

**Methodology Advisory Committee (MAC) paper: CAI screen design
project**

Jane Barresi, David Finlay, Alanna Sutcliffe
Information Management Division
Australian Bureau of Statistics
June 2002

1.1 Introduction

1 As the ABS reintroduces Computer Assisted Interviewing (CAI) for our household surveys, a number of important issues are currently being considered to enable a successful transition. One of these issues under consideration is determining the appropriate screen layout for interviewers. This paper summarises the issues pertinent to designing a suitable screen layout to assist in the minimisation of measurement error. The paper will cover the following main issues:

- purpose of this paper;
- history of CAI in Australian Bureau of Statistics (ABS);
- CAI screen presentation in the ABS;
- principles of good screen design (for example, mixed case text and context); and
- proposed usability testing methodology.

2 This paper provides a summary of the project being undertaken in the Population Surveys Branch to assess the impact of screen design on the collection of data and the methods for assessing its suitability for interviewers. Specifically, rather than recommending the key features of optimum screen designs for ABS screens generally, we have progressed to recommend a screen design layout which optimises the interviewer's ability to interpret the screen presentation, read out the relevant question text with accuracy and at the same time follow any specified instructions such as showing a prompt card or recording multiple responses. We are now evaluating this screen design through usability testing (Hansen and Fuchs, 1997).

1.1.1 Purpose of MAC paper

3 The purpose of this paper is to seek feedback from the MAC on our usability testing methodology. Considering the diverse methodologies available for evaluating screen design, we would specifically like the MAC to focus on whether you think the chosen methodology is appropriate?

1.2 Background

1.2.1 What is CAI

4 CAI involves the use of the computer to prepare an electronic questionnaire which is then used to conduct the survey interview. Basically, the questions and related explanatory material are displayed to the interviewer, or respondent (if self-administered), on the computer screen. Responses are then keyed into the computer by the interviewer (or respondent). CAI in the ABS environment tends to be CAPI (Computer Assisted Personal Interviewing) or CATI (Computer Assisted Telephone Interviewing), however other innovative approaches are also being used by other statistical organisations (eg. palm held computer devices are regularly used for surveying at Office for National Statistics in the United Kingdom).

1.2.2 History of CAI in the ABS

5 The ABS begun using CAI for Special Supplementary Surveys (SSS) in the early 1990's, however in 1996 it was decided by ABS management that CAI for ABS household surveys should be discontinued primarily due to cost considerations. During the next few years CAI was gradually wound down until it was revived again in 1999 subsequent to a ABS business case proposing CAI's re-introduction after having careful consideration of its advantages and disadvantages. Since this time, ABS has successfully re-introduced CAI for ABS household surveys and is currently planning for the CAI conversion of the Labour Force Monthly Population Survey (MPS) to CAI in the near future. The latter will be done in a way that minimises the risk of an impact on the key output statistics due to methodological change.

6 The software that has been adopted for use in CAI household surveys is Blaise, and is used by a number of the international statistical agencies. With the advent of Windows technology in Blaise and more powerful notebook computers, it has provided more potential in screen design than was previously possible under the original DOS version. The current Blaise software has certain characteristics which cannot be altered but also has extensive features for screen presentation which can be used to implement a wide range of desired outcomes.

CAI Screen Presentation in the ABS

7 The objective of developing a suitable screen layout for CAI is to optimise the interviewer's ability to interpret the screen presentation, read out the relevant question text with accuracy and at the same time follow any specified instructions such as showing a prompt card or recording multiple responses. This objective is no different from that which applies in the current paper questionnaire environment.

8 A literature search was carried out to find relevant international material which dealt with screen design for CAI and Blaise (Bushnell, 2000; Couper et al, 2000; Degeral, 2000; Fuchs, 2000; Hansen, Couper and Fuchs, 2000; Hansen, et al, 2000; Kelly, 1998; Pierzchala, 1998; Pierzchala and Farrant, 2000, Pierzchala and Manners, 2001). Contact was also made with major statistical agencies (eg. United States Bureau of the Census), who make use of CAI, to identify current screen design practices. A sample of Blaise instruments used in some of these agencies was also obtained, and then the Blaise software was scrutinised closely to identify the various adjustments that are possible.

