

# Screen Layout Standards at Statistics Finland

*Vesa Kuusela, Survey Research Unit, Statistics Finland*

## 1. Introduction

Probably all professional survey organizations apply a standardized interview method in their data collection. A question presented in an interview is supposed to initiate a measurement process and the given answer is the result of measurement. However, a survey question may be a very sensitive meter affected by various disturbances. For example, the result may vary considerably depending on how the question is presented. There is a wealth of examples how even small, almost unnoticeable, changes in the wording of a sentence may change its meaning. In surveys, where many interviewers ask questions, error variation due to different interviewing styles is always a serious threat to the credibility of surveys. The main purpose of the standardized interview method is to minimize inter-interviewer variation but it reduces intra-interviewer variation, as well.

The interviewing method is composed of several different topics but in brief, it means that questions are written in a specific format and interviewers should follow the format as instructed. One point here is that all questions should be asked exactly as they have been written. In other words, interviewers should follow strictly the wording of a question and use auxiliary information only on a specified style. An implication of the method is that interviewers have been trained to present – and understand – questions in similarly basing on this writing style. Another implication is that the questionnaires have to be written and formulated in such a style that interviewers may be expected to follow the specifications. One should bear in mind that an interview is an interactive discussion between two people and in that situation, a badly phrased question cannot be asked as such.

An essential precondition in carrying out standardized interviews is an explicitly defined writing style of the questions. That is the most important reason, why the layout and style of CAI questionnaires should be standardized. Reversibly, without a standard layout standardized interviewing cannot be expected. If the questions on different questionnaires are presented in varying styles, interviewers probably ask questions in varying styles, as well. That will increase inter-interviewer variation immediately and in the long run it will gradually lead to corruption of the method interviewers have been trained to apply.

Interviewers' working ergonomics is another reason to standardize screen layouts. If several authors make questionnaires and all were allowed to make the screens as they like, probably screens would be composed basing on different principles. In practice that would mean that interviewers needed to learn different styles. Coping with them might become very demanding and in any case, it would be waste of interviewers' time – and nerves.

The layout and presentation standards also facilitate authors' work. When there is a good and easy to apply standard the authors don't have to design the screens anew every time they make a questionnaire. Sometimes a problem also is that authors do not have at their disposal similar laptops as interviewers have. Computer displays may differ considerably and a design, which looks good on one screen, may be nearly useless on another. It is difficult only to imagine what a certain design would look like on another screen but imagination is not needed if one can apply a ready-made and tested solution.

A minor, but not meaningless, reason to standardize question layout is the external image of the agency. Respondents only rarely see the screen of the computer but sometimes they do, and they form partly their opinion on what they see. Interviewers also embrace the idea of the organization and its style of working by the tools they are given.

The need for standardization of questionnaires is not new issue, but the need becomes more prominent with graphical user interfaces (GUI). The paper questionnaires have already been standardized for a long time and on text-based screens, partly the same standards could be applied as on paper questionnaires. Therefore, the need of standardization of text-based screens was much lesser than on graphical screens, and on the other hand, there were not many options in designing text-based screens. Graphical user interfaces have changed the situation dramatically. For example, Windows Blaise provides almost endless amount of possibilities for variations

In year 2000, at the same time when Statistics Finland acquired new laptops for interviewers, the entire CAI information system was converted to Windows and renovated comprehensively. Naturally, also Windows Blaise was taken in use. Soon the need to define new layout and presentations standards to comply with implication of the new environment became obvious. A group of few people with different backgrounds was collected to design a new standard.

## **2. Aims of design**

At Statistics Finland, a standardized interview method was adopted already long before CAI started, as probably in all professional survey organizations, and the interview method has been an inherent part of interviewer training all the time. In other words, the some presentation standards were already defined (and applied) when CAI was introduced, and interviewers were experienced in applying it. The standards were designed mainly for paper questionnaires but they could be followed reasonably well still when using text based screens and DOS software.

The new layout and presentation standards were outlined to compose rather a new design than only a refinement of the previous one. Graphical User Interfaces provide plenty of new features and possibilities, which could improve the functionality of the questionnaires when applied deliberately. That is, GUI also includes features, which could focus questions better and decrease inter-interviewer variation. An attempt to apply only the existing methods had been a waste of those possibilities. In addition, some of previous solutions were too laborious and cumbersome to accomplish in designing CAI questionnaires.

