

# Exploring the Use of Interactive Voice Response as a Population Health Tool

Stephen Ross Corkrey  
BSc (Hons) Adelaide

A thesis presented to the University of  
Newcastle in candidacy for the degree  
of Doctor of Philosophy.

July 2002

*I hereby certify that the work embodied in this thesis is the result of original research and has not been submitted for a higher degree to any other University or Institution.*

(Signed)

---

# Acknowledgements

I would like to thank my supervisor, Dr. Lynne Parkinson, for her unstinting support and intellectual integrity. I would also like to thank Associate Professor Kate D'Este for her critical examination of the thesis drafts, and Associate Professor Bob Gibberd for his advice and involvement.

I am also grateful to Judy Foulkes, Maria Rees and Sue Green, for volunteering their voices, Su Htun and Lucy Bates for numerous discussions about cervical screening, Penninah Oberdorfer for her interesting conversation, Amanda Turner for supplying her network skills, Karen Gillham for her assistance in obtaining hysterectomy rates, Sally Burrows and the other statisticians for their friendship, and the CATI interviewers, Megan Freund, Denise Kelly, Jenny Jackson, Beverly Parker, Jennifer Parsons, and Sally Mayhew, for their professionalism and indomitable humour. I am also grateful to the University for providing the support needed for the development of GEIS and to Dr S.J. Smith and Professor R.J. MacDonald, for their advice.

This project was made possible by grants from Hunter Medical Research, The University of Newcastle RMC scheme, and a University of Newcastle Postgraduate Scholarship, with additional operational funding from the Hunter Centre for Health Advancement.

Lastly, my greatest thanks are due to my wife, Judy Foulkes, whose support cannot be understated. Judy also edited the manuscript.

# Contents

<b>Acknowledgements</b>	<b>ii</b>
<b>List of Tables</b>	<b>vii</b>
<b>List of Figures</b>	<b>x</b>
<b>Synopsis</b>	<b>xiv</b>
<b>Thesis note</b>	<b>xvii</b>
<b>Glossary of abbreviations</b>	<b>xviii</b>
<b>1 Review of IVR studies 1989–1999</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Aims . . . . .	5
1.3 Methods . . . . .	5
1.4 Results and discussion . . . . .	6
1.4.1 Study area . . . . .	8
1.4.2 Touchphone prevalence . . . . .	9
1.4.3 Voice assessment . . . . .	11
1.4.4 Acceptability . . . . .	13
1.4.5 Response rate . . . . .	13
1.4.6 Sampling method . . . . .	14
1.4.7 Sample characteristics . . . . .	17
1.5 Conclusions . . . . .	20
<b>2 Asking sensitive questions</b>	<b>22</b>
2.1 Introduction . . . . .	22
2.2 Aims . . . . .	22
2.3 Burden of illness . . . . .	23
2.4 Sensitive questions . . . . .	24
2.5 Review of methods of asking sensitive questions . . . . .	26
2.5.1 Methods . . . . .	26
2.5.2 Results . . . . .	26
2.5.3 Conclusions . . . . .	42
2.6 An alternative method: Interactive voice response (IVR) . . . . .	42
2.7 Conclusions . . . . .	44



<b>3</b>	<b>GEIS</b>	<b>46</b>
3.1	Introduction . . . . .	46
3.2	Aims . . . . .	47
3.3	Using GEIS to create an interview . . . . .	47
3.4	Implementation of GEIS . . . . .	49
3.4.1	The GEIS windows . . . . .	49
3.4.2	Implementing IVR within GEIS . . . . .	59
3.4.3	Selecting interviewees . . . . .	60
3.4.4	Status codes . . . . .	61
3.4.5	Data quality considerations . . . . .	64
3.4.6	Call scheduling . . . . .	65
3.4.7	Call cost . . . . .	67
3.5	Data management in GEIS . . . . .	67
3.5.1	Data storage in GEIS . . . . .	67
3.5.2	Record locking . . . . .	68
3.5.3	Multiple data access . . . . .	69
3.6	GEIS scripts . . . . .	70
3.6.1	Script structure . . . . .	70
3.6.2	GEIS script item types . . . . .	71
3.6.3	Self-protection statements . . . . .	74
3.6.4	Answer-quoting in question texts . . . . .	76
3.7	Conclusions . . . . .	77
<b>4</b>	<b>Developing and testing the IVR system</b>	<b>78</b>
4.1	Introduction . . . . .	78
4.2	Aims . . . . .	79
4.3	Testing the outbound mode . . . . .	79
4.3.1	Introduction . . . . .	79
4.3.2	Pretest One aims . . . . .	80
4.3.3	Methods . . . . .	81
4.3.4	Results . . . . .	85
4.3.5	Discussion . . . . .	87
4.3.6	Conclusions of Pretest One . . . . .	92
4.4	Testing the inbound mode . . . . .	93
4.4.1	Introduction . . . . .	93
4.4.2	Pretest Two aims . . . . .	94
4.4.3	Methods . . . . .	95
4.4.4	Results . . . . .	99
4.4.5	Discussion . . . . .	103
4.4.6	Conclusions of Pretest Two . . . . .	104
4.5	Community pilot of the inbound and outbound methods . . . . .	104
4.5.1	Introduction . . . . .	104
4.5.2	Pilot One aims . . . . .	105
4.5.3	Methods . . . . .	105
4.5.4	Results . . . . .	110
4.5.5	Discussion . . . . .	115
4.5.6	Conclusions of Pilot One . . . . .	116
4.6	Community pilot of the Hybrid method . . . . .	117
4.6.1	Introduction . . . . .	117
4.6.2	Pilot Two aims . . . . .	118

4.6.3	Methods . . . . .	118
4.6.4	Results . . . . .	125
4.6.5	Discussion . . . . .	128
4.6.6	Conclusions of Pilot Two . . . . .	129
4.7	Conclusions . . . . .	130
<b>5</b>	<b>Telephone interviewing methods</b>	<b>131</b>
5.1	Introduction . . . . .	131
5.2	Aims . . . . .	132
5.3	Methods . . . . .	133
5.3.1	Design . . . . .	133
5.3.2	Setting . . . . .	133
5.3.3	Sample . . . . .	133
5.3.4	Procedure . . . . .	135
5.3.5	Apparatus . . . . .	140
5.3.6	Measures . . . . .	141
5.3.7	Analysis . . . . .	144
5.4	Results . . . . .	152
5.4.1	Comparisons between interview methods . . . . .	152
5.4.2	Comparison of the results to the National Drug Strategy Household Survey . . . . .	164
5.5	Discussion . . . . .	172
5.5.1	Comparisons between interview methods . . . . .	172
5.5.2	Comparison of the results to the National Drug Strategy Household Survey . . . . .	178
5.6	Conclusions . . . . .	179
<b>6</b>	<b>IVR brief intervention: A pilot study</b>	<b>181</b>
6.1	Introduction . . . . .	181
6.1.1	Cervical cancer burden of illness . . . . .	182
6.1.2	Brief interventions . . . . .	183
6.1.3	Strategies to increase cervical screening . . . . .	183
6.1.4	Use of IVR in brief interventions . . . . .	185
6.2	Aims . . . . .	187
6.3	Methods . . . . .	188
6.3.1	Design . . . . .	188
6.3.2	Setting . . . . .	188
6.3.3	Sample . . . . .	189
6.3.4	Procedure . . . . .	189
6.3.5	Apparatus . . . . .	194
6.3.6	Measures . . . . .	195
6.3.7	Analysis . . . . .	197
6.4	Results . . . . .	202
6.4.1	Participation rates . . . . .	203
6.4.2	Demography . . . . .	204
6.4.3	IVR self-report screening status . . . . .	206
6.4.4	Barrier messages . . . . .	207
6.4.5	Information messages . . . . .	207
6.4.6	IVR and follow-up CATI self-report screening status . . .	209
6.4.7	Acceptability . . . . .	213

6.4.8	Cost . . . . .	217
6.5	Discussion . . . . .	217
6.5.1	Participation rates . . . . .	217
6.5.2	Demography . . . . .	219
6.5.3	IVR self-report screening status . . . . .	219
6.5.4	Barrier messages . . . . .	219
6.5.5	Information messages . . . . .	220
6.5.6	IVR and follow-up CATI self-report screening status . . . . .	221
6.5.7	Acceptability . . . . .	223
6.5.8	Cost . . . . .	224
6.6	Conclusions . . . . .	224
<b>7</b>	<b>IVR brief intervention: A randomized controlled trial</b>	<b>226</b>
7.1	Introduction . . . . .	226
7.2	Aims . . . . .	227
7.3	Methods . . . . .	227
7.3.1	Design . . . . .	227
7.3.2	Setting . . . . .	228
7.3.3	Sample . . . . .	228
7.3.4	Procedure . . . . .	237
7.3.5	Apparatus . . . . .	242
7.3.6	Measures . . . . .	242
7.3.7	Analysis . . . . .	244
7.4	Results . . . . .	247
7.4.1	Participation rates . . . . .	247
7.4.2	Demography . . . . .	249
7.4.3	Efficacy . . . . .	250
7.4.4	Cost per additional screen . . . . .	252
7.5	Discussion . . . . .	253
7.5.1	Participation rates . . . . .	253
7.5.2	Demography . . . . .	255
7.5.3	Efficacy . . . . .	255
7.5.4	Cost per additional screen . . . . .	263
7.6	Conclusions . . . . .	265
<b>A</b>	<b>Appendix for Chapter 3</b>	<b>320</b>
A.1	GEIS compiler steps . . . . .	320
A.2	IVR equipment . . . . .	323
A.3	Interview STATUS variable codes . . . . .	323
A.4	The GEIS call-scheduling algorithm . . . . .	323
A.5	GEIS data sets and catalogues . . . . .	325
A.5.1	The DEFINIT Catalogue . . . . .	325
A.5.2	The DISTANCE data set . . . . .	325
A.5.3	The SCRIPT data set . . . . .	327
A.5.4	The CONFID data set . . . . .	329
A.5.5	The CONTROL data set . . . . .	329
A.5.6	The ANSWERS data set . . . . .	332
A.5.7	The PROTECT data set . . . . .	333
A.5.8	The CALLS data set . . . . .	333
A.5.9	The COMBRESP data set . . . . .	333

A.5.10 The local data sets . . . . .	336
A.5.11 The ILOG data set . . . . .	336
A.5.12 The INTRVS data set . . . . .	336
A.5.13 The STANDARD and PHRASES data sets . . . . .	336
A.5.14 The AUTOCALL data set . . . . .	337
A.6 The share server . . . . .	337
A.7 GEIS script syntax . . . . .	340
A.7.1 Item texts . . . . .	340
A.7.2 Answer-quoting in IVR . . . . .	340
A.7.3 Self-protection statements . . . . .	341
A.7.4 Names and functions of GEIS script items . . . . .	342
<b>B Appendix for Chapter 4</b>	<b>346</b>
B.1 Pretest One . . . . .	346
B.1.1 Sampling program for Pretest One . . . . .	346
B.1.2 Participant instructions . . . . .	347
B.2 Pretest Two . . . . .	353
B.2.1 Program to generate dial-in code from a user network name	353
B.2.2 Participant instructions . . . . .	354
B.3 Pilot One . . . . .	357
B.3.1 Sampling program for Pilot One . . . . .	357
B.3.2 Program to calculate distances from Wallsend . . . . .	359
B.3.3 Information letters and instruction sheets . . . . .	361
B.4 Pilot Two . . . . .	367
B.4.1 Sampling program for Pilot Two. . . . .	367
B.4.2 Information letter sent to participants . . . . .	369
B.4.3 Use of the AUTOCALL data set within the Hybrid method.	372
<b>C Appendix for Chapter 5</b>	<b>373</b>
C.1 The five-item AUDIT instrument . . . . .	373
C.2 Sampling program . . . . .	374
C.3 Letters to participants . . . . .	378
<b>D Appendix for Chapter 6</b>	<b>384</b>
D.1 Cessnock pilot sampling programs . . . . .	384
D.1.1 IVR phase sampling program . . . . .	384
D.1.2 CATI phase sampling program . . . . .	386
D.2 Cessnock pilot information letter . . . . .	387
D.3 Cessnock pilot script technical details . . . . .	387
<b>E Appendix for Chapter 7</b>	<b>391</b>
E.1 Sampling program for the IVR message . . . . .	391
E.2 Information letter . . . . .	400
E.3 Data request to the Health Insurance Commission . . . . .	403