9 Following a review of the above material, a paper has been prepared which discusses and sets out the detailed recommendations for a screen design to be used for CAI at the ABS (these are summarised in section 1.3). This project has developed screen design recommendations that concide as much as possible with current ABS survey design practices, particularly in relation to the MPS paper version. However we expect that any additional CAI screen design features (eg. help option, amount of context to include, etc.) will be considered on a survey-to-survey basis by the relevant ABS project boards.

1.2.3 Usability Testing CAI instruments

10 The introduction of CAI by a number of statistical organisation's has raised a number of additional questionnaire design issues not addressed in the development of paper instruments. Research and other organisation's experience have shown that the advent of CAI over paper questionnaires can produce effects between the two modes of administration. Specifically, these effects have been theorised to be driven by changes in the interviewers interaction with the collection instrument; the reaction of the respondent to perceived confidentiality of the CAI environment; and changes in the process supporting the interview (Couper, 2000; Fuchs, 2000; Hansen, 2000; Hansen, Couper and Fuchs, 1998; Schaeffer, 1995). This has seen the introduction of new research methods, such as usability testing, to evaluate how easy or difficult it is for users to interact with CAI instruments and systems (Hansen and Fuchs, 1997).

11 Usability testing emphasises the cognitive and interactional aspects of computer use, addresses the ease or difficulty a user has interacting with hardware or software and provides evidence of how accurately users can accomplish their tasks by interacting with an application (Bosley, Conrad and Uglow, 1998). Ease of use is believed to be determined in large part by the design of the screen layout (Hansen, Couper and Fuchs, 1998).

1.3 Screen Design Recommendations

12 There are many aspects of screen presentation, covering a range of items such as font type, size, colour, emphasis and positioning of text and other objects (icons, buttons etc). Each of these aspects then need to be considered for each type of field element such as the question text itself, the instructions, the response set and data entry elements. This paper is too short to provide the reasoning behind each recommendation other than to say that they have been considered in the context of the research material and the constraints, if any, which apply to CAI at the ABS (eg the use of Blaise and certain defined screen size).

13 The key features of the recommended screen design are:

- general principle is to keep the screen free from "clutter" to enable the interviewer to focus on the important elements necessary to carry out interviewing functions;
- ensure that questionnaire elements displayed on the screen have a purpose;
- context and navigation information to be displayed discreetly;
- design primarily for use of the keyboard, and limited mouse use; and
- fill text (eg. names are automatically inserted into the question) is to be included in a way that retains the flow of the question.

14 For detail on the recommended design please refer to Attachment A

1.4 Usability Testing Methodology

15 Many of the techniques available for pretesting paper-and-pencil questionnaires, can be used to evaluate CAI systems and survey instructions, including the effectiveness of CAI layout and design. However, instead of focussing on respondents understanding of the questions, usability evaluation focuses on the

interviewers interaction with the CAI system and survey instrument. One method used in Human-Computer Interaction (HCI) research to evaluate usability is the laboratory based usability test.

16 The laboratory based usability test involves people being observed in a controlled setting as they use computer systems. There are less costly methods of usability evaluation (Nielsen and Mack, 1994). However, usability testing is the only method that allows for user involvement and is effective in identifying serious and recurring usability problems (Hansen, Couper and Fuchs, 1998).

1.4.1 Usability Walkthroughs

17 This is a quick and flexible method for collecting user feedback regarding the design of user interface mock-ups and prototypes. During walkthroughs, users are presented screen designs associated with a workflow scenario and are prompted to respond to screen design and interactions. We intend to conduct two small rounds of usability walkthroughs, conducting individual sessions with 6 interviewers in each round (12 interviewers altogether), to establish what the main usability issues are. We have recently completed the first round of usability walkthroughs (see preliminary results described below) in which we aimed to achieve the following objectives.

Broad Objectives

- gaining user (ie. interviewer) input to, and evaluation of, screen design; and
- user validation of designs.