The one of the most important aims of the design work was that result should function well on the field interviewers' laptops. The LCD displays of laptops are more sensitive instruments than CRT displays of desktop computers. In addition, laptops are often used in settings that are more difficult to control. Three other general aims and preconditions were also set for the designers

1. to provide interviewers with easily comprehensible screens were important parts are easy to find
2. to keep amount of work reasonable required from the authors
3. to retain as much as reasonable the features of the older standard

A danger, when having all the possibilities that the Windows environment provides, is to get too exited of them, or to exaggerate the number of elements which need special attention. The resulting screen would be restless if it contains many different colors and several fonts, in addition to other graphical elements. On

an incoherent and tangled screen, even important parts may be difficult to notice or they may be even lost amidst the great number of different details.

On the other hand, some features, which Blaise and Windows environment provide, would bring about very informative and clear screen designs but some features had required too much detailed editing from authors. For example, if those answer categories, which should be read to respondent, were indicated using a different color than in question, individual attributes had to be attached to each answer category. In a long questionnaire, that would considerable increase editing.

As there already existed a standard, which interviewers were trained to apply, it was considered unnecessary to give up all the conventions. Therefore, those elements of the old standards were retained which were found good and working and which were easily applicable.

## 2.1 Prevailing conventions

The prevailing interview method at Statistics Finland already defined some rules for the compilation of questions. The style how a question is written includes, besides the question itself, also instructions for interviewers. They tell in which manner a question should be read to a respondent and how to continue in case the respondent does not understand the question. A major distinction was between presenting a factual and an attitudinal questions. The main rules say in brief:

- 1) A question text should always be read up to question mark;
- 2) First time a question should be read exactly as it is written;
- 3) If respondent does not understand the question,
  - a) in a question concerning a **fact**, interviewer should specify or explain it by his or her own words;
  - b) in question concerning **opinion, attitude, or knowledge**, interviewer should specify or explain the question only using instructions given on the screen.

These rules are obvious and their realization does not necessarily call for any graphical solutions. For example, attitudinal question was indicated with letter M (abbreviation from Finnish) after the question number. This kind of conventions could be copied from the prevailing standard.

## 2.2 Elements of a CAI question

Questions, which appear in a CAI questionnaire, have a relatively common basic structure. There are a limited number of different elements on a question screen, which should be discernible to direct interviewer's attention to them and hence to facilitate and to improve interviewer's performance.

**Question identification and attributes**, like question name or number, type of the question, attributes of the question, use of display card, and previous answer to the question in longitudinal surveys

**Question text and answer categories**. The elements of a question, which should be taken into consideration, are

- the part of the text which should always be read
- the part of the text, which is read conditionally to specify the question
- emphasized parts of the question text indicating key words
- indicator whether answer categories should be read or not

**Instructions for the interviewer** are needed for several different purposes:

- Complementary or specifying instructions for questions
- Definitions and concepts
- Instructions how answers should be marked
- Directive instructions
- Very important directive instructions

The CAI software takes automatically care of some issues, which were essential in paper questionnaires. For example, routing instructions, which were inherent elements in paper questionnaires, are not needed at all. Another example is a multiple answer question, which is handled by the system and does not need additional instructions. In addition, the default screen of DEP already has a lot of information, which is useful for interviewers.

### **3. Layout and presentation standards**

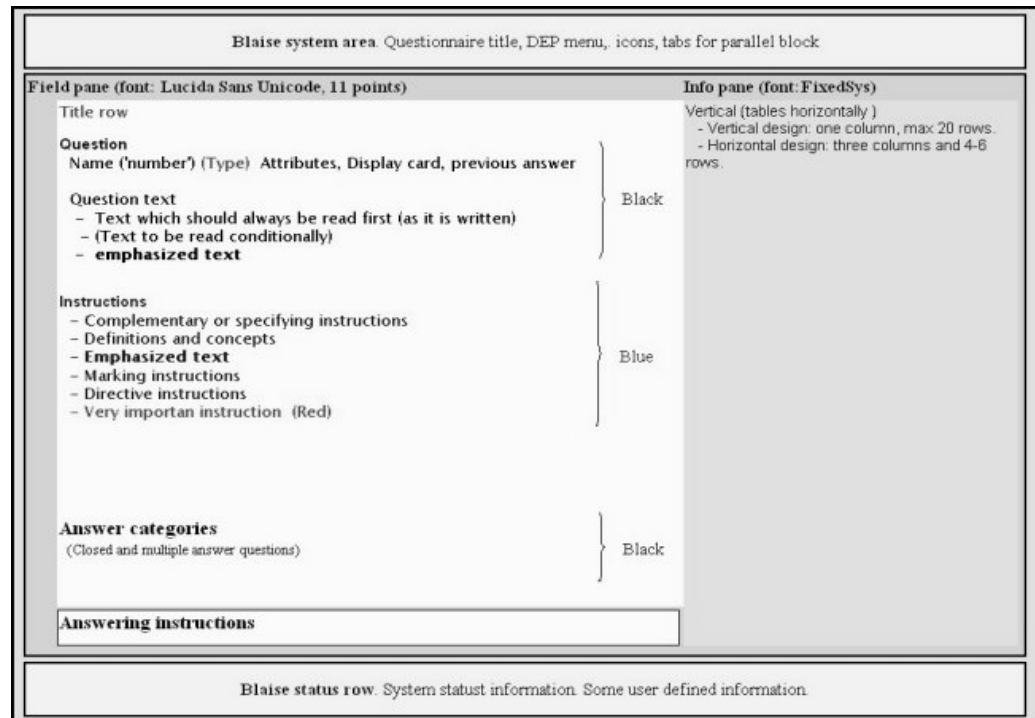
The standard, which was eventually agreed on, defined the division of DEP screen, menu items, writing style for questions and instructions, and the use of colors and fonts (see figure 1).

In most cases, the DEP screen is divided horizontally. Only on large tables, vertical division should be applied. The reason for horizontal division is that on a 1024\*768 pixel screen rows are very long if they go from margin to margin. Horizontal division makes rows shorter depending, of course, on how the division is done.

The reason why a horizontal division was selected was that a short text line takes easier a shape and is therefore easier to read. The horizontal division also provides more lines for the question. This arrangement allows for more possibilities for modifications in question writing and layout design.

All Blaise system texts visible in Dep were translated in Finnish as well as the Dep menu. The default Dep menu was slightly edited to meet the needs of interviewers and to comply with the information system. In practice that meant that only those menu items are shown which interviewers needed. In addition, some specific design features were taken in use. For example, parallel blocks are shown as tabs, and the line between question text and answer categories was left away.

Figure 1: Schematic presentation of a CAI screen designed for layout standard.



Only one font (Lucida Sans Unicode, 11 points) is used throughout the field pane, and another (Fixedsys) is used in the info pane. Emphasized words are written in bold in every instance, including questions, answer categories and instructions. Text that is read conditionally, e.g. if respondent did not understand the question, is surrounded by parentheses.

Whether answer categories should be read to the respondent or not is indicated by placing a question mark after the question text or after the last category to be read, or verbally in the instructions (as previously). Several other alternatives were tested, as well, but they all appeared to be too laborious to carry out in everyday authoring practice.

Blaise default color scheme is used except in Field pane texts. Colors make the distinction between the question text and instructions. Both the question and related answer categories are shown in black.

All instructions, wherever they appear on the screen, are blue. Although there are many different types of instructions, they were not discriminated by any means. The informative value of discriminating between different instruction types was reflected against the graphical image that had resulted. Many different colors would have been too distracting and it was given up. Different fonts might have been possible but that was not considered beneficial enough because of increased need of editing.

However, there is one exception to the previous rule. Red is reserved (only) for very important instruction, and to indicate whether the question is an attitudinal one. Red color should be used very sparingly to preserve its effect.

#### 4. Discussion

It is not sufficient only to define screen layout standards. They also have to be widely known and accepted to become a part of everyday questionnaire authoring. The two critical groups here are those who make the questionnaires and

researchers. Both groups have to understand the importance of everyone applying the same standards. The authors of the questionnaires have to know the technical details and how to apply them – and take them as a natural part of questionnaire design. Researchers, on the other hand, have to understand the importance of the standards and learn to require their use in every questionnaire.

At Statistics Finland, the defined screen layout standards were published on a paper printout, which was presented in a meeting to all those people who deal with CAI questionnaires in a way or another. In addition, a ready-made Modelib was made available for everyone.

The defined standards have been in use for some time in all questionnaires of Statistics Finland. The feedback from interviewers has been mainly positive. In addition, authors workload has remained bearable. Hence, the result may be regarded to satisfy the requirements, which were set. However, the defined standard should not be regarded as a final end result. Rather question is of a continuous process.

The Blaise Modelib editor proved to be a surprisingly cumbersome tool and its behavior remained somewhat vague. Especially, the definition of the place and size of screen elements was difficult. During the designing and testing of different screen layouts a need for a WYSIWYG type modelib editor came up many times.

The screen layout is not the only part of a questionnaire, which needs to be standardized. Some structural and syntactical standards are equally important. For example, then an adequate naming convention of questions and blocks helps navigation within a large questionnaire. These kinds on standards are not as visible as the screen layout but also they could make interviewer performance more accurate and easier.

Finally, one should notice that layout and presentation standards are only a small, but important, factor influencing the survey quality. Concepts, question wording and question order have a much deeper and wider effect on survey instruments and survey process (not to speak about sampling and non-response). However, adequate standards make good instruments work better – and poorly defined standards make inferior instruments work worse.