ATTACHMENT B

Usability Testing Report

1 Introduction

1 Through an extensive literature review, and consultation with international statistical organisations, Population Survey Development (PSD) have developed a suitable screen design for use at the ABS. In particular, our findings indicated that a screen layout which optimises the interviewer's ability to carry out their functions (ie. read the relevant question text and specified instructions with accuracy) is an important issue to consider in enabling a smoother transition from paper to CAI (Couper et al, 2000; Hansen, et al, 2000). Based on the research findings and recommendations from other international statistical organisations, we developed a screen design for use in ABS household surveys. We would now like to evaluate this screen design, and this is typically achieved through usability testing (Hansen, Couper and Fuchs, 1998).

2 Screen Design Recommendations

2 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 report 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).

3 The key features of the recommended screen design are:

General matters

- general principle is to keep the screen free from "clutter" to enable the interviewer to focus on the important elements necessary to carry out the interviewing functions;
- ensure that questionnaire elements displayed on the screen have a purpose;
- context and navigation information to be displayed discretely (ie. in a way that does not interupt the interviewers main task of asking questions);
- design for the keyboard and limited mouse use; and
- fill text to be included in a way that retains the flow of the question.

Top half of the screen (known as the Info Pane)

• question text to be 12 point San Serif bold mixed case and black in colour, to be displayed at the top left hand corner of the screen (with a small margin on the left and top) and flow from left to right up to about 60% of the way across the screen;

- interviewer instructions and information text to be 11 point Sans Serif bold mixed case and blue in colour, be indented one tab stop with the main action word shown in upper case;
- response categories (if applicable) are to be shown below the question text in a vertical list in the same text format as the question text;
- if insufficient space exists to show the full response set in a single list then a second column (or third) can be used to display additional categories in the set.

Bottom half of the screen (known as the Form Pane)

- text in this part to be 11 point San Serif normal (ie not bold) mixed case and black in colour;
- the screen is to show field descriptions, entered values and relevant labels for the entered responses in the context of the flow of the questionnaire; and
- section headings are to be included to provide context.

<u>Other</u>

- a limited number of icons to be displayed on the icon bar;
- other useful information (eg question identification) to be shown discretely in the status bar; and
- the Windows task bar is to be hidden.
- 4 A sample screen showing the recommended design is shown in Appendix 1.

3 Usability Walkthroughs/Lab Tests

This is a guick and flexible method for collecting user feedback regarding the 5 design of user interface mock-ups and prototypes. During walkthroughs, users (or groups of 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/labs tests using 6 interviewers each (12 interviewers altogether), to establish what the main usability issues are. This report focuses on presenting the results from the first round of usability walkthrough testing which was completed on May 1 and 2, 2002 at the ABS office in Sydney. This involved asking interviewers to provide feedback on the usability of the screen design developed according to the literature (screen A) when compared to a screen layout based on the Blaise default design (screen B). The screen designs were programmed with the Labour Force component from the Monthly Population Survey. Interviewers were individually tested for approximately one-hour each. During the testing session they were asked to complete up to 5 different household scenario's on both screens interchangeably (ie. interviewers were randomly allocated to begin the scenario using either screen A or screen B).

Broad Objectives

- gaining user input to, and evaluation of, screen design; and
- user preferences for screen design.

Detailed Objectives

- Assess the readability of questions and interviewer notes (ie. font size, mixed case, bold, indentation/margins, font colour) between the two screen designs. Note, interviewer notes refer to several different actions interviewers are required to perform, such as handing over a prompt card, or probing the respondent, etc.
- 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. The speed bar and status bar are described in Attachment A of the MAC paper.
- Assess ease or difficulty in navigation through the instruments.

Materials:

- video tape;
- video equipment;
- prepare scenario's and probe/evaluation questions;
- CAI notebooks (with MPS Labour Force Survey); and
- enclosed room with chairs and table

Procedure

i Introduction of test, obtaining consent to video tape, instructions on using the notebooks.

ii Turn video camera on

iii Review each scenario with user (respondent played by facilitator, interviewer uses interface)

iv Users tell the facilitator about the actions they would take and aspects they like or do not like (during or after the scenario). Facilitator can prompt and probe for feedback on main testing issues after each scenario is completed ("Does this make sense?", "What would you prefer to see here?", "Where would you go from here?"

6 Note, the testing is interactive, where the facilitator can interrupt users to probe behind the issues and is careful to make the participant feel at ease and encourage and confirm their responses.