Detailed Objectives

- Assess the readability of questions (ie. font size, mixed case, bold, indentation/margins, font colour) between the two screen designs;
- Assess the readability of interviewer notes (eg. to show prompt cards, mark only one response category, etc.) between the two screen designs;
- Assess the use of icons and upper case words in the interviewers' instructions;
- Assess the use and preferences for fill text to be blended with question text (ie. in terms of colour, mixed case for respondents names, etc);
- Assess use of question numbers (as field tags);
- Assess use of displaying page numbers on screen;
- Assess use of field descriptions in Form Pane, and section headings;
- Assess use of help pop-up box, and what information interviewers would like to be displayed in the help function;
- Assess use of speed bar and status bar (ie. whether these should be displayed on screen) and if so, what information would be useful to display in them; and
- Assess ease or difficulty in navigation through the instruments.

18 The advantages of usability walkthroughs are as follows:

- Fast (informal and easier to set up than more formal usability testing. The focus is on collecting user feedback on design, rather than on objective measures of performance);
- Flexible (can be completed with any level of prototype or design representation, paper prototypes, screen captures or slide shows, electronic mock-ups or prototypes);

- Puts users in the role of the evaluator (ie. users feel very comfortable when asked to review or evaluate a design. The focus is on the interface and not the user); and
- Identify potential user interface issues early (ie. help identify potential usability issues early in the design process when they are easier and less costly to fix. Because they are fast and relatively easy to set up and conduct, they provide feedback on a just-in-time basis).

19 Disadvantages are as follows:

- Utilises subjective data only; and
- Subjective data includes ratings of likes/dislikes, ease/difficulty of use, etc., where as objective data includes task performance, task time, error rates, help references, etc.

20 Although usability walkthrough testing utilises subjective data only, we believe this testing provides valuable information and feedback on the design quickly and efficiently. In particular, this form of usability testing is important in this early stage of the evaluation process because it allows us to get an overall indication of the design's usability in a timely fashion, as well as helping to identify potential usability issues when they are easier and less costly to fix. The findings from the first round of usability walkthrough testing indicates that our screen design is preferred by users in comparison to a design which is based on a Blaise default screen design. The interviewers reported our screen design to be easier to read, navigate through the instrument, and believed it to be more similar to the paper version of the same questionnaire because our screen design allowed them to see preceding and forthcoming questions (see Attachment B).

21 At present we are preparing for the second round of usability walkthroughs. Upon completion of this next round, we intend to refine our screen design based on these findings and then intend to conduct further usability testing in the laboratory and in the field (both of which are described below).

Usability testing in the Laboratory:

22 This is a means for evaluating how well a system measures against its usability criteria. Testing is usually conducted in a specially equipped laboratory with video and sound recording facilities, a test room separated from an observation room by one-way glass, and with software logging and editing facilities. The aims include: to uncover potential trouble spots in the user interface; to identify issues when there is still opportunity to resolve them (identify them in a usability test before waiting to do it in the field); and to obtain both subjective and objective data.

23 The advantages of this type of testing include:

- It considers not only the tasks that users perform but the context in which they are performed;
- enables continual upgrading and maintenance of existing systems; provides an assessment of new software and hardware;
- tests current systems, providing bench-marking criteria for the design of successive versions; and
- helps resolve development and design conflicts amongst team members and stakeholders.

24 There are also disadvantages or limitations that should be highlighted. These include:

- It is not a design method (but it is user centred);
- Users can indicate usability problems but are not able to consider design solutions;
- Time required to set-up and conduct proper test (usability walkthroughs and evaluations may be more 'time efficient');
- Takes the user out of the context of their normal working environment; and
- Usability testing does not assess the user in their natural setting (the field).

Usability Field Test:

25 The usability field test focuses on evaluating the interviewers ability to carry out their functions using the CAI screen design on the MPS. Specifically, the series of usability walkthroughs and laboratory testing of our design will assist in the refinement of the new screen design before the field test and identify the particular usability issues that will need addressing. Similar to usability testing in the laboratory, the aim of the field usability test is to collect both subjective and objective data with a large sample of respondents in a natural setting (ie. the survey field).