# List of Tables

1.1	Articles and study areas for IVR studies: 1989–1999. . . . .	7
1.2	Touchphone prevalence, voice assessment, acceptability, and response rates for IVR Studies: 1989–1999. . . . .	10
1.3	Summary of the relationship between vocal characteristics and survey parameters. . . . .	12
1.4	Sampling method, subject remuneration, and subject reuse for IVR studies: 1989–1999. . . . .	15
1.5	Sample characteristics for IVR studies: 1989–1999. . . . .	18
2.1	Methods for asking sensitive questions and studies identified in relation to each method. . . . .	27
3.1	GEIS variables used to select interviewees. . . . .	62
3.2	Principle STATUS codes. . . . .	63
3.3	Example of STATUS codes in successive call attempts. . . . .	63
3.4	The major GEIS data sets. . . . .	68
3.5	The major GEIS script item types. . . . .	73
4.1	Pretest One assessment results summary. . . . .	88
4.2	Pretest One agreement with feeling comfortable with the system. . . . .	88
4.3	Pretest One summary of respondents' comments. . . . .	89
4.4	Number and duration of IVR messages in Pretest One and Pretest Two. . . . .	100
4.5	Pretest Two assessment results summary. . . . .	101
4.6	Pretest Two agreement with feeling comfortable with the system. . . . .	101
4.7	Pretest Two summary of respondents' comments. . . . .	102
4.8	Number and duration of IVR messages in Pretest One, Pretest Two, and Pilot One. . . . .	111
4.9	Pilot One interview outcomes by mode. . . . .	112
4.10	Pilot One IVR and follow-up interview outcomes. . . . .	113
4.11	Pilot One acceptability by mode and IVR interview outcome. . . . .	114
4.12	ARIA categories, scores, and descriptors. . . . .	119
4.13	Pilot Two CATI interview outcomes. . . . .	125
4.14	Number and duration of IVR messages in Pretest One, Pretest Two, Pilot One, and Pilot Two. . . . .	126
4.15	Pilot Two IVR interview acceptability questions. . . . .	127
4.16	Pilot Two demographic characteristics of follow-up cases. . . . .	128

5.1	Outcome rates and sample sizes by telephone interview method.	153
5.2	Unweighted respondent demographic composition for each nominal telephone interview method. . . . .	154
5.3	Weighted respondent demographic composition by nominal telephone interview method compared to the Australian population.	155
5.4	Mean number of calls by method. . . . .	157
5.5	Item non-response rate by method. . . . .	161
5.6	Mean cost by method. . . . .	161
5.7	Alcohol, marijuana, amphetamines and heroin consumption by method. . . . .	162
5.8	Median acceptability scores (ease, enjoyable, stressful, likeable) for the interview methods. . . . .	163
5.9	Preference and honesty ratings of interview methods. . . . .	164
5.10	Drinking status by method for males and females, compared to the NDSHS results. . . . .	166
5.11	Alcohol quantity by method for males and females, compared to the NDSHS results. . . . .	167
5.12	Frequency of alcohol consumption by method for males and females, compared to the NDSHS results. . . . .	169
5.13	Consuming marijuana by method for males and females, compared to the NDSHS results. . . . .	170
5.14	Consuming marijuana by method by age and sex, compared to the NDSHS results. . . . .	170
5.15	Consuming amphetamines by method for males and females, compared to the NDSHS results. . . . .	172
5.16	Consuming heroin by method for males and females, compared to the NDSHS results. . . . .	173
6.1	Brief intervention studies: 1998–2001. . . . .	184
6.2	Pilot study messages played to women by cervical screening status.	199
6.3	Pilot study barrier messages available to under-screened women.	200
6.4	Pilot study participation rates by the last domain completed. . .	204
6.5	Pilot study demographic composition of listeners to the IVR call from the IVR and CATI follow-up samples, and non-listeners in the CATI follow-up sample, compared to ABS census data. . . .	205
6.6	Pilot study agreement ( $\kappa$ ) between IVR and CATI demographic variables. . . . .	206
6.7	Pilot study barrier messages selected by age group of women. . .	208
6.8	Pilot study acceptability of the IVR call. . . . .	213
6.9	Pilot study preference for alternative IVR call topics. . . . .	215
6.10	Pilot study alternative health topics for IVR calls nominated by participants. . . . .	216
6.11	Over-report of Pap tests by women reported in cervical screening studies. . . . .	223
7.1	Intervention and control group assignment, number of eligible women, Index of Socio-Economic Disadvantage, and number of selected households for selected postcodes. . . . .	236
7.2	Participation rates by last domain completed and post code, and in comparison to the Cessnock pilot. . . . .	248

7.3	Age, education, marital status, country of birth, employment status, and household income of listeners to the IVR message compared to the ABS census. . . . .	249
7.4	Effect per $10^4$ women by age group, previous screening, and effect period. . . . .	253
7.5	Calculation of cost for the IVR brief intervention. . . . .	254
7.6	Methods used, effects, and effect determination methods observed in cervical screening intervention studies. . . . .	256
7.7	Age effects, screening status effects, and cost per additional screen observed in cervical screening intervention studies. . . . .	258
7.8	Estimated percentage effect rates for varying efficacy and listening percentages. . . . .	259
7.9	Cost per additional screening for cervical screening methods identified in literature search studies compared to this study. . . . .	264
A.1	A complete list of valid codes for the STATUS variable. . . . .	324
A.2	The structure of the DEFINIT catalogue. . . . .	326
A.3	The contents of the DISTANCE data set. . . . .	326
A.4	Structure of the SCRIPT data set. . . . .	327
A.5	Codes used by the SURVTYPE variable in the CONFID data set. . . . .	329
A.6	Structure of the CONFID data set. . . . .	329
A.7	Structure of the CONTROL data set. . . . .	330
A.8	Call codes used by the CALLRSLT variable in the CONTROL data set. . . . .	331
A.9	Call codes used by the ARESULT variable in the CONTROL data set. . . . .	332
A.10	Structure of the ANSWERS data set. . . . .	333
A.11	Structure of the PROTECT data set. . . . .	334
A.12	Structure of the CALLS data set. . . . .	334
A.13	Structure of the COMBRESP data set. . . . .	335
A.14	Structure of the ILOG data set. . . . .	336
A.15	Structure of the INTRVS data set. . . . .	337
A.16	Structure of the STANDARD data set. . . . .	338
A.17	Structure of the PHRASES data set. . . . .	338
A.18	Structure of the AUTOCALL data set. . . . .	339

# List of Figures

1.1	Typical IVR hardware set up on a personal computer. . . . .	2
1.2	Number of IVR and CATI studies published per year: 1989–1999. . . . .	8
3.1	The procedure to develop a script within GEIS. . . . .	48
3.2	The GEIS control panel. . . . .	50
3.3	The GEIS script import window. . . . .	51
3.4	The GEIS script import results window. . . . .	52
3.5	The GEIS phrase editor for IVR messages. . . . .	52
3.6	The GEIS compiler window. . . . .	53
3.7	The GEIS interviewer log-in window. . . . .	55
3.8	The GEIS respondent selection window showing pull-down list used to select a respondent. . . . .	55
3.9	The GEIS interviewing window. . . . .	56
3.10	An example of a GEIS report window. . . . .	57
3.11	The GEIS respondent status code maintenance window. . . . .	58
3.12	The GEIS output window. . . . .	59
3.13	IVR call scheduling loop. . . . .	66
3.14	Example of record-locking in multiple GEIS data sets. . . . .	69
3.15	Example of a GEIS script item. . . . .	71
3.16	Example of the logical relationships between four questions, their self-protection statements, and option-code values. . . . .	75
4.1	Pretest One frequency of the number of calls required to complete an interview and to contact a respondent. . . . .	86
4.2	Keypad of an Interquartz facility phone: Model 9837. . . . .	97
4.3	Keypad of an Interquartz facility phone: Model IQTEL300. . . . .	97
4.4	Pilot One general verbatim comments by IVR call outcome. . . . .	115
4.5	Pilot Two domains handled by the interviewer and GEIS running in IVR mode. . . . .	121
5.1	Assignment of interviewer or RVS and domains (technology, alco- hol, drugs, demography) for each of the methods CATI, Hybrid I, Hybrid II, and IVR. . . . .	137
5.2	Interview durations shown as box plots for the whole interview, the part conducted by the interviewer, and the part conducted by the RVS. . . . .	156
5.3	Percentage change in age by number of calls to complete interviews	157



5.4	Percentage change in proportion of female respondents by number of calls to complete interviews with $\pm 2.5\%$ limits. . . . .	158
5.5	Percentage change in AUDIT score by number of calls to complete interviews with $\pm 2.5\%$ limits. . . . .	158
5.6	Percentage change in proportion consuming alcohol by number of calls to complete interviews with $\pm 2.5\%$ limits. . . . .	159
5.7	Percentage change in proportion consuming marijuana by number of calls to complete interviews with $\pm 2.5\%$ limits. . . . .	159
5.8	Percentage change in proportion consuming amphetamines by number of calls to complete interviews with $\pm 2.5\%$ limits. . . . .	160
5.9	Percentage change in proportion consuming heroin by number of calls to complete interviews with $\pm 2.5\%$ limits. . . . .	160
6.1	Determination of cervical screening status used within the Pap status domain. . . . .	198
6.2	Pilot study texts of recorded questions. . . . .	210
6.3	Pilot study self-report screening status for successive interviews of 18–49 aged women from IVR and CATI follow-up. . . . .	211
6.4	Pilot study self-report screening status for successive interviews for 50–69 aged women from IVR and CATI follow-up. . . . .	212
7.1	The Hunter region and NSW, Australia. . . . .	229
7.2	Statistical local areas of the Hunter region. . . . .	230
7.3	Postcodes selected overlaid with statistical local areas of the Hunter region. . . . .	238
7.4	Effect by month following the intervention and screening status for women aged 20–49 years. . . . .	250
7.5	Effect by month following the intervention and screening status for women aged 50–69 years. . . . .	251
A.1	Hardware set up. . . . .	323
A.2	Program used to run the share server. . . . .	339
A.3	Example of a program used to access data on the share server. . . . .	339
A.4	Example of a question text and answer-quoting for an IVR interview. . . . .	341
B.1	E-mail sent to Pretest One participants. . . . .	348
B.2	Instructions sent to Pretest One participants. . . . .	349
B.3	Assessment sheet used by Pretest One participants. . . . .	350
B.4	E-mail sent to Pretest Two participants. . . . .	354
B.5	Assessment sheet sent to Pretest Two participants. . . . .	355
B.6	Information letter sent to Pilot One participants in the outbound calling mode. . . . .	362
B.7	Information letter sent to Pilot One participants in the inbound calling mode, page one. . . . .	364
B.8	Instruction sheet sent to all Pilot One participants. . . . .	366
B.9	Information letter sent to Pilot Two participants. . . . .	370
B.10	Instruction sheet sent to Pilot Two participants. . . . .	371
B.11	Example of states of the AUTOCALL data set during the stages of the Hybrid method. . . . .	372

C.1	Letter sent to households assigned to the CATI method. . . . .	379
C.2	Letter sent to households assigned to the Hybrid methods. . . . .	380
C.3	Instruction sheet included with letter sent to households assigned to the Hybrid methods. . . . .	381
C.4	Letter sent to households assigned to the IVR method. . . . .	382
C.5	Instruction sheet included with letter sent to households assigned to the IVR method. . . . .	383
D.1	Information letter sent to households in the Cessnock pilot. . . .	388
D.2	Instruction sheet sent to households in the Cessnock pilot. . . .	389
E.1	Information letter sent to households. . . . .	401
E.2	Instruction sheet sent to households. . . . .	402

# Synopsis

The research described in this thesis reviewed previous uses of Interactive Voice Response (IVR), developed appropriate software, and employed IVR to obtain self-report of sensitive issues in surveys and conduct brief public health interventions.

Chapter 1 introduces IVR and describes a systematic critical review of the use of IVR. IVR is a telephone interviewing technique where the human speaker is replaced by a high quality recorded interactive script to which the respondent provides answers by pressing the keys of a touch-telephone (touchphone). IVR has numerous advantages, including: economy, autonomy, confidentiality, access to certain population groups, improved data quality, standardised interviewing, multi-lingual interfaces, and detailed longitudinal assessments. Despite this, there have been few survey applications of IVR compared to alternative methods such as Computer-Assisted Telephone Interview (CATI). There has not been any evaluation of the use of IVR for asking sensitive questions in surveys or as a tool for health promotion at the community level.

A literature review, described in Chapter 2, was conducted to identify other methods of asking sensitive questions. The literature review identified 19 different methods. Those methods that were most successful were those that provided

the greatest degree of anonymity to the respondent. It was suggested that IVR may be a suitable method for community surveys.

As described in Chapter 3, a custom software called Generalized Electronic Interviewing System (GEIS) was developed. This provided both CATI and IVR interviewing modes. As described in Chapter 4, it was found that the response rate obtained using IVR was unacceptably low, and an alternative interviewing method, the Hybrid method was developed. In the Hybrid method the interview was initiated by the interviewer but completed using IVR with GEIS.

As described in Chapter 5, the IVR, CATI and Hybrid methods were used to investigate self-reported rates of alcohol and drug consumption within a telephone household survey of 2880 households. The self-report rates were compared to the National Drug Strategy Household Survey (NDSHS). Response rates did not differ significantly between the CATI and Hybrid methods, however the response rate for IVR was significantly less than the other methods. The Hybrid and IVR methods obtained significantly higher self-report consumption rates for alcohol and marijuana, and significantly higher hazardous drinking scores using Alcohol Use Disorders Identification Test (AUDIT).

In Chapter 6 a pilot of an IVR cervical screening brief advice interface is described. A total of 5000 households were contacted by the IVR system. The system randomly selected an eligible woman aged 18–69 per household and determined her cervical screening status. A total of 661 women listened to the IVR message. The IVR call was shown to be acceptable and inexpensive compared to a mail pamphlet intervention.

In Chapter 7 a randomized controlled trial of an IVR cervical screening brief advice involving 17,008 households is described. Cervical screening rate data

were obtained from the Health Insurance Commission (HIC) for a period spanning six months before and following the intervention. The cervical screening rate was increased in the intervention postcodes by 0.43% compared to the control postcodes, and the increase was greater for older women at 1.34%. This was a desirable outcome since this group is considered to be an at-risk group. The overall conclusion was that IVR technology could be feasibly used to contact women to deliver brief interventions aimed at increasing cervical screening rates, but the cost per screen was likely to be high. It is suggested that an IVR system could be linked to cervical screening registers to more directly and economically contact women, and provide an efficacious complementary approach to the existing letter reminder system.