Results:

7 A summary of interviewer characteristics is displayed in Table 1 below. As presented in Table 1, the interviewers involved in the usability walkthrough testing were all female and ranged in age from 41 to 72 years. This is largely consistent with the demographic profile found the ABS interviewer population. The interviewers

tested ranged in the ABS interviewing experience, with two interviewers relatively inexperienced (commenced interviewing < 6 months ago), another interviewer with 12 years experience and the remaining interviewers with fairly extensive ABS interviewing experience. Almost of all of the interviewers tested (5 out of 6) had previous or current CAI experience, however the amount of CAI experience did vary across the interviewers.

Table 1: Summary of main interviewer characteristics for round 1 of usability walkthrough testing (n=6)

Interviewer Characteristics	Measurement Units
Age	Years
Range	41-72
Mean & Median	55
	Number
Sex	
Female	6
English as a first spoken language	6
Education	
Secondary school certificate	3
Trade certificate/apprenticeship	1
Post-graduate diploma	2
ABS interviewing experience	
<12 months	2
1-10 years	1
10-20 years	2
>20 years	1
Experience with CAI	
Yes	5

8 This next section provides a summary of results found against each detailed testing objective.

User input to, and evaluation of, screen design:

9 For simiplicity in reporting the new screen is referred to as "screen A" and the Blaise default screen is referred to as "screen B". In addition, in this report when referring to the "majority of interviewers", this means that at least 4 out of the 6 interviewers responded in a particular way.

Assess the readability of questions and interviewer notes (ie. font size, mixed case, bold, indentation/margins, font colour) between the two screen designs.

- All interviewers reported that they had no difficulty reading the text on both screens A and B. However, they all reported that screen A was the easier of the two screens to read from.
- The majority of interviewers preferred the mixed case text presented in screen A. However this was with the exception of one interviewer (aged 72 years, with 18 years interviewing experience) who preferred the upper case text because "it was familiar". However, she also noted she had no trouble reading the mixed case text.
- All interviewers expressed a preference for the text colour displayed on screen A (ie. black, bolded for questions and blue for interviewer notes).
- Font size was reported by all to be easy to read, with the exception of one

interviewer (with 22 years interviewing experience) who reported that it might be better in a larger font. She noted however that she was overdue for a new prescription on her reading glasses.

• When interviewers were asked to view and then make a comparison between screen A and B on the readability of the first Labour Force question, they all reported that reading screen A was faster and easier. They commented that there seemed to be less to read on screen A because of the indentation in the text in comparison to the way the same question was presented on screen B.

Assess the use of icons and upper case words in the interviewers' instructions.

- When presented the advanced screen version for comment, the majority of interviewers reported that they preferred the use of symbols and short upper case words to aid the reading of notes to interviewers.
- They added that usually interviewers do not read interviewer notes during an interview because they are quite lengthy and it interferes with interview flow. However, they reported that perhaps the use of symbols and short upper case words would help them to quickly read the notes. One interviewer (aged 72 years, with 18 years experience interviewing) was the exception to this, she expressed a preference for maintaining the more lengthy interviewer notes (eg. "Interviewer: show prompt card 9") because it was familiar to her. She did note that she would be able to learn the symbols if they were introduced.

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).

- The majority of interviewers reported a preference for black fill text. However noted that if question text is to be displayed in mixed case then respondent names need to be in mixed case as well (not lower case, which occurred by accident on a few occasions when interviewers forgot to press shift for the first letter).
- One interviewer (aged 61 years, with 22 years interviewing experience) suggested that some fills (eg. dates) should be highlighted in a different colour so that the fill text stands out. However, she noted that there would a temptation only to notice the fill text, and therefore keeping fill text colour in accordance with the question text colour would aid with reading the whole question.

Assess use of question numbers (as field tags)

• Two interviewers noted that they did presently use question numbers for sequencing and/or knowing their place in the survey during interviewing (ie. how far they were through the survey). However, the remaining four interviewers reported that they they did not presently use question numbers while interviewing with paper questionnaires. Rather, they explained that they tend to remember, or refer to, questions by their content or meaning, instead of the question number itself.

Assess use of displaying screen page numbers.

• There was a mixed response across interviewers on this issue. Some interviewers reported that there was a need to have page numbers displayed

somewhere on the screen, mainly so that they know their place in the survey or to quickly transfer to another screen. Whereas, other interviewers reported no need to display page numbers in a computerised instrument, especially if there is no current way of easily transferring to specific screen page number (as a navigation tool).

Assess use of field descriptions in Form Pane, and section headings.

• All interviewers reported a preference for screen A where context was displayed in the form pane. They noted this to have several advantages, such as it defining your place in the survey instrument during interview, identifying and correcting mistakes easily, professional presentation of the screen A design, ability to review consistency of responses during the interview to reduce error, etc.

Assess use of help pop-up box, and what information interviewers would like to be displayed in the help function.