26 The field test will involve the administration of our screen design during the MPS field test in October and November 2002. The screen design will be evaluated in both CAPI and CATI modes of administration, which will involve testing interviewers' reaction to the screen layout, the interviewer-respondent interaction and evaluating management information files. Specifically, the interviewers' reaction to the screen layout will be assessed through the completion of brief evaluation forms at regular intervals (ie. after every 5 interviews) during the field test, and then an in-depth group discussion at the interviewer debriefing session upon the field test completion. In addition, the interviewer-respondent interaction (or interview flow) will be assessed through behaviour coding (either observation of live interviews or video-taped interviews). It is intended that the behaviour coding frame used would be based on the Conversation Analytic (CA) technique used by Hansen (2000) in her usability study of the NHIS (see Attachment C). Finally, the interview flow will also be assessed from the management information files which provide the opportunity to collect quantitative data on the number of missed questions, errors, time to read questions and navigation through the instrument.

27 On a final note, the usability field test does not provide a quantitative comparison of the statistics resulting from CAI version of the MPS. As mentioned earlier, the proposed screen design is based on research and recommendations from international statistical organisations to minimise any statistical impact.

Method Source: The Hiser group. (1995). Designing with users: The key to success. Trailmoss- The Hiser Consulting Group.

1.5 Summary

28 We have developed a screen layout which optimises the interviewers' ability to carry out their functions, which we intend to implement in the MPS and other ABS household surveys over time. We expect that any additional screen design features (eg. help option, amount of context to include, etc.) will be considered and tested on a

survey by survey basis.

1.6 References

Bushnell, D. (2000). From DOS to windows: Usability issues for interviewers. International BLAISE User Group- Kinsale May 2000.

Couper, M.P., Beatty, P., Hansen, S.E., Lamias, M., Marvin, T. (2000). CAPI Design Recommendations. *Report submitted to the Bureau of Labour Statistics*. Interface Design Group, Survey Methodology Program, Survey Research Center, University of Michigan.

Degeral, H. (2000). The process of making a new CAI-operation in Statistics Norway. International Blaise User Group. Kinsale.

Finlay, D. (1999). *International Experience in the Introduction of Computer Assisted Interviewing (CAI) into Labour Force Surveys*, ABS Internal document (unpublished).

Fuchs'. M. (2000). Screen design and question order in a CAI instrument Results from a usability field experiment. *Survey Methodology*, 26(2): 99-207.

Fuchs, M., Couper, M.P., and Hansen, S.E. (2000). Technology effects: do CAPI or PAPI interviews take longer? *Journal of Official Statistics*, 16(3); 273-286.

Hansen, S.E., Couper, M.P., and Fuchs, M. (1998). *Usability evaluation of the NIHS instrument*. Presented at the annual meeting of the American Association of Public Opinion Research, May, St. Louis, MO.

Hansen, S.E., Beatty, P., Couper, M.P., Lamias, M., Marvin, T. (2000). The effect of CAI screen design on user performance: Results of an experiment. *Report submitted to the U.S. Bureau of Labour Statistics*. Interface Design Group, Survey Methodology Program, Survey Research Center, University of Michigan.

Kelly, M. (1998). Producing an error-free CAI instrument- is it possible? In. IBUC 98 5th International Blaise Users Conference, Lillehammer, 11-13 november 1998

Pierzchala, M. (1998). Optimal screen design in Blaise. Actes de la 4 Conference Internationale: des Utilisateurs de BLAISE (Proceedings of the fourth international blaise users conference, (5-7 May 1997)- Centre de conferences de Bercy - Paris.

Pierzchala, M., and Farrant, G. (2000). Helping non-Blaise Programmers to specify a Blaise instrument. In International Blaise User Group- Kinsale May 2000.

Pierzchala, M., and Manners, T. (1998). Producing CAI instruments for a program of surveys. In Couper, M.P., Baker, R.P., Bethlehem, J., Clark, C.Z.F., Martin, J., Nicholls, W.L., O'Reilly, J.M. (1998). *Computer Assisted Survey Information Collection*. Wiley Series in probability and statistics. New York.

Pierzchala, M., and Manners, T. (2001). Revolutionary Paradigms of Blaise. In

Proceedings of the 7th international Blaise users conference. IBUC (September 12-14, 2001). Westat. USA.

The Hiser Group. (1995). *Designing with users: The key to success*. Trailmoss Pty Ltd trading as The Hiser Consulting Group. Australia.