr>
<tr>
<td>16 Since what date have you been performing your national service?</td>
<td>day month year</td>
<td></td>
</tr>
<tr>
<td>17 Until what date will you be performing your national service?</td>
<td>day month year</td>
<td></td>
</tr>
<tr>
<td>18 Tick as appropriate. Then skip questions 19 to 22, and proceed to question 23.</td>
<td>I am following officers training</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>I am following NCO training</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I have submitted an application for temporary leave</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>none of these</td>
<td>23</td>
</tr>
<tr>
<td>19 Have you had your medical for national service?</td>
<td>yes no</td>
<td>20</td>
</tr>
<tr>
<td>20 Have you passed your medical?</td>
<td>yes no</td>
<td>21</td>
</tr>
<tr>
<td>21 Tick as appropriate</td>
<td>postponement requested</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>postponement granted</td>
<td></td>
</tr>
<tr>
<td></td>
<td>exemption requested</td>
<td></td>
</tr>
<tr>
<td></td>
<td>exempted</td>
<td></td>
</tr>
<tr>
<td></td>
<td>liable to special conscription</td>
<td></td>
</tr>
<tr>
<td></td>
<td>none of these</td>
<td></td>
</tr>
<tr>
<td>22 On what date must you start your national service according to your call-up card?</td>
<td>day month year</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. Example of modification: the unclear route (see Figure 2) has been clarified.
Figure 4. An example of form design based on a three-column approach.

Results

To find out the effects of our modifications, we compared the answers noted by the respondents on the original and on the revised versions. Just as in the case of the original versions, we arranged for the revised versions to be completed by subjects belonging to the target groups. Eighty-six subjects were involved, none of whom had completed the original versions. All subjects worked with the same situation sketches, as in the case of the original versions. Once again we asked the subjects to think aloud; the remarks were recorded on audio tape.

Figure 5 shows some combined results. They relate to the seven revised forms with the relevant explanatory notes. Five of the forms were from the Ministry of Education and Science, the other two were from the tax authorities. The effectiveness of the original of each form was compared with that of the new version that had been modified in accordance with our criteria.

Figure 5 shows that a considerable improvement has been made, calculated over all subjects and forms. The proportion of correctly completed forms rose from 12.3 percent to 52.2 percent. Clearly this was a considerable improvement, but in absolute terms the results were still a bit disappointing. Almost 50 percent of the forms were still not properly completed.

One point should be noted, though, in connection with these figures. All forms, original and new, were tested to the extreme. We arranged the situations in the sketches so that there was a real chance that the subjects would be confronted with problems. It was hardly surprising, therefore, that these problems actually

<table>
<thead>
<tr>
<th>Explanatory notes</th>
<th>Questions</th>
<th>Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>If you are a married woman, only fill out your maiden name here.</td>
<td>1 Your surname:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 Your initials:</td>
<td></td>
</tr>
<tr>
<td>If you are married but not living together with your spouse, tick the no box.</td>
<td>3 Are you married?</td>
<td>□ yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>□ no</td>
</tr>
</tbody>
</table>

A form was regarded as correctly completed if it did not contain any fatal errors, i.e., missing or incorrect replies that could lead to incorrect decisions when the information was processed, or that could make it impossible to make a decision. If an error did not directly affect the procedure, it was not treated as fatal. Nor was it a fatal error if the correct answer was apparent from the rest of the form or from an annex.
occurred. As a result, however, no direct conclusions could be drawn about the precise percentage of forms that would have been properly completed in reality. Nonetheless, in view of the size of the difference, it seems likely that a similar improvement would be achieved if the new forms were used in practice.

In analyzing the results, we tried to determine not only the effect of the modified version, but also the influence of possible variables relating to the subjects, such as educational level, personal income, and experience with completing forms. We discovered an interaction effect between education and version. It seems that people with a poor education benefit to a much greater extent from our versions of the forms (from 6.9% correct to 70.8% correct) than do people with a better educational background (from 14.3% correct to 47.0% correct).

How can these results be explained? In view of our efforts, the fact that the variable version did have an effect pleased us, but did not really come as a surprise. All the modifications were intended to increase the ease with which the forms could be completed, and this appears to have worked, at least up to a point. What is more difficult to explain is the interaction between educational level and version. Our impression is that the errors made by the subjects with a better educational background in completing the original versions were quite often due to nonchalance, and to an underestimation of the difficulty of the job in hand. Subjects with a poor educational background, however, seem to get more frequently into difficulties with the original version because of a lack of prior knowledge and inadequate proficiency in reading and arithmetic. If these suppositions are correct, we may have succeeded mainly in solving the problems of lack of knowledge and proficiency, but we will have been less successful in dealing with matters of attitude, such as the kick-and-rush strategy.