- All interviewers reported a preference for the availability of a help option, however some qualified this by saying it would be most helpful with MPS supplmentary surveys and large household surveys rather than the MPS Labour Force survey.
- The interviewers reported that they liked the idea of displaying prompt card categories within the help box.
- In addition, some made other suggestions that might be included in the help box, such as to have prompts for interviewers, definitions, etc., where appropriate.

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.

- Upon seeing the advanced screen version where no status or speed bar's are displayed, interviewers reported the screen to be much more clearer and appeared wider, brighter or larger.
- In relation to the speed bar *per se*, all interviewers reported using only the keyboard during CAI. Therefore, being able to activate the speed bar (which involves the mouse) was a redundant function for them. This was with the exception of one interviewer (with 22 years experience interviewing and is currently working on the CAI version of the General Social Survey). Although she uses the keyboard predominantly, she noted that to exit from the survey form she uses the speed bar (as taught this in current training).
- However, during testing, she noted she found it easier to press ENTER to exit.

Assess ease or difficulty in navigation through the instruments.

 All interviewers preferred screen A for navigating through the instrument, reporting that it was much quicker and efficient in comparison to navigating with screen B. They noted this was mainly due to the availability of contextual information being displayed in the form pane (ie. context of preceding and forthcoming questions). Two interviewers noted that this helped them to feel more in control of the interview process.

<u>User preferences for screen design:</u>

10 All interviewers reported at the end of the testing session that they preferred screen A to screen B. The main reasons for their preference of screen A were:

- the bolded, black text with indentation and margins allowing for easier reading on the screen;
- interviewer notes in blue and indented;
- contextual information presented in the form pane which allowed for easier navigation and the ability to identify and correct mistakes more efficiently;
- the screen looks clearer and wider thereby aiding visibility; and
- interviewer notes being briefer (especially in the advanced version), more efficient to read.

Recommendations for screen design (Monthly Population Survey- MPS)

- 11 Maintain the following screen A features:
 - mixed case, black, bolded question text with size 14 font;
 - indentation and margins;
 - black fill text;
 - blue interviewer notes, indented and brief (using symbols and/or upper case short words);
 - context displayed in form pane; and
 - hide speed bar on screen (with possiblity of it being available later if interviewers use the mouse more).

12 These screen design recommendations 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.

Recommendations for round two usability walkthrough testing

- 13 Test following features further in round two:
 - screen page number and question numbers able to be accessed from the status bar if required (eg. in resolving field queries with ABS Central Office staff);
 - content of help box;
 - test using the "finger pointing" symbol in comparison with the "arrow" symbol for interviewer; directions about prompt cards; and
 - test a different symbol for the help option if possible.

14 Maintain a similar procedure where interviewers work through scenarios, however incorporate the following points:

- evaluate screen design with both interviewers and non-interviewers;
- use an MPS supplementary survey to test interviewers reaction with a survey they are not as familiar with; and
- conduct some testing sessions outside in natural lighting conditions.

Reference

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.

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.

Appendix 1

Figure 1. Sample screen showing new layout features such as interviewer instructions, context information (screen A).

ann Arone Bengale Opton Heb		
4 2		
READ statement for the first person or if the respondent changes.	to be interviewed in the household,	
I would like to ask about last week Monday the 15th and ending Sund	, that is, the week starting lay the 21st of April.	
PRESS Enter to continue		
The state of the second st	LARSON AND A AND A DATA AND AND AND AND AND AND AND AND AND AN	
Whether working	Seeking PT work	
Whether working Introduction Worked last week Family business work	Whether seeking work Seeking FT work Seeking PT work Waiting to start work	
Whether working Worked last week Family business work Absent from work	Whether seeking work Seeking FT work Seeking PT work Waiting to start work	
Worked last week Family business work Absent from work	Whether seeking work Seeking FT work Seeking PT work Waiting to start work	
Whether working Morked last week Family business work Absent from work	Whether seeking work Seeking FT work Seeking PT work Waiting to start work	
Whether working Introduction Worked last week Family business work Absent from work	Whether seeking work Seeking FT work Seeking PT work Waiting to start work	
Whether working Worked last week Family business work Absent from work		

Figure 2. Sample screen showing Blaise default layout features such as interviewer instructions, context information (screen B).

The Northly Population Survey
fann groven Herigde Octore Help
interviewer; Read statement for the first person to be interviewed in the household, or if the respondent changes.
WOULD LIKE TO ASK ABOUT LAST WEEK, THAT IS, THE WEEK STARTING MONDAY THE 15TH AND ENDING SUNDAY THE 21ST OF APRIL.
Interviewer: Press Enter to continue
Production
20/153 Q18 LotFore-Wok-al-to

Figure 3. Sample screen showing new layout features such as interviewer instructions, context information (screen A).

-					
What was the main reas week?	on you worked les	s than 35 hours last			
⊂ 1. Holiday/Flextime/St ⊂ 2. Own illness or injur ⊂ 3. Standard work arra	udy/Personal reaso y/Sick leave ngements/Shift wo	ons rk/RDO			
4. Stood down/On she 5. Bad weather/Plant I 6. Began/left/lost job o 7. On strike/Locked o 8. Other	ort time/Insufficient oreakdown during week ut/Industrial disput	work e			
4. Stood down/On she 5. Bad weather/Plant I 6. Began/leftilost job o 7. On strike/Locked o 8. Other	ort time/Insufficient oreakdown Juring week ut/Industrial disput	e Time off	2	No	
4. Stood down/On she 5. Bad weather/Plant I 6. Began/left/lost job o 7. On strike/Locked o 8. Other Hours worked Hours worked	ort time/Insufficient oreakdown luring week ut/Industrial disput	e Time off Extra hours	2	No No	
4. Stood down/On she 5. Bad weather/Plant I 6. Began/left/lost job o 7. On strike/Locked o 8. Other Hours worked troduction	ort time/Insufficient preakdown luring week ut/Industrial disput	e Time off Extra hours Actual hours (main)	2	No No	
4. Stood down/On she 5. Bad weather/Plant I 6. Began/left/lost job o 7. On strike/Locked o 8. Other Hours worked Hours worked Mon 1 Whether worked Tue 2	Yes No	e Time off Extra hours Actual hours (main) Actual hours (all)	222	No No	
4. Stood down/On she 5. Bad weather/Plant I 6. Began/left/lost job o 7. On strike/Locked o 7. On strike/Locked o 8. Other Hours worked Hours worked Mon 1 Whether worked Mon 1 Whether worked Tue 2 Whether worked Wed 1	Ves Ves Ves Ves Ves	e Time off Extra hours Actual hours (main) Actual hours (ali) Usual hours	22	No No	
4. Stood down/On she 5. Bad weather/Plant I 6. Began/leftilost job o 7. On strike/Locked or 8. Other Hours worked Hours worked Mon 1 Wrether worked Mon 1 Wrether worked Tue 2 Wrether worked Tue 2	Yes No Yes No No	e Time off Extra hours Actual hours (main) Actual hours (ali) Usual hours Usual hours	2 2 25 35	No No	
4. Stood down/On she 5. Bad weather/Plant I 6. Began/leftilost job (7. On strike/Locked of 8. Other Hours worked	Ves No Yes No Yes No Yes	e Time off Extra hours Actual hours (main) Actual hours (ali) Usual hours Usual hours Usual hours Usual hours	2 2 25 35	No No	
4. Stood down/On she 5. Bad weather/Plant I 6. Began/left/lost job o 7. On strike/Locked o 7. On strike/Locked o 8. Other Hours worked Hours worked Hours worked Mon 1 Whether worked Tue 2 Whether worked Tue 2 Whether worked Tu 2 Whether worked Fn 1 Whether worked Sat 2	Ves No Yes No Yes No Yes No Yes No	e Time off Extra hours Actual hours (ali) Usual hours (ali) Usual hours Usual hours Usual hours	2 2 25 35	No No	

Figure 4. Sample screen showing Blaise default layout features such as interviewer instructions, context information (screen B).

	In Nonthly Population Survey		
Image: Standard work arrangements/Shift work/RDO C 6. Began/left/lost job during week C 7. On strike/Locked out/Industrial dispute C 8. Other C 8. Bed weather/Plant breakdown	Enne Aneven Merigade Options Help		
I. Holiday/Flextime/Study/Personal reasons C. 6. Began/left/lost job during week C. 7. On strike/Locked out/Industrial dispute Standard work arrangements/Shift work/RDO C. 8. Other C. 8. Other C. 8. Other C. 9. Other C. 9			
1. HolidayiFlextime/StudyiPersonal reasons 6. Began/left/lost job during week 7. On strike/Locked out/Industrial dispute 8. Other 3. Standard work arrangements/Shift work/RDO 8. Other 8. Other 4. Stood down/On short time/Insufficient work 5. Bad weather/Plant breakdown 043	WHAT WAS THE MAIN REASON YOU WORKED LESS T	THAN 36 HOURS LAST WEEK?	
043	1. Holiday/Flextime/Study/Personal reasons 2. Own illness or injury/Sick leave 3. Standard work arrangements/Shift work/RDO 4. Stood down/On short time/insufficient work 5. Bad weather/Plant breakdown	C 6. Began/left/lost job during week C 7. On strike/Locked out/Industrial dispute C 8. Other	_
	043		