# Thesis note

The main argument of this thesis is given in Chapters 1 through 7. Supporting materials for each chapter are shown in Appendices A through E.

Additional materials are provided in the document Supplementary Materials, included on the accompanying Compact Disk (CD) in the file **Supplementary-Materials.pdf**. The CD is inserted inside the back cover of the thesis. Sections of the supporting materials are referred to within the body of the thesis and these take the form “Supplementary Materials Section 5.1.1”. In particular, the survey scripts are included in the Supplementary Materials document due to their large size. This file is in Portable Data Format and may be viewed or printed with an Adobe® Acrobat® viewer or equivalent.

The GEIS software and its instruction manual are provided on the CD.

# Glossary of abbreviations

**AAPOR** American Association for Public Opinion Research.

**ABS** Australian Bureau of Statistics.

**AIDS** Acquired Immune Deficiency Syndrome.

**ARIA** Accessibility/Remoteness Index of Australia.

**ASGC** Australian Standard Geographical Classification.

**AUDIT** Alcohol Use Disorders Identification Test.

**CAPI** Computer-Assisted Personal-Interview.

**CASI** Computer-Assisted Self-Interview.

**CATI** Computer-Assisted Telephone Interview.

**CD** Census Collection District.

**CDRGP** Context-Determined Rule-Generated Pseudonym.

**GEIS** Generalized Electronic Interviewing System.

**HAHS** Hunter Area Health Service.

**HAREC** Hunter Area Research Ethics Committee.

**HCHA** Hunter Centre for Health Advancement.

**HIC** Health Insurance Commission.

**HREC** University of Newcastle Human Research Ethics Committee.

**IVR** Interactive Voice Response.

**LGA** Local Government Area.

**NDSHS** National Drug Strategy Household Survey.

**NLI** ABS National Localities Index.

**NSW** New South Wales.

**NWPC** Number of women aged 18–69 in each postcode.

**PABX** Private Automatic Branch Exchange.

**PCA** Principal Components Analysis

**PTR** Pap Test Register.

**RVS** Recorded Voice System.

**SAQs** Self-Administered Questionnaires.

**SD** Statistical Division.

**SED** Index of Socio-Economic Disadvantage.

**SEIFA** Socio-Economic Indexes for Areas.

**SLA** Statistical Local Area.

**SPS** Self-Protection Statement.

**SSD** Statistical Subdivision.

**STD** Subscriber Trunk Dialling.

**S/T** States and Territories.



# Chapter 1

## Interactive Voice Response: Review of studies 1989–1999\*

### 1.1 Introduction

The need for accurate data on behaviours of most interest to public health research has led to a variety of data collection techniques, the most common of which is the survey. Numerous methods have been used to collect survey data including face-to-face, self-administered, mail, and telephone modes. Interactive Voice Response (IVR) is a relatively recent development in telephone surveying, which has yet to be formally tested in population surveys. This chapter discusses the range of uses IVR has had to date, critically reviews the work done so far, and identifies areas where further work is required, particularly in respect to population surveys.

In IVR, the human speaker is replaced by a high quality recorded interactive script to which the respondent provides answers by pressing the keys of a

---

\*A paper based on material in this chapter has been accepted for publication in *Behavior Research Methods, Instruments, & Computers*

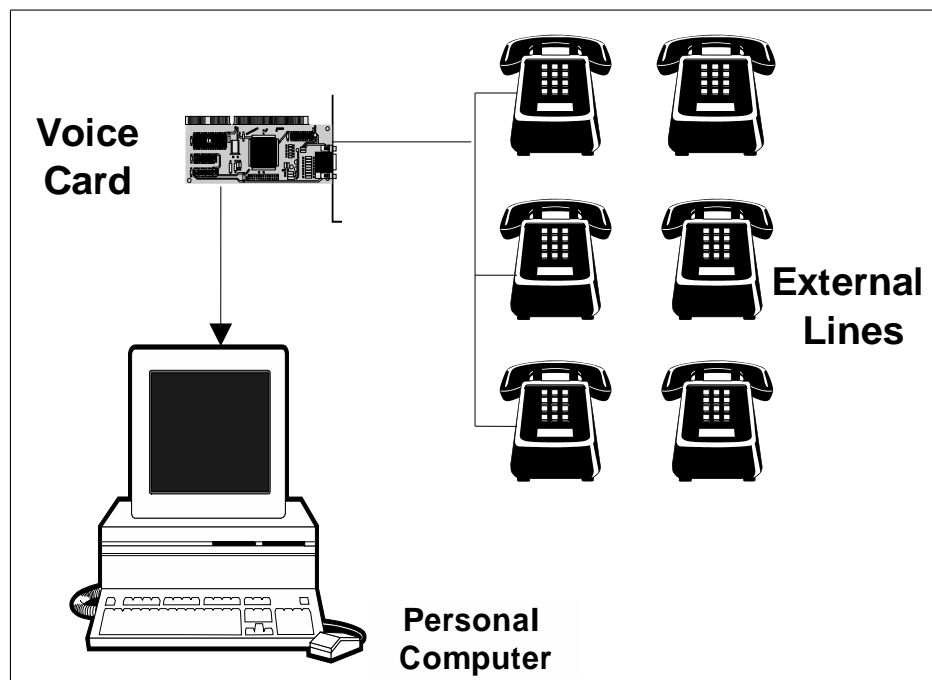


Figure 1.1: Typical IVR hardware set up on a personal computer.

touch-telephone (touchphone). It differs from a Computer-Assisted Telephone Interview (CATI) by lacking an interviewer to read the questions and enter the answers into the computer.

An IVR system typically involves running software on a personal computer in which a voice card has been installed (see Figure 1.1). The voice card is connected to one or more telephone lines and allows the computer to make or receive calls, play or record speech, or respond to keys pressed. The computer may ask questions of the caller by playing sound files that contain recorded questions. The caller responds by pressing keys on their touchphone.

Many attributes of CATI are shared by IVR, including those listed by Nicholls II<sup>1</sup>:

- Skipping or branching between questions can be automatically determined based on answers to earlier questions;

- Options and question order can be randomized;
- Interviews can be interrupted and resumed at a later time;
- Questions can have single choice responses as well as open-ended responses;
- Responses can be checked as they are entered for consistency with previous answers and using range limits;
- Feedback can be provided on inappropriate responses;
- Calls can be automatically scheduled; and
- Interviews and record-keeping are automatic.

However, IVR can offer considerable advantages over existing survey methods:

- An IVR system can run automatically and continuously;
- Respondents may proceed through an interview at their own rate;
- With the involvement of field workers, and the use of cellular telephones, IVR can allow access to hard to reach groups, such as sex industry workers, drug users, and homeless people (e.g. Alemagno et al.<sup>2</sup>);
- Data entry errors are reduced since the data is directly keyed in by the respondent;
- There is no interviewer bias and inter-rater reliability is irrelevant;
- There is an opportunity for interview standardization through the use of high quality recorded voices and standardized pronunciation;

- IVR benefits from the use of audio recordings to present questions, which could reduce respondent literacy difficulties<sup>3,4</sup> and allow access to hard to reach groups such as non-English speakers<sup>5–7</sup>.
- Since IVR uses a standardized interface and any assessment is conducted by a computer, interview accuracy is improved;
- The cost per interview is likely to be small<sup>8</sup>;
- More precise and detailed longitudinal assessments should be possible<sup>9,10</sup>; and
- Its greater confidentiality may result in a lower response bias<sup>3</sup>.

Other recent methods include web and e-mail surveys that may share many of the above attributes. However, these are currently limited by low response rates<sup>11</sup> and poor frame coverage for general populations<sup>12</sup>, although a mixed-mode approach that combines e-mail with conventional mail questionnaires may reduce coverage error<sup>13</sup>.

There is the potential to use IVR to conduct automated screenings of large populations to identify particular target groups. These interviews may then be followed-up in more detail using other means, once informed consent has been given. The complete anonymity and increased perception of confidentiality provided by IVR may also have particular benefit when used to assess sensitive behaviours, as more confidential interviewing methods have been shown to produce greater reporting levels of sensitive behaviours<sup>3</sup>.

## 1.2 Aims

This chapter critically reviews the literature on Interactive Voice Response applications from 1989–1999. The focus of the review is on factors that may affect its acceptability, reliability, and feasibility, particularly in respect to its potential for population surveys and population health research.

The aims of this chapter are to describe from the studies:

1. Areas where IVR has been applied;
2. Prevalence of touchphones;
3. Voice assessment methods;
4. Acceptability measures;
5. Response rates for survey applications;
6. Sampling methods; and
7. Sample characteristics.

## 1.3 Methods

The business, medical and statistical literature were examined for IVR studies by searching six online databases: Psycinfo<sup>14</sup>, Medline<sup>15</sup>, Sociofile<sup>16</sup>, Health Reference Center<sup>17</sup>, Current index to statistics<sup>18</sup>, ABI Inform<sup>19</sup>, IDEAL<sup>20</sup>, and Social Sciences Citation Index<sup>21</sup>. The following search terms were used either solely or in combination: “interactive voice response, bias, IVR, telephone, interview, computer, computer assisted, CATI, interview mode, data collection

method, face to face, survey, personal, response, post, mail, response effect, mode effect, data grabber, touch tone”.

Abstracts of references located by the search were examined and only those making reference to Interactive Voice Response (IVR) or a synonym were retained. Articles discussing the use of speech recognition or those that referred only to technological descriptions or marketing applications were discarded. Only articles in peer-reviewed journals were retained.

Full copies of all relevant articles published between 1989–1999 were obtained. The bibliographies of these articles were searched for further references and then treated as described above. The process terminated when no more articles were identified.

All articles were then coded using a data extraction sheet with the headings: study area, touchphone prevalence, voice assessment, acceptability, response rate, sampling method, and sample characteristics (headings are defined in *Results and Discussion* (Section 1.4)).

Two of the databases, Medline<sup>15</sup> and Current index to statistics<sup>18</sup>, were also examined for CATI papers published in journals using the search terms: “Computer Assisted Telephone, CATT”.

## 1.4 Results and discussion

The identified studies are shown in Table 1.1. The reviewed studies described IVR applications from 1989 onwards and no earlier work in refereed journals was known, except for a technical description of a clinical trial<sup>51</sup> published in 1977. Since 1994, the rate of publication of IVR articles only modestly increased,

Table 1.1: Articles and study areas for IVR studies: 1989–1999.

<i>Study</i>	<i>No.</i> <sup>a</sup>	<i>Study Area</i>
Havice(1989)	8	Media surveys.
Havice(1990)	22	Survey refusal rates.
Havice(1990)	23	Survey of student accommodation living conditions.
Havice and Banks(1991)	24	Media surveys.
Leirer et al.(1991)	25	Medication nonadherence in the elderly.
Phipps and Tupek(1991)	26	Business survey.
Patel and Babbs(1992)	27	In-home symptom monitoring for cardiac patients.
Baer et al.(1993)	28	Assessment of obsessive compulsive disorder.
Alemi et al.(1994)	29	Monitoring drug client treatment after-care.
Linkins et al.(1994)	30	Reminders for childhood immunization.
Tanke and Leirer(1994)	31	Automatic reminders for tuberculosis clinic attendance.
Baer et al.(1995)	32	Assessment of depression risk.
Mundt et al.(1995) <sup>b</sup>	10	Longitudinal monitoring of alcohol consumption.
Mundt et al.(1995) <sup>b</sup>	9	Longitudinal monitoring of tobacco and alcohol consumption.
Perrine et al.(1995) <sup>b</sup>	33	Monitoring of alcohol consumption.
Schneider et al.(1995)	34	Automatic smoking self treatment cessation program.
Searles et al.(1995) <sup>b</sup>	35	Longitudinal monitoring of alcohol consumption.
Albisser et al.(1996)	36	Automatic diabetes monitoring and feedback.
Alemagno et al.(1996)	2	Substance abuse survey among the homeless.
Alemi et al.(1996)	37	Support system for pregnant women using cocaine.
Harding et al.(1997)	38	Symptoms monitoring for irritable bowel syndrome.
Kobak et al.(1997)	39	Assessment of psychological conditions.
Kobak et al.(1997)	40	Assessment of psychological conditions.
Mundt et al.(1997)	41	Psychological performance assessment.
Piette and Mah(1997)	42	Symptom monitoring for diabetes.
Greist et al.(1998) <sup>c</sup>	43	Obsessive compulsive disorder treatment program.
Kadison et al.(1998)	44	Assessment of breast cancer risk.
Marks et al.(1998) <sup>c</sup>	45	Obsessive compulsive disorder treatment program.
Meneghini et al.(1998)	46	Diabetes management system.
Mundt et al.(1998)	47	Depression risk assessment.
Osgood-Hynes et al.(1998)	48	Depression self-treatment program.
Bachofen et al.(1999)	49	Obsessive compulsive disorder self-treatment program.
Kobak et al.(1999)	50	Assessment of depression and anxiety.

<sup>a</sup> Bibliography number.<sup>b,c</sup> These studies reported on aspects of the same work.

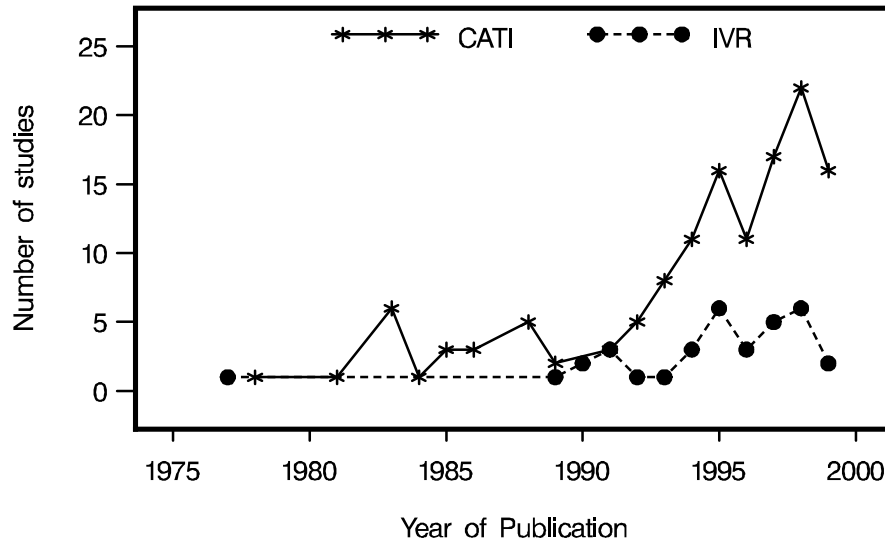


Figure 1.2: Number of IVR and CATI studies published per year: 1989–1999.

but since 1992 the number of published CATI studies increased markedly (Figure 1.2). The scarcity of IVR applications probably resulted from its more recent development and it seemed likely that the number of IVR applications would increase as the technology matured. However, application of IVR to surveys had been infrequent with only five (15%) of the reviewed studies<sup>8,22–24,26</sup> being in this area (See Table 1.1). Reasons for this are addressed in this chapter.

#### 1.4.1 Study area

Despite its apparent advantages, IVR had been applied in limited study areas. In the reviewed studies IVR was used primarily in health areas, for example, alcohol and drug monitoring, as well as survey applications to non-health areas (Table 1.1). The health-related studies principally described applications that were algorithmic, such as usage of the Yale-Brown Obsessive Compulsive scale<sup>28</sup>, or provided services, such as the Breast Cancer Telephone Risk Assessment system<sup>44</sup> or insulin monitoring<sup>36</sup>.



Such applications presupposed an interested subject. Although there were several passive systems described, no studies actively screened individuals for particular characteristics, such as health risks. This was despite the apparent advantage of IVR to conduct automated and economical screening for conditions such as alcohol abuse. The only studies that sought estimates of prevalence of specific characteristics were opinion surveys.

Therefore the usefulness of IVR in population surveys and screening programs was yet to be explored.

### 1.4.2 Touchphone prevalence

Since IVR usually requires the subject to have access to a touchphone, the prevalence of touchphones was of particular interest (see Table 1.2). Not all touchphones might be acceptable, since some might not be appropriately configured or use incompatible technology.

Where touchphones are common an IVR system might provide a useful service. In Australia, non-ownership of telephones was associated with households with lower incomes, the elderly, or young adults<sup>52</sup>. It might be speculated that touchphones have similar socioeconomic associations.

The prevalence of touchphones in households was unimportant to some studies. For example, Linkins et al.<sup>30</sup>, Tanke and Leirer<sup>31</sup>, and Leirer et al.<sup>25</sup> described reminder services that would work equally well with rotary and touchphones; and Alemagno et al.<sup>2</sup> used outworkers equipped with cellular phones. Only a small number of studies provided estimates of the prevalence of touchphones, but these were all obtained from anecdotal sources, such as contacts in telephone companies. Havice<sup>22</sup> reported a touchphone prevalence of 95%, while

Table 1.2: Touchphone prevalence, voice assessment, acceptability, and response rates for IVR Studies: 1989–1999<sup>a</sup>.

<i>No.</i> <sup>b</sup>	<i>Touchphone</i> <sup>c</sup>	<i>Voice</i>		<i>Acceptability</i> <sup>d</sup>	<i>Response rate</i>
		<i>Assessed</i> <sup>e</sup>	<i>Number</i> <sup>f</sup>		
8	-	no	1	-	7.5
22	95	no	-	-	6.8
23	-	no	1	-	18
24	75–98	no	1	-	-
25	-	no	1	-	
26	-	yes	-	yes	100
27	-	no	-	-	
28	-	no	-	yes	
29	-	no	-	yes	
30		no	-	-	
31		no	> 1	yes	
32	-	yes	2	yes	
10	-	no	-	-	
9	-	no	-	-	
33	-	no	-	-	
34	-	no	1	-	
35	-	no	-	-	
36	-	no	-	-	
2	-	no	-	-	
37	-		> 1	-	
38	-	no	-	yes	
39	-	no	-	yes	
40	-	no	-	-	
41	-	no	-	-	
42	-	no	1	yes	
43	-	no	-	yes	
44	-	no	-	yes	
45	-	no	-	-	
46	-	no	1	-	
47	-	no	-	-	
48	-	no	-	yes	
49	-	no	-	-	
50	-	no	-	yes	

<sup>a</sup> A blank cell indicates that the characteristic was not applicable to the variable. A dash (-) indicates that the characteristic was not reported.

<sup>b</sup> Bibliography number.

<sup>c</sup> Percentage touchphone prevalence.

<sup>d</sup> Was acceptability measured?

<sup>e</sup> Were voices used in the study assessed?

<sup>f</sup> The number of voices used in the study.

Havice and Banks<sup>24</sup> reported estimates between 75% and 98%. None of the remaining studies reported a touchphone prevalence.

It was concluded that touchphone prevalence and its relationship to the demographic composition of populations was still not known.

### 1.4.3 Voice assessment

One obvious aspect of interest to IVR research was the choice of voice. The voice is the only means by which an IVR system can engage the subject's interest sufficiently long to meet a study's objectives.

Despite the importance of voice, only Baer et al.<sup>32</sup> and Phipps and Tupek<sup>26</sup> assessed voices used by their systems. Baer et al.<sup>32</sup> did this by means of a focus group, and Phipps and Tupek<sup>26</sup> used a follow-up CATI.

Most studies did not report how many speakers were used to record the voices, but seven (21%) studies<sup>8,23–25,34,42,46</sup> used a single voice. Baer et al.<sup>32</sup> alternated the voices of a male and a female professional actor, reportedly to maintain the subject's interest. Meneghini et al.<sup>46</sup> used a synthetic voice. Tanke and Leirer<sup>31</sup> used an unstated number of multiple female voices speaking in different languages. Alemi et al.<sup>37</sup> used ad hoc recordings by members of staff.

In general, the studies failed to show any consideration of the standardization of voices, any effort in voice characterization, any measure of the possible effect of changing the number of speakers within scripts, or the choice of voice on the reporting levels and response rates.

This was despite work in the survey literature (Table 1.3) indicating that the voice quality of interviewers could affect response rates<sup>53,54</sup>, reporting rates<sup>55</sup>, and the perception of credibility<sup>56</sup>. Self-report rates can be influenced by inter-

Table 1.3: Summary of the relationship between vocal characteristics and survey parameters.

<i>Characteristic</i>	<i>Effect of voice</i>	<i>No.<sup>a</sup></i>
Response rates	Interviewers with high pitch variation, fast, loud voices, and clear pronunciation have higher response rates.	53,54
Self-report rates	Self-report rates can be influenced by interviewers dropping or raising the tone of their voices at the end of questions.	55
Interviewer credibility	Loudness increases perception of credibility, but intonation only increases perception of credibility if the subject feels personally involved in the discussion.	56
Interviewer persuasion	A faster speaking pace is perceived as being more persuasive than a slower pace.	57,58
Interviewer confidence	A soft speaker may be perceived as less assured than a speaker with a high volume.	59

<sup>a</sup> Bibliography number.

viewers dropping or raising the tone of their voices at the end of questions<sup>55</sup>. A faster speaking pace might be perceived as being more persuasive than a slow pace<sup>57,58</sup>, and a soft speaker might be perceived as less assured than a speaker with a high volume<sup>59</sup>.

A closer attention to voice would have two benefits:

1. Standard assessments of voices using suitable scales<sup>53,60</sup> would allow greater comparability of studies; and
2. IVR might also allow the effect of voice type on response rate to be systematically explored in a population survey context.

When IVR is employed, standard assessment of voices is essential. Oksenberg et al.<sup>53</sup> have developed a scale that would be useful for this purpose.

#### 1.4.4 Acceptability

The use of IVR is likely to be novel to most people and therefore its acceptability needs to be carefully assessed. It might be sufficient to assess new IVR applications by whether the subjects found it acceptable, but when comparing IVR to another method, comparative acceptability should be measured. In order to assess the degree of acceptability between different studies there needs to be some equivalence in methods.

Only twelve (36%) studies<sup>26,28,29,31,32,38,39,42–44,48,50</sup> attempted to measure acceptability (Table 1.2), and there was little agreement in the methods used. Of these, three (25%) compared the acceptability of IVR with other methods<sup>32,39,50</sup>. Baer et al.<sup>32</sup>, Kobak et al.<sup>39</sup>, and Kobak et al.<sup>50</sup> found that IVR was less preferred than alternative methods (face-to-face or CATI). However, Kobak et al.<sup>39</sup> and Kobak et al.<sup>50</sup> also found that compared to the other methods, IVR was the least embarrassing method. The remaining nine studies all reported satisfaction with IVR, but without comparison to other methods.

Accordingly, the acceptability of IVR was an issue still to be explored.

#### 1.4.5 Response rate

Response rate is important in survey applications<sup>61</sup>. Even if a random sampling method is used, obtaining a very low response rate can produce the equivalent of a volunteer sample. The effect is to widen the uncertainties of estimates in proportion to the non-response rate<sup>62</sup>. For IVR to be considered as a viable survey method, an estimate of the expected response rate is needed.

Of the studies describing survey applications, Havice and Banks<sup>24</sup> did not report a response rate nor the number of attempted contacts, and Alemagno

et al.<sup>2</sup> used a volunteer sample for which a response rate is meaningless (Table 1.2).

The response rate obtained by Havice<sup>8</sup> for an IVR survey on media ratings was 7.5% and for two CATI surveys was 55.2% and 74.1%. Although these were apparently quite discrepant, the low rate for IVR was predominantly due to the automated system not attempting to recontact unanswered numbers. Therefore, this rate was not a useful measure of what might be achieved when using IVR with follow-up of unanswered numbers.

Havice<sup>22</sup> reported a response rate of 6.8% for a survey on an unstated topic, while Havice<sup>23</sup> reported a response rate of 18% for a survey on housing.

Apart from one business survey<sup>26</sup>, it appeared that the response rates with IVR were unusually low and variable. It was concluded that reasonable estimates of response rates were not available from past research.

### 1.4.6 Sampling method

The use of non-random sampling methods affects the generality of conclusions and statistical test validity, and can prevent the estimation of the response rate that may be obtained by a survey method. However, many reviewed studies did not use formal random samples. There were nineteen (57%) studies that used convenience samples<sup>9,10,25,27–31,35–42,46,47,49</sup>, nine (27%) that used volunteer samples<sup>2,32–34,43–45,48,50</sup> and five (15%) that used random sampling<sup>8,22–24,26</sup> (Table 1.4).

Apart from a business survey<sup>26</sup>, studies in which calls could be made into an IVR system suffered from lack of randomness due to self-selection. Some studies required subjects to ring into the system to obtain assessments. For example,

Table 1.4: Sampling method, subject remuneration, and subject reuse for IVR studies: 1989–1999<sup>a</sup>.

<i>No.</i> <sup>b</sup>	<i>Method</i> <sup>c</sup>	<i>Remuneration</i> <sup>d</sup>	<i>Reused</i> <sup>e</sup>
8	R	no	no
22	R	no	no
23	R	no	no
24	R	no	no
25	C	yes	no
26	R	no	
27	C	no	no
28	C	no	no
29	C	yes	no
30	C	no	no
31	C	no	no
32	V	no	no
10	C	yes	yes
9	C	yes	yes
33	V	yes	no
34	V	no	no
35	C	yes	yes
36	C	no	no
2	V	yes	no
37	C	no	no
38	C	no	no
39	C	yes	no
40	C	yes	yes
41	C	yes	yes
42	C	no	no
43	V	no	no
44	V	no	no
45	V	no	no
46	C	no	no
47	C	no	yes
48	V	no	no
49	C	no	no
50	V	no	no

<sup>a</sup> A blank cell indicates that the characteristic was not applicable to the variable.

<sup>b</sup> Bibliography number.

<sup>c</sup> Sampling method: R=Random sample; V=volunteer sample; C=convenience sample.

<sup>d</sup> Were subjects remunerated?

<sup>e</sup> Were subjects reused from earlier studies?

Baer et al.<sup>32</sup> used IVR to detect depression and Perrine et al.<sup>33</sup> attempted to measure alcohol consumption. Both studies used volunteer samples. Having callers dial into an IVR system may have required considerable commitment by the subjects, restricting the generality of their application.

A total of seven (70%) of the ten studies<sup>8,22–25,27–31</sup> in which the systems initiated the calls also used convenience rather than random samples.

Where possible, populations should be sampled randomly to allow for unbiased and generalizable results. For example, a study of a service meant for diabetic patients should use a random sample from a population of diabetic patients. Only the studies by Havice<sup>8,22,23</sup> and Havice and Banks<sup>24</sup> used random sampling.

A total of eight (42%) of the studies using convenience samples<sup>9,10,25,29,35,39–41</sup> and two (22%) of those using volunteer samples<sup>2,33</sup> also remunerated their subjects for their cooperation (Table 1.4). Remunerated subjects are presumably more likely to be cooperative than unpaid subjects, so a study using remuneration may provide overly optimistic evidence in favour of IVR. Some studies did not report the nature or degree of remuneration, but of those that did, amounts ranged from \$U25 to \$US750.

A total of six (18%) studies<sup>9,10,35,40,41,47</sup> reused subjects from earlier studies, and reused subjects were more likely to be remunerated. This suggested that subjects of some studies were sourced from regularly-used pools. Members of such pools were likely to systematically differ from subjects selected at random from a population due to their experience of scientific trials and motivation to participate, again severely limiting the generalizability of the study outcomes.



### 1.4.7 Sample characteristics

The demographic composition of a sample is an important aid to determining whether the results of a study can be generalized to the wider population or target group. Five (15%) studies<sup>22,34,36,38,46</sup> did not provide demographic compositions of their samples, others deliberately restricted their samples to all male or all female, and some used samples with unusual demographic compositions (Table 1.5).

Two (6%) studies, Alemi et al.<sup>37</sup> and Kadison et al.<sup>44</sup>, used entirely female samples because they concerned breast cancer and pregnancy. Three (9%) studies, Mundt et al.<sup>9,10</sup>, and Searles et al.<sup>35</sup>, used all male samples to examine alcohol and tobacco consumption. The remaining studies did not report whether their samples were representative of the reference populations. Alemagno et al.<sup>2</sup> used a sample that was 76% male, which could be representative of the population under study (homeless people), although this was not reported. This lack of definition made the reported demographic compositions difficult to interpret. Demographic characteristics should routinely be compared to local population expectations, but none of the studies provided such comparisons.

Most studies did not report comparisons of demographic characteristics between IVR and other modes (Table 1.5). Since many studies used unusual samples and did not attempt to confirm the representativeness of their samples, no conclusions could be made about the applicability of IVR systems to differing demographic groups.

Where possible, studies should report whether the demographic characteristics of participants differ from non-participants, and whether those who prefer IVR differ from those who prefer an alternative. Few of the reviewed studies

Table 1.5: Sample characteristics for IVR studies: 1989–1999<sup>a</sup>.

<i>No.</i> <sup>b</sup>	<i>n</i> <sup>c</sup>	<i>Age</i>		<i>Males</i> <sup>d</sup>	<i>Mar.</i> <sup>e</sup>	<i>Emp.</i> <sup>f</sup>	<i>Cmp.</i> <sup>g</sup>
		<i>Median</i>	<i>Mean</i>				
8	575	28	35	38	-	-	yes
22	424	-	-	-	-	-	no
23	1228	-	-	51	-	-	yes
24	338, 147	-	37, 35	40, 41	-	-	yes
25	16	-	71	45	-	-	no
26	465						
27	3	-	-	67	-	-	no
28	18	-	38	44	-	-	no
29	42	-	38	48	-	-	no
30	8002	-	-	58			yes
31	2008	19	-	54	-	-	yes
32	1812	-	37, 24	46	-	-	no
10	51	28	-	100	33	88	no
9	51	28	-	100	33	88	no
33	31	36	-	50	100	94	no
34	37, 975	-	-	-	-	-	no
35	51	28	-	100	33	88	no
36	204	-	-	-	-	-	no
2	207	-	-	76	7	32	no
37	179	27	27	0	-	31	no
38	270, 370	-	-	-	-	-	no
39	51	-	44	43	-	10	no
40	200	-	41	33	-	68	no
41	10, 7	33	-	100	42	100	no
42	74	-	-	95	53	11	yes
43	40	-	35	53	-	-	no
44	343	35	-	0	-	100	no
45	63	-	33	54	-	-	no
46	184	-	-	-	-	-	no
47	367	-	36	35	-	60	no
48	41	-	42	29	44	-	no
49	21	-	31	57	-	-	no
50	113, 72	-	44, 42	-	-	-	no

<sup>a</sup> A blank cell indicates that the characteristic was not applicable to the variable.

A dash (-) indicates that the characteristic was not reported.

<sup>b</sup> Bibliography number.<sup>c</sup> Sample Size.<sup>d</sup> Percentage of males in the study's sample.<sup>e</sup> Percentage of subjects in the study's sample who were married.<sup>f</sup> Percentage of subjects in the study's sample who were employed.<sup>g</sup> Were the sample demographic characteristics compared between IVR and another mode?

attempted this or did so scientifically.

Linkins et al.<sup>30</sup> reported that the demographic breakdown of respondents in their IVR and control groups did not differ significantly using a sample size of 8002. Tanke and Leirer<sup>31</sup> found no significant effect due to age or sex in the effectiveness of the IVR system using a sample size of 2008. Piette and Mah<sup>42</sup> reported, without statistical tests, some descriptive data suggesting that older subjects were more likely to listen to at least two health promotion IVR messages, but this could also have been due to older people hanging up more slowly. Although Albisser et al.<sup>36</sup> did not present any demographic characteristics of participants in an insulin monitoring system, they reported, without any statistical test, that older participants were slower to learn the use of the system.

Havice and Banks<sup>24</sup> compared a pair of surveys, each consisting of two arms: IVR and CATI. The respondents' age and education for IVR participants were both significantly less than CATI participants. However, since the response rates were not reported these comparisons were not very useful.

Havice<sup>8</sup> found that the reported age and sex of respondents from an IVR survey was comparable to two CATI surveys, but did not provide any statistical test. In addition, the IVR and CATI surveys covered different topics, and ages were recorded in different ways in each mode making comparisons difficult.

While Havice<sup>23</sup> reported that the results of an IVR survey of university students were comparable to official University records, the age data from the university records was calculated using years enrolled rather than the actual age, making this comparison unreliable. The gender ratio however did not appear to differ significantly within the limits of the sample size ( $n=224$ ).

Given these limitations, the reviewed studies gave little or no indication of the applicability of IVR in the wider community.

## 1.5 Conclusions

A *prima facie* impression of IVR suggested it might be useful in population surveys and population health studies since it shared many characteristics with CATI as described in Section 1.1, such as automatic skipping and response validation, while also possessing certain advantages, particularly automatic operation, greater economy, and standardized interviews. It was expected that IVR should be especially useful for assessing sensitive behaviours since it is likely to be seen as a more confidential method by subjects.

Only five (15%) previous studies<sup>8,22–24,26</sup> attempted to use IVR for surveys. IVR has tended to be used in applications that required considerable motivation from subjects and typically involved algorithmic or passive applications.

Most projects ignored issues vital to a scientific exploration of IVR, such as the prevalence of touchphones, characterization and standardization of voice, and standard measures of acceptability. It was unclear what response rates IVR might obtain since the response rates reported by survey studies were quite variable with some being very low.

Many studies were unhelpful in evaluating IVR for survey applications since they used samples that were distinctly non-random, probably unrepresentative or undescribed, or simply too small, any of which might have invalidated the analyses reported therein.

It was clear that for IVR work to progress these deficits needed to be ad-

dressed. There needed to be further work to explore the areas in which IVR may be effectively applied, such as, for example, active screening of individuals in a household setting for health-related risks or interventions to modify risky-behaviours in population settings. Further work was also required to assess response rates and acceptability in a population survey context, while also using standard methods for voice characterization, acceptability, and sample selection.

It was thought that a potential application for IVR would be to ask questions about sensitive issues, such as risky health-related behaviours within a survey context, an area not covered by the reviewed studies. This and other methods of asking sensitive questions are discussed in the next chapter.

## Chapter 2

# Asking sensitive questions: Review of studies

### 2.1 Introduction

This chapter examines the problem of asking sensitive questions, previous methods of handling them, and the advantages of using Interactive Voice Response (IVR) to ask them.

### 2.2 Aims

The aims of this chapter are to discuss:

1. The need to ask questions in public health research about sensitive behaviours;
2. The issues involved in asking sensitive questions;
3. Methods that have been used to ask sensitive questions; and
4. An alternative method, IVR.

## 2.3 Burden of illness

Some of the behaviours of most interest to public health research, due to their obvious and high burden of illness (for example, adolescent smoking, alcohol abuse, illicit drug use, domestic violence, sexual behaviours), are behaviours that are often very difficult to gauge. As described in this chapter, these are sensitive and personal behaviours where self-report is subject to at least a strong socially desirable reporting bias<sup>63,64</sup> (defined in Section 2.4), and sometimes legal consequences.

However, sensitive behaviours are behaviours which are often difficult to assess other than by self-report. As discussed later in this chapter, a number of alternative methods for assessing behaviours have been used, but these remain problematic. The effects of these behaviours can sometimes be measured, for example alcohol-related deaths, domestic violence deaths or rates of Acquired Immune Deficiency Syndrome (AIDS). But these are endpoint outcomes. For successful public health intervention, risk behaviours need to be assessable at the earliest possible time, while they remain modifiable and before they result in serious harm.

For example, alcohol abuse and illicit drug use have serious consequences for the health of the individual and well-being of the community, and are associated with a strong political imperative for intervention. Alcohol is involved in half of all violent crime and its use is relatively high in Australia<sup>63</sup>. Hazardous and harmful alcohol use in Australia in 1992 caused the loss of 3600 lives and 55,450 person-years at an average of 15.2 years of life per death. Also attributable to alcohol were 71,593 hospital episodes and 731,169 hospital bed days<sup>64</sup>. Illicit drug use resulted in 488 deaths and 17,899 person-years lost at an average of

36.7 years of life lost per death, 5390 hospital episodes, and 40,522 hospital bed days<sup>64</sup>.

Researchers and health practitioners need access to reliable data on the true extent of these sensitive behaviours so they can lobby for funding to develop effective interventions to modify these behaviours, to ensure that intervention attempts target the groups where these behaviours are most problematic, and to monitor changes in the prevalence of these behaviours, especially after intervention efforts.

## 2.4 Sensitive questions

A question is considered sensitive if it raises concerns about disapproval or other consequences such as legal sanctions for reporting truthfully. Those people with the most sensitive information are often the least likely to disclose<sup>65</sup>. Asking people about sensitive behaviours may increase the rate of non-response. Alternatively, respondents may give socially acceptable answers resulting in biased responses. A definition of socially desirable reporting is: *a respondent's tendency to over-report the frequency of performing a socially desirable behaviour, or under-report the frequency of performing a socially undesirable one*<sup>65,66</sup>. Questions about, for example, alcohol and illicit drug use, are likely to be subject to socially desirable reporting<sup>67</sup>.

For the most sensitive questions the respondent may be concerned that they will be labelled as an alcoholic or drug addict, that their employment may be adversely affected, their insurance capacity may be restricted, or they may be exposed to legal consequences. Although these concerns may be unsupported,



the respondent's responses are still likely to be modified.

Apart from socially desirable reporting, response bias to sensitive questions can also result from respondents attempting to maintain a good impression with the interviewer<sup>68</sup>, from questions that increase anxiety<sup>66</sup>, or respondent acquiescence<sup>69</sup>. Maintaining a good impression with the interviewer is known as impression management, and acquiescence is the tendency to agree with or defer to the interviewer. Since acquiescence can be associated with the respondent's culture, acquiescence may result in misleading ethnicity or cross cultural associations with the study outcome<sup>69,70</sup>. Impression management and acquiescence are not directly related to sensitive questions and so are not considered further here. However, since both are related to the respondent's interaction with the interviewer they are likely to be modified by any method that also reduces socially desirable reporting.

Not all areas are likely to be equally sensitive so that the degree of socially desirable reporting may also vary. Fleming et al.<sup>71</sup> found that respondents were least likely to be concerned about reporting smoking or fitness levels, slightly more concerned about weight, alcohol or drugs use, and very concerned about reporting sexual behaviour.

Socially desirable reporting also has socioeconomic correlates<sup>68,70,72,73</sup>. Clancy and Gove<sup>73</sup> found that socially desirable reporting was more likely in females than males, but Welte and Russell<sup>68</sup> found that socially desirable reporting increased with age, decreased with education, but did not differ between males and females. Carr and Krause<sup>72</sup> also found that socially desirable reporting decreased with increasing social status. Ross and Mirowsky<sup>70</sup> reported that socially desirable reporting increased with decreasing socioeconomic status and

increasing age.

## 2.5 Review of methods of asking sensitive questions

### 2.5.1 Methods

To identify methods of asking sensitive questions a literature search was conducted. The databases Psycinfo<sup>14</sup>, Medline<sup>15</sup>, Sociofile<sup>16</sup>, and Health Reference Center<sup>17</sup> were searched for articles in peer-reviewed journals to 1999 using the search terms: ‘sensitive questions’, ‘social approval’, ‘self-report’, ‘social desir\$’, and ‘interview\$’, where ‘\$’ stands for any combination of characters.

The articles were obtained and their bibliographies searched for additional articles that may have discussed sensitive questions. These articles were then obtained and treated in the same way. The process terminated when no more additional articles relating to methods of asking sensitive questions were found. The articles were examined for methods of handling sensitive questions, with a particular focus on the difficulties and merits incurred in their use within a survey context.

### 2.5.2 Results

Table 2.1 shows the methods identified and the corresponding articles located in relation to each method identified. The methods are detailed below:

Table 2.1: Methods for asking sensitive questions and studies identified in relation to each method.

<i>Method</i>	<i>Articles</i>
<b>DIRECT METHODS</b>	
<b>Face-to-face interviewing</b>	Aquilino <sup>74</sup> , Aquilino and Lo Sciuto <sup>75</sup> , Axinn <sup>76</sup> , Bradburn et al. <sup>66</sup> , Cannell and Fowler <sup>77</sup> , Carr and Krause <sup>72</sup> , Colombotos <sup>78</sup> , David <sup>79</sup> , Dohrenwend et al. <sup>80</sup> , Dohrenwend et al. <sup>81</sup> , Fendrich and Vaughn <sup>82</sup> , Fidler and Kleinknecht <sup>83</sup> , Gove and Geerken <sup>84</sup> , Greist et al. <sup>85</sup> , Henson et al. <sup>86</sup> , Hochstim <sup>87</sup> , Johnson and Delamater <sup>88</sup> , Kissinger et al. <sup>89</sup> , Mensch and Kandel <sup>90</sup> , Moum <sup>91</sup> , Presser and Stinson <sup>92</sup> , Richman et al. <sup>93</sup> , Schwarz et al. <sup>94</sup> , Seng <sup>95</sup> , Singer and Kohnke-Aguirre <sup>96</sup> , Skinner and Allen <sup>97</sup> , Stember and Hyman <sup>98</sup> , Waterton and Duffy <sup>99</sup> , Webster <sup>100</sup> , Welte and Russell <sup>68</sup>
<b>Self-Administered Questionnaires</b>	Akers et al. <sup>101</sup> , Ash and Abramson <sup>102</sup> , Boekeloo et al. <sup>4</sup> , Cannell and Fowler <sup>77</sup> , Corey <sup>103</sup> , DeLamater and MacCorquodale <sup>104</sup> , Evans et al. <sup>105</sup> , Fendrich and Vaughn <sup>82</sup> , Gfroerer and Hughes <sup>106</sup> , Johnson and Delamater <sup>88</sup> , Leutgart and Armstrong <sup>107</sup> , Liu et al. <sup>108</sup> , Makkai and McAllister <sup>109</sup> , Malvin and Moskowitz <sup>110</sup> , Moum <sup>91</sup> , Olson <sup>111</sup> , Petzel et al. <sup>112</sup> , Richman et al. <sup>93</sup> , Schwarz et al. <sup>94</sup> , Singer <sup>113</sup> , Skinner and Allen <sup>97</sup> , Smith <sup>114</sup> , Turner et al. <sup>115</sup>
<b>Mail surveys</b>	Ayidiya and McClendon <sup>116</sup> , Fuller <sup>117</sup> , Hebert et al. <sup>118</sup> , Hochstim <sup>87</sup> , Salvendy <sup>119</sup> , Siemiatycki <sup>120</sup> , Wildman <sup>121</sup>
<b>VALIDATION METHODS</b>	
<b>Biochemical validation</b>	Akers et al. <sup>101</sup> , Gillies et al. <sup>122</sup> , Keso and Salaspuro <sup>123</sup> , Latkin and Vlahov <sup>124</sup> , Magura et al. <sup>125</sup> , Peachey and Kapur <sup>126</sup> , Yoshino and Kato <sup>127</sup>
<b>Official records</b>	David <sup>79</sup> , Fendrich et al. <sup>128</sup> , Hardt and Peterson-Hardt <sup>129</sup>
<b>Collateral reports</b>	Maisto et al. <sup>130</sup>
<b>Diaries</b>	Matt et al. <sup>131</sup>
<b>Surreptitious identification</b>	Corey <sup>103</sup> , Erdos and Regier <sup>132</sup>

Continued on the next page.

Table 2.1 continued from the previous page.

*Method Articles*

---

### QUESTION PRESENTATION METHODS

#### **Bogus pipeline**

Bauman and Dent<sup>133</sup>, Campanelli et al.<sup>134</sup>, Evans et al.<sup>105</sup>, Jones and Sigall<sup>135</sup>, Lowe et al.<sup>136</sup>, Murray et al.<sup>137</sup>, Tourangeau et al.<sup>138</sup>

#### **Counterbiasing information**

Raghubir and Menon<sup>65</sup>

#### **Indirect questioning**

Bégin and Boivin<sup>139</sup>, Fisher<sup>140</sup>, Jo et al.<sup>141</sup>, Salvendy<sup>119</sup>

#### **Time use**

Presser and Stinson<sup>92</sup>

#### **Confidentiality assurance**

Frey<sup>142</sup>, Hardt and Peterson-Hardt<sup>129</sup>, Reamer<sup>143</sup>, Singer<sup>113</sup>, Woods and McNamara<sup>144</sup>

---

### ANONYMOUS METHODS

#### **Anonymity**

Akers et al.<sup>101</sup>, Ash and Abramson<sup>102</sup>, Becker and Bakal<sup>145</sup>, Fuller<sup>117</sup>, Leutgart and Armstrong<sup>107</sup>, Liu et al.<sup>108</sup>, Malvin and Moskowitz<sup>110</sup>, Olson<sup>111</sup>, Rosen<sup>146</sup>, Wildman<sup>121</sup>

#### **Context-Determined Rule-Generated Pseudonym (CDRGP)**

Carifo and Biron<sup>147</sup>

#### **Randomized response**

Abul-Ela et al.<sup>148</sup>, Akers et al.<sup>101</sup>, Bégin and Boivin<sup>139</sup>, Boruch<sup>149</sup>, Campbell and Joiner<sup>150</sup>, Duffy and Waterton<sup>151</sup>, Fidler and Kleinknecht<sup>83</sup>, Folsom et al.<sup>152</sup>, Goodstadt et al.<sup>153</sup>, Greenberg et al.<sup>154</sup>, Levy<sup>155</sup>, Reinmuth and Geurts<sup>156</sup>, Warner<sup>157</sup>

#### **Item counting**

Maxim<sup>158</sup>

---

### COMPUTER METHODS

#### **Computer interviewing**

Boekeloo et al.<sup>4</sup>, Erdman et al.<sup>159</sup>, Greist et al.<sup>85</sup>, Kissinger et al.<sup>89</sup>, Kohlmeier et al.<sup>160</sup>, Koson et al.<sup>161</sup>, Matt et al.<sup>131</sup>, Skinner and Allen<sup>97</sup>, Supple et al.<sup>162</sup>, Tourangeau and Smith<sup>3</sup>, Waterton and Duffy<sup>99</sup>

#### **Telephone interviewing**

Aquilino<sup>74</sup>, Aquilino and Lo Sciuto<sup>75</sup>, Clancy and Gove<sup>163</sup>, Colombotos<sup>78</sup>, DeLamater and MacCorquodale<sup>104</sup>, Fendrich et al.<sup>128</sup>, Fendrich and Vaughn<sup>82</sup>, Frey<sup>142</sup>, Gfroerer and Hughes<sup>106</sup>, Hebert et al.<sup>118</sup>, Henson et al.<sup>86</sup>, Hochstim<sup>87</sup>, Ivis et al.<sup>164</sup>, Kegeles et al.<sup>165</sup>, Moum<sup>91</sup>, Schwarz et al.<sup>94</sup>, Siemiatycki<sup>120</sup>, Smith<sup>114</sup>

---

**Direct methods**

**Face-to-face interviewing** The most direct method is with face-to-face interviewing. However, response bias is often reported with face-to-face interviewing relating to sensitive issues. The response bias in face-to-face interviewing can result from socially desirable reporting when respondents are asked sensitive questions<sup>79,90,93</sup>. David<sup>79</sup> found that families on welfare support only reported 38% of their true income when interviewed face-to-face.

Mensch and Kandel<sup>90</sup> used longitudinal data obtained from three large national US surveys of drug use. They found that lifetime drug use (except marijuana) was increasingly under-reported with successive interviews. Lifetime reports should not have logically declined with successive interviews and the use of substances such as cocaine was not likely to be forgotten. However, using face-to-face interviewing, the later interviews resulted in under-report compared to earlier reports. Mensch and Kandel<sup>90</sup> noted that the under-reporting occurred when the same interviewers were used with successive interviews, but did not occur when different interviewers were used. This indicated that it may have been the familiarity with the interviewer that reduced reporting.

Other sources of bias have been noted with face-to-face interviewing. Response bias in face-to-face interviewing occurs when the respondent's characteristics interact in some way with the interviewer's demographic characteristics (age<sup>91</sup>, sex<sup>76,100</sup>, ethnicity<sup>100</sup>), or familiarity with the interviewer<sup>90</sup>. Response bias also increases as the social distance (income and ethnicity) between the interviewer and respondent increases<sup>80,81</sup>. Lastly, Singer and Kohnke-Aguirre<sup>96</sup> found that interviewers who expected respondents to under-report, or that the interview would be difficult, tended to obtain higher item nonresponse rates.

Stember and Hyman<sup>98</sup> also found that the magnitude of interviewer effects depended on question-form and were more important with respondents interested in the survey topic.

The combination of socially desirable reporting and interviewer-mediated response bias suggested that alternative methods should be used to examine sensitive issues.

### **Self-Administered Questionnaires** Self-Administered Questionnaires (SAQs)

are completed by the respondent without direct observation by an interviewer. Self-report for sensitive issues can be higher with SAQs than with face-to-face interviewing<sup>77</sup>, perhaps because of the method's greater perceived anonymity.

Self-report can be enhanced further with non-identified rather than identified questionnaires<sup>102,108,109,111</sup>. When used within longitudinal studies or to further emphasize the anonymity of the questionnaire, the more sensitive questions can be placed within sealed booklets<sup>109</sup>.

**Mail surveys** Mail surveys lack an interviewer and this might reduce socially desirable reporting<sup>116</sup>. Respondents are also able to answer questions in any order, change answers and take their time to complete the questionnaire.

However, Ayidiya and McClendon<sup>116</sup> found no order and primacy effects, and considerable numbers of no-opinion answers in a mail survey. This indicated that responses to mail surveys did not necessarily benefit from respondents being able to peruse the questionnaire, and the rate of non-committal responses might increase.

Mail surveys can also produce high item nonresponse rates<sup>87,120,166</sup> and low response rates<sup>120,167</sup>. Respondents with literacy or language deficiencies might

have difficulty completing the questionnaire and be less likely to return it<sup>168</sup>.

### Validation methods

**Biochemical validation** Self-report of substance abuse can be confirmed by biochemical tests that detect the presence of metabolites. However, the agreement of a self-reported result with a biochemical test result depends on: (a) the characteristics of the substance<sup>169</sup>, (b) sensitivity and specificity of the test<sup>169</sup>, (c) the time lag between testing and consumption<sup>169</sup>, (d) the potential presence of small amounts of metabolites in individuals who have not consumed the substance<sup>101</sup>, and (e) knowledge of normal ranges<sup>101</sup>. The cost of biochemical tests<sup>170</sup> and reduced response rates might also restrict their use with surveys. Lastly, they are only useful for asking sensitive questions on drugs and similar issues for which biochemical tests are available.

**Official records** Another means of validating a response is to compare the respondent's answers to official records, such as the respondent's medical records<sup>128</sup>, unemployment benefit records<sup>79</sup>, and police arrest history<sup>129</sup>. However, obtaining these data can be costly, they might not be available for the behaviour or individuals under study<sup>128</sup>, and they might require the respondent's consent. Obtaining consent might be difficult for sensitive issues in a general survey context. However, at a group level, data for some issues, such as cervical screening rates, might sometimes be available in summarized form.

**Collateral reports** If the sensitive behaviour can be observed by someone familiar with the respondent then their report can be used to validate the self-report. For example, a respondent's alcohol self-report might be verified by

comparing it with the respondent's husband's or wife's report. Collateral report is usually correlated with and exceeds self-report<sup>130</sup>.

However, subjects might deceive their partners<sup>127</sup>. While this method might be of use in monitoring alcoholics, it is not of use where the behaviour is unlikely to be observed by the partner. It is also less applicable to surveys than a clinical or treatment study.

**Surreptitious identification** Respondents are more likely to report accurately under genuinely confidential reporting conditions<sup>3</sup>. The identification of respondents is necessary in longitudinal surveys, and numbered questionnaires might be needed for administrative purposes<sup>132</sup>. However, such codes and identifying information on questionnaires might suggest a lack of confidentiality to respondents.

To avoid this problem surreptitious identification marks have been used. One method was to print questionnaires on different weights or colours of paper, which allowed for group-matching but not individual identification<sup>121</sup>. In another study, invisible fluorescent ink was used to mark questionnaires<sup>132</sup>, and in another discreet marks or holes were made in questionnaires so that they could be matched to separately obtained identification slips<sup>146</sup>. Such methods are clearly unethical and might also lead to distrust of researchers.

**Diaries** To overcome memory problems in reporting past events, Matt et al.<sup>131</sup> asked respondents to keep daily diaries. The recording of information in this manner might be less subject to socially desirable reporting since they are recorded by respondents outside the interview context. When respondents were subsequently interviewed, their self-reports were found to exceed the diary



reports<sup>131</sup>. However, this method might incur considerable respondent burden and expense.

### **Question presentation methods**

**Bogus pipeline** The bogus pipeline method involves convincing respondents that their self-report could be verified by a measuring device that might not actually exist<sup>135</sup>. It is assumed that respondents report more accurately if they think that their response could be independently validated<sup>134</sup>.

Some studies found that the bogus pipeline increased self-report of smoking, drugs, sexual behaviour, and alcohol<sup>105,133,136–138</sup>. However, no bogus pipeline effect was found by Akers et al.<sup>101</sup> or Campanelli et al.<sup>134</sup> in surveys of alcohol and smoking in children. A bogus pipeline also resulted in greater negative affect than direct questions<sup>135</sup>.

Whether the bogus pipeline effect actually results in accurate self-report remains unclear<sup>134</sup>. It might require special conditions to work<sup>137</sup>, and may cause a systematic over-reporting bias<sup>101</sup>.

**Counterbiasing information** Since bias can result from socially desirable reporting, one approach, called counterbiasing, has been to attempt to convince the respondent that a particular socially undesirable behaviour is more common and acceptable than is actually the case<sup>65</sup>. However, counterbiasing is only likely to work with behaviours where respondents are not well-informed about their actual prevalence. Where it does work, the effect might be sensitive to the degree of counterbiasing. This means that different studies could obtain incongruent results if their application of counterbiasing is not the same. There is also the concern that suggesting undesirable behaviours were common might

imply their acceptability and lead to respondents modifying their own behaviour as a result<sup>68</sup>.

**Indirect questioning** An early method of reducing socially desirable reporting was to ask the respondent questions of their opinion of other people's behaviour rather than their own<sup>140</sup>. For example, if the direct question is

“Do you shop lift?”

the indirect question would become

“Do you think other people shop lift?”

It is assumed that the respondent would find it easier to discuss sensitive issues if asked about other people's actions, but that their answers would still reflect their own behaviour<sup>139</sup>. Indirect questioning has been found to reduce socially desirable reporting<sup>140</sup>. The method was subsequently extended by Jo et al.<sup>141</sup> who found that combining indirect and direct questioning using structural equation modeling also reduced socially desirable reporting.

However, interpretation of indirect questions can be difficult since it can not be determined whether the respondent's answers refer to their personal attributes or their opinions of others. Bégin and Boivin<sup>139</sup> found that respondents asked indirect questions reported what they thought others would do, rather than themselves.

Using indirect questions might also affect response rates. In a mail survey on religion and transportation, Salvendy<sup>119</sup> obtained higher response rates using a direct rather than an indirect questionnaire.

**Time use** Another means of indirectly asking a question is to request the respondent to enumerate all activities that occur within a given time frame. For example, Presser and Stinson<sup>92</sup> found that

“Please list what you did on Sunday”

gave a more accurate estimate of religiosity than simply asking

“Do you attend a religious institution?”

The time-use method might reduce socially desirable reporting by having the respondent concentrate on recalling historical information rather than being concerned with the impression that they may make on the interviewer.

However, the time-use method presumably suffers from a reliance on the respondent’s memory. If a careless interviewer prompted the respondent with suggestions for possible activities the respondent might still respond in what they perceived was a socially desirable manner.

**Confidentiality assurance** A simple assurance of confidentiality is often used in interviews in an attempt to reduce socially desirable reporting. Singer<sup>113</sup>, Frey<sup>142</sup> and Reamer<sup>143</sup> all found that a confidentiality assurance had no effect on response quality. A confidentiality assurance can also increase item nonresponse rates<sup>113</sup> and subject apprehension if the respondents are suspicious of the researcher’s intent<sup>143</sup>. Woods and McNamara<sup>144</sup> found that stating that answers would not be kept confidential had the effect of reducing data quality. However, this does not imply that a confidentiality assurance can increase data quality. It appears that a simple confidentiality assurance is an insufficient technique for ensuring accuracy for sensitive questions.

**Anonymous methods**

**Anonymity** Ensuring that the interview is genuinely and clearly anonymous reduces socially desirable reporting<sup>3,145</sup>. Anonymity can be easily and overtly achievable if questionnaires lack any identifying codes, although respondents might still be suspicious if the questionnaire is collected by an interviewer. Similarly, electronic methods might be genuinely anonymous if no identifying information is requested.

Anonymous questionnaires must lack administrative code numbers<sup>121</sup>, and signatures or contact details should not be requested<sup>111</sup>. It was noted that anonymity had a greater effect in household surveys than business surveys<sup>121</sup>. This was possibly because a survey received at home might be perceived as more intrusive than one received at work<sup>121</sup>.

While anonymous methods are clearly suitable for cross-sectional studies, they can be more difficult to use in longitudinal studies. This is because the information used in longitudinal studies to match responses from various times can not be collected without also retaining identifying information<sup>147</sup>.

**Context-Determined Rule-Generated Pseudonym (CDRGP)** The Context-Determined Rule-Generated Pseudonym (CDRGP) method allows anonymous interviewing to be used in longitudinal studies<sup>147</sup>. The respondent can be asked during the first contact to create a pseudonym that would be used for all subsequent contacts. If the respondent is asked to create a pseudonym themselves they might subsequently forget it and create another when interviewed on a second occasion. This would lead to attrition of respondents throughout a longitudinal study.

Using the CDRGP method, the respondent does not need to remember the pseudonym since it is regenerated each time according to a fixed set of rules. The pseudonym typically consists of a series of letters constructed from the first letter of the respondent's name, first letter of their birth month, an "M" or "F" depending on their sex, and so on. Although ingenious, the use of respondent-generated identification codes has still led to considerable attrition over time<sup>110</sup>.

**Randomized response** An anonymous interview can reduce socially desirable reporting but a respondent might still be suspicious that their answers might be identifiable. Randomized response avoids this by ensuring that self-report is demonstrably anonymous.

In randomized response the respondent randomly selects one of a pair of matched questions to answer, submits the answer to the interviewer, but does not reveal which question has been answered. One of the pair of questions is sensitive and the other innocuous. An example of a question-pair is<sup>152</sup>:

"I was born in the month of April."

"I had an automobile accident during last year in which I was at fault."

The selection of which question to answer is made using a randomization device such as a dice. Which question has been answered is known only to the respondent. The interviewer simply records the answer, and since they are unaware of which question was answered the respondent's status remains confidential.

While individual respondents' answers remain confidential Warner<sup>157</sup> showed that the proportion of responses by all respondents to the sensitive question

could be subsequently recovered using Equation 2.1. In the equation, the proportion of persons with the sensitive attribute ( $\hat{\pi}$ ) can be obtained given the observed proportion of ‘yes’ answers ( $\hat{\lambda}$ ) and the probability of the sensitive question being answered ( $p$ ), which is fixed by the appropriate use of a randomization device, such as a dice. For example, one way to obtain  $\hat{\lambda} = 2/6$  using a dice would be to ask the respondent to answer the sensitive question when the roll of a dice resulted in ‘1’ or ‘2’, and to answer the non-sensitive question when the roll resulted in ‘3’, ‘4’, ‘5’, or ‘6’. The method assumes that respondents always answer honestly since they can be assured their personal status remains unknown.

$$\hat{\pi} = \frac{\hat{\lambda} - (1 - p)}{2p - 1} \quad (2.1)$$

where:  $\hat{\lambda}$  = the observed proportion of ‘yes’ answers,

$p$  = the probability of the sensitive question being answered ( $p \neq 1/2$ ), and

$\hat{\pi}$  = the estimate of the proportion of persons with the sensitive attribute.

The method was subsequently further developed<sup>148,149,151,152,156</sup>, including variations for obtaining quantitative data such as aggregated response<sup>171</sup>. Randomized response has been extensively used to obtain self-report with sensitive issues<sup>83,101,139,153,154,156</sup>.

However, randomized response has considerable disadvantages:

- The respondent must manipulate a randomization device such as a dice.

This makes the method unwieldy in modes other than face-to-face interviewing, especially when there are several sensitive questions. However, it has been used in Computer-Assisted Telephone Interview (CATI)

surveys<sup>172</sup>;

- The method might be unacceptable or confusing to respondents<sup>171</sup>;
- In many studies, once individuals admit to a sensitive behaviour they are asked additional questions to further probe aspects of the behaviour. This can not be done using randomized response since the respondent's status with respect to the sensitive behaviour is unknown. If additional information is required then the probe questions would need to be asked of all respondents, which may be burdensome and confusing. If the probe questions are also sensitive then they would also have to use randomized response;
- The method is of little use in longitudinal studies since the respondent's status is unknown<sup>147</sup>;
- The method is difficult to administer<sup>147</sup>; and
- A large sample size is required because the estimates of prevalences have a larger variance than when using direct questions<sup>150</sup>.

**Item counting** The item counting method is another anonymous method<sup>158</sup>. The respondents are divided into two groups. Individuals in both groups are presented with a series of items and asked how many of the items are applicable to them. For one group the items in the series refer to innocuous and commonplace behaviours. In the other group the series contains an additional item that refers to a single sensitive behaviour. Since respondents are only asked to report how many items apply to them the method is anonymous. The difference in the mean number of items between the two groups gives the prevalence of the

sensitive behaviour.

Unlike randomized response, a randomization device is not required so the method can be used in a variety of modes, including telephone interviewing.

However, the method has some disadvantages:

- As with randomized response it requires large sample sizes<sup>171</sup>;
- Probe questions can not be directed to particular individuals;
- It is difficult to apply within longitudinal studies since the respondent's status is unknown;
- The questions are more complex than direct questions;
- Each sensitive behaviour must be embedded within a series of insensitive ones. This lengthens the interview compared to one where direct questions are asked; and
- Some respondents might become suspicious when presented with the series containing the sensitive item<sup>171</sup>.

### Computer methods

**Computer interviewing** In the Computer-Assisted Personal-Interview (CAPI) and Computer-Assisted Self-Interview (CASI) methods, interviews are conducted by presenting the questions on a computer screen and entering answers directly into the computer via a keyboard or touchscreen.

By automating interviews, CASI and CAPI can standardize questions<sup>89</sup>. While CASI can eliminate interviewer variation, CAPI might allow the interviewer to maintain a rapport with the respondent<sup>159</sup>. However, the perceived



anonymity of CASI might allow respondents to report sensitive information more readily<sup>89,93,99,162</sup>. CASI might also be accompanied by audio recordings of questions<sup>3,4,108,115</sup> and video recordings of a person asking the question<sup>89</sup>.

Computer interviewing has obtained higher self-report for alcohol consumption<sup>99</sup>, more accurate sexual behaviour<sup>3,89</sup>, more accurate drug use<sup>162</sup>, and reduced embarrassment<sup>3,85,159</sup>. Skinner and Allen<sup>97</sup> and Koson et al.<sup>161</sup> reported that the methods were perceived as less threatening than a personal interview. Greist et al.<sup>85</sup> found good agreement with traditional methods of assessing psychological conditions. Boekeloo et al.<sup>4</sup> obtained less item nonresponse using computer interviewing than with paper questionnaires.

Computer interviewing might eliminate data entry and reduce staffing, but the hardware and programming costs might be considerable, and the methods' portability might be limited.

**Telephone interviewing** The use of a telephone permits interviewing to be conducted anywhere within the range of the public telecommunications system. While not necessarily a computer interviewing method, most modern telephone surveys are done with a Computer-Assisted Telephone Interview (CATI) approach<sup>1</sup>.

Telephone interviewing is perceived as more anonymous than face-to-face interviewing<sup>86,94,120</sup>, but less confidential than SAQs<sup>94,120</sup>. Aquilino and Lo Sciuto<sup>75</sup> and Gfroerer and Hughes<sup>106</sup> found that self-report of drug use was under-reported by telephone interview compared to SAQs. Moum<sup>91</sup> found no difference in identification of psychiatric disorders between telephone and face-to-face modes, but the use of SAQs obtained higher rates than either telephone or face-

to-face. Aquilino<sup>74</sup> found that the differences in reporting by telephone rather than face-to-face may be more important with some subpopulations, such as black Americans.

Telephone interviewer characteristics such as voice<sup>54</sup>, sex and age<sup>94</sup> can also influence socially desirable reporting. Fendrich and Vaughn<sup>82</sup> found that under-reporting was associated with increasing interviewer familiarity with the respondent.

### 2.5.3 Conclusions

Of the direct methods, the more confidential methods, such as SAQs and mail surveys, obtained lower socially desirable reporting, but could produce high item non-response rates. The validation methods were impractical and expensive for survey applications, and the question presentation methods were problematic. The anonymous methods reduced socially desirable reporting, but were impractical and expensive, and may not be applicable to longitudinal studies that track individuals. The computer-based methods reduced socially desirable reporting and telephone interviewing could result in interviewer bias. It also remained unclear whether CATI reduced socially desirable reporting.

## 2.6 An alternative method:

### Interactive voice response (IVR)

One way to ameliorate many of the problems associated with asking sensitive questions is to improve the technology used for data collection. Interactive Voice Response (IVR) is a relatively new and promising technique for telephone

interviews that enables the rapid, accurate and timely collection of data with a confidentiality similar to that of SAQs.

In IVR the human telephone interviewer is replaced by a high quality recorded voice interview to which the respondent provides answers by pressing the keys of a touch-telephone. The system can be programmed to call nominated numbers or to accept calls. Using IVR, interviews can be conducted at the subject's convenience allowing access to populations not available during business hours, permitting a greater logistic freedom. As with CATI, subjects can be anywhere within the range of modern telecommunications systems. There is no interviewer bias and transcription errors are reduced since the data are directly keyed by the respondent.

Many people have already encountered IVR when dealing with large corporations or government agencies. In 1998 the proportion of households in Australia paying bills or transferring funds by telephone ranged from 27% to 44% between states and territories<sup>173</sup> indicating that Australians were increasingly familiar with IVR technology. Some universities now use IVR to provide course and career information, provide examination results, library loan control, re-enrollment and distance-learning information<sup>174–176</sup>.