**DISCUSSION**

The modifications that we made to the forms were not confined to strictly defined categories. We followed a number of guidelines, but it proved necessary
in the case of each form to take account of technical and organizational parameters, to strike compromises between the various requirements, and to cut various Gordian knots. Sometimes the most simple phrases of a question had to be rejected because they were not legally watertight, and at other times the explanatory notes had to be limited because there was insufficient space on the paper (in some cases an extra A4 sheet can cost hundreds of thousands of dollars in extra printing and postage costs). Similarly, sometimes it was not possible to use the clearest graphic design because it was incompatible with the chosen production method; in other cases the order of the questions had to be adapted to fit in with the procedure by which the answers were processed. As a result, the contents and presentation of the forms as ultimately modified were not only determined by the specific guidelines of our research, but also depended on ad hoc alterations.

A partial ad hoc approach of this kind forms a threat to the internal validity of the research, and hence to the ability to generalize from the results. It is unclear to which modifications (or combinations of them) a given effect must be attributed. It is also unclear how much influence is exerted by the skills of the people who modified the forms. How would the forms have looked if the modification criteria had been applied by other people?

Clearly, the research carried out was limited in this respect. However, we feel that it is a useful supplement to the laboratory research referred to previously, in which versions of forms that differ from one another in one or in a few controlled respects are compared. In practical research such as ours, which attaches a high priority to ecological validity, it is possible to ascertain whether the ideas that can be obtained from laboratory research are feasible in practice and can—in combination with other modifications—produce real improvements. Conversely, practical research can also generate a number of specific questions that can only be dealt with in laboratory research.

One of the questions which emerges from our research, and should be dealt with in laboratory research, is the way in which respondents can be encouraged to pay attention to the explanatory notes. Many problems were due to the lack of background knowledge of the people completing the forms, and were compounded by their habit of ignoring the explanatory notes. Various techniques were used in the revised forms to counter this habit. The basic principle was the adoption of the three-column approach described above. However, the way in which the information was provided in the left-hand column varied. Sometimes the requisite explanation was given there, and at other times reference was made to a particular passage in a separate document. In one form, there was no space for three columns and a different system of presentation was chosen. Each term that required explanation was immediately followed by the symbol (T), which meant that the term was dealt with in a sheet with separate explanatory notes (Toelichting) accompanying the form. Systematic research into the effects of these variants would certainly be desirable.
More research would also be needed concerning the influence of the chosen research method on the behavior of the respondents. As previously mentioned, we arranged for the subjects to complete the forms on the basis of situation sketches, among other things in order to enable us to compare the results of the original forms with those of the modified forms. The obvious question is whether respondents act differently when using situation sketches, and whether they encounter problems different from those met by people completing the forms on the basis of their own situation. The same question can be asked about the instruction to think aloud [17]. No clear impression of the effect that thinking aloud has on the way in which forms are completed can be gained either from the literature on thinking aloud or from the interviews we conducted later with our subjects. Since the thinking-aloud method—whether or not on the basis of situation sketches—is increasingly being recommended as an aid in the development and revision of documents, validation research is urgently required [18, 19].

The last matter that we wish to raise here is the extent to which it is possible to generalize on the basis of our observations. The behavior of our subjects is remarkably similar in some respects to that of users of product instructions and computer manuals, who also tend to read carelessly and try to perform the functional tasks without seriously studying the information in the manual [20]. It seems worthwhile to examine the extent to which this behavior characterizes the way in which instructional texts of this nature are used, and how—in view of this behavior—such texts can be improved.

ACKNOWLEDGMENT

We are grateful to Gerard Nas for his comments on an earlier version of this article.

REFERENCES

5. S. Huismans and E. Siegerist, Burger en overheid; Een analyse van wat de Nederlandse burger vindt van de overheid, Alphen a/d Rijn, 1974.
7. C. Jansen et al., De GSD en zijn cliënten; Verslag van een onderzoek naar opinies van cliënten over de Gemeentelijke Sociale Dienst van Hengelo(O), PvdA, Hengelo, 1984.

Other Articles On Communication By These Authors


Direct reprint requests to:

Dr. Carel Jansen
The University of Utrecht
Trans 10 3512 JK Utrecht
The Netherlands