IVR is a relatively new area in public health research, but it appeared from the previous research discussed in Chapter 1 to have considerable promise as a standardized and confidential tool for assessing community rates of screening behaviours and for delivering health-related messages and reminders.

Overall, it was considered that IVR offered considerable advantages over the other methods:

- Although not a genuinely anonymous method such as the randomized

response or item counting methods, IVR was likely to be perceived as confidential, which should reduce socially desirable reporting<sup>67,85,177</sup>;

- Non-sampling errors would be reduced since the data are directly keyed by the respondent<sup>26,51,178</sup>;
- There is no interviewer effect either at a single time, nor accumulating over time, as might occur in face-to-face interviewing<sup>90</sup>;
- IVR could be used to track consenting individuals within longitudinal assessments<sup>10,33,35,179,180</sup>;
- Probe questions could be used when individuals reporting sensitive behaviours were identified<sup>33</sup>;
- An IVR system could run automatically and continuously enabling large populations to be economically screened or interviewed<sup>44,181</sup>;
- Respondents could listen to a message at their own rate rather than being hurried, as might occur in a CATI<sup>94</sup>;
- Since IVR used a standardized interface and assessment was conducted by a computer, diagnostic accuracy was likely to be improved<sup>178</sup>; and
- The cost per contact was likely to be small<sup>178,182–184</sup>.

## 2.7 Conclusions

The considerable advantages of IVR over existing methods for dealing with sensitive questions merited investigation. IVR technology had not yet been tested

for acceptability, cost-effectiveness and feasibility in the general community setting. It was anticipated that it would provide a research tool similar to CATI, but also with an improved accuracy, more up-to-date monitoring of public health behaviours, and at a lower cost.

It was decided that a comparison should be made of IVR and another method to ask sensitive questions. The closest interviewing method to IVR was CATI. The next chapter describes the development of a software tool to provide both IVR and CATI functionality.

## Chapter 3

# The Generalized Electronic Interviewing System (**GEIS**)

### 3.1 Introduction

At the initiation of this research no software was known that could provide both Computer-Assisted Telephone Interview (CATI) and Interactive Voice Response (IVR) functionality, nor that could provide the complexity that was expected in the surveys to be conducted. Accordingly, a custom software called Generalized Electronic Interviewing System (GEIS) was developed for this study by the candidate.

Because the exact nature of the functionality that would be required was not clear at the time, a very general scripted system was developed capable of supporting the two modes, CATI and IVR. Unlike a programmed interviewing system, a scripted system allowed scripts to be easily and quickly modified to adapt to changing research needs, and also allowed interview scripts to be completely and openly documented. For example, no aspect of an interview, such as question-jumps, was written in program code. Instead, all aspects of the interview were incorporated into the script.

Since the data would be analyzed in SAS/STAT®software<sup>185</sup>, *GEIS* was written with SAS/AF®software<sup>186</sup> to simplify data access. However, Visual Voice® Pro version 5.0 software<sup>187</sup> was needed to control the IVR hardware.

## 3.2 Aims

The aims of this chapter are to describe:

1. Using *GEIS* to create an interview;
2. The implementation of *GEIS*;
3. The storage of data within *GEIS*; and
4. How scripts are written using *GEIS*.

## 3.3 Using *GEIS* to create an interview

*GEIS* eliminated the need for programming. Instead, the survey designer specified all details of CATI and IVR questionnaires in the script. The overall procedure to develop a script in *GEIS* is shown in Figure 3.1. In this study, the candidate was the survey designer.

The process began with the creation of a data set containing the respondent contact information and an interview script. The *CONFID* dataset (see Table 3.4 of Section 3.5.1) contained the respondents, their telephone numbers, and addresses. Each respondent was assigned a unique identification code using the *ID* variable. The interview script was then prepared using a text editor. The script had to obey *GEIS* syntax conventions (Section 3.6). Apart from defining the

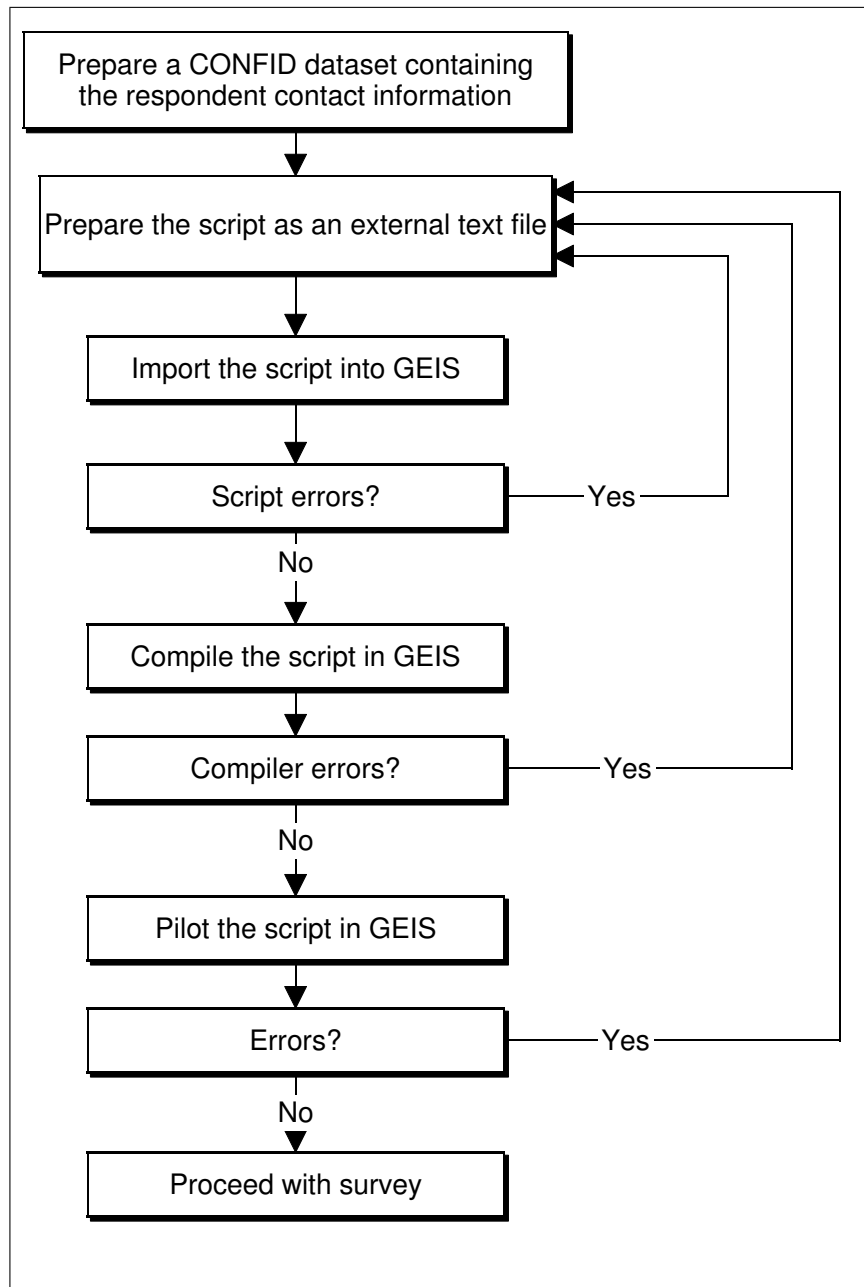


Figure 3.1: The procedure to develop a script within GEIS.



interview within **GEIS** the script also served as an open, standard, and publishable method of documenting interviews. Other data sets might sometimes be needed, such as the responses given by individuals during a previous interview that might be used to select questions for the next interview.

The script was imported into **GEIS** (Section 3.4.1). Errors in syntax were reported by **GEIS** at this stage. These were corrected and the script import attempt repeated until no errors occurred.

The script was then compiled. The compiler created the data sets needed by the project and performed numerous checks (Appendix Section A.1) that had to be passed before the script was permitted to be used. These checks were sufficiently detailed that most scripts ran reliably on the first attempt.

The script was then piloted (Section 3.4.1). This step had a series of goals, including finding errors not detected by the compiler, refining the wording of questions, piloting the interview with volunteers, and training interviewers.

Once the script testing had been completed, the project data sets were transferred to a share server (Section 3.5.3) where they might be used by CATI interviewers and the IVR system to collect data.

## 3.4 Implementation of **GEIS**

### 3.4.1 The **GEIS** windows

**GEIS** consisted of a series of windows described below.

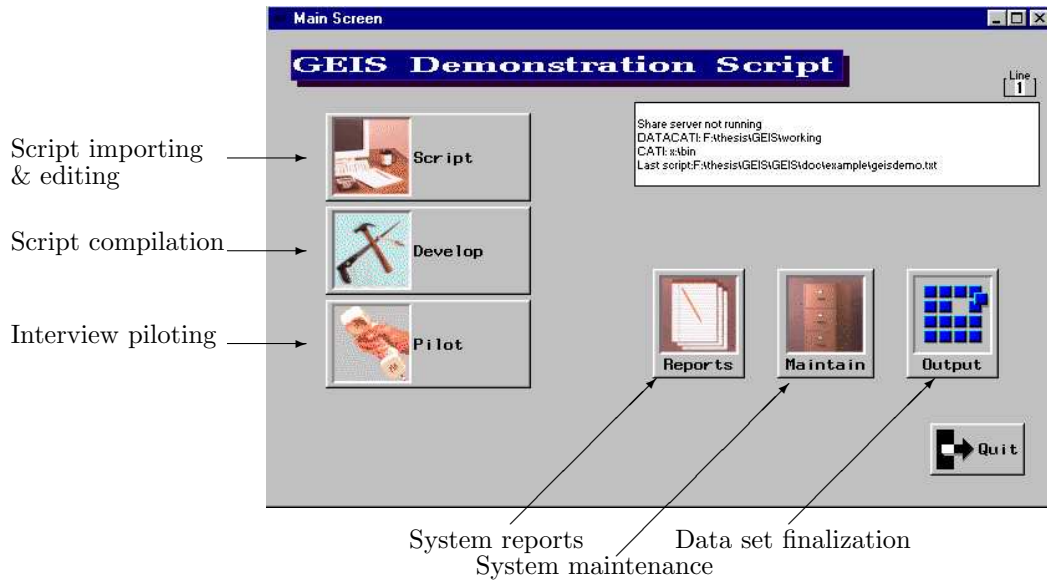


Figure 3.2: The GEIS control panel.

### Control panel

The GEIS control panel (Figure 3.2) allowed access to all functions. The primary facilities available from the control panel were: (a) script importing and editing, (b) script compilation, (c) interview piloting, (d) system reports, (e) system maintenance, and (f) data set finalization. Each of these facilities was obtained by clicking the corresponding buttons in the control panel. They are described in detail below.

### Script importing and editing

Scripts were imported into GEIS from the script import window (Figure 3.3). The script import window displayed the script file name as well as the names of any sub-scripts. A script could be edited in the internal editor or by using an external text editor or word processor.

When a script was imported a window (Figure 3.4) was displayed to show the results of the import process. Errors and warnings were displayed in differently

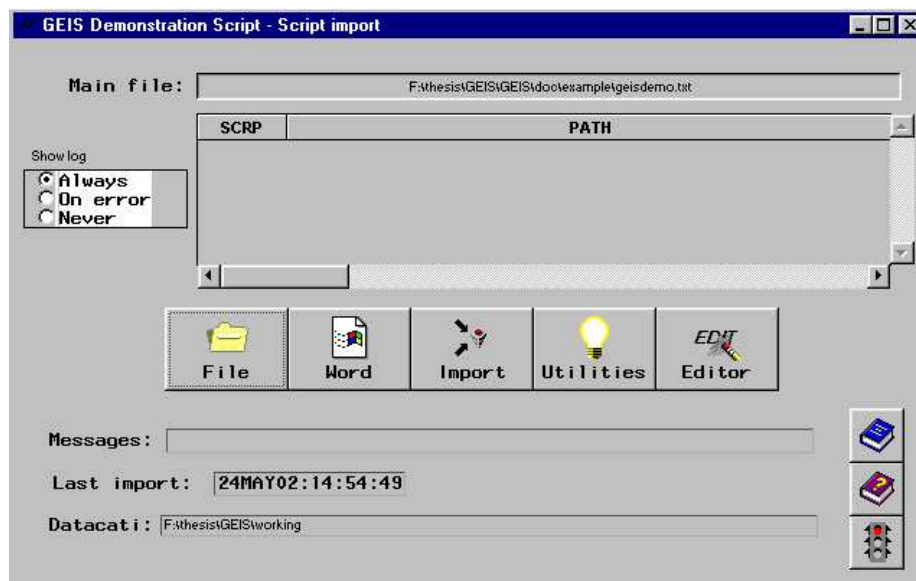


Figure 3.3: The GEIS script import window.

coloured texts.

Other utilities available from the script import window were script exporting, flow-charting, script viewing, and editing the recorded messages using the phrases editor. The phrases editor (Figure 3.5) allowed the user to record the IVR messages for questions in a script. The text for each question was displayed on the screen. The phrases editor window was used to assign the recorded message file to the question.

### Script compiling

The compiler window was opened from the control panel by selecting the ‘**Develop**’ button (Figure 3.2). An imported script was compiled from the compiler window (Figure 3.6). This process involved creating data sets and catalogues, checking the script for errors, and updating the GEIS log.

The GEIS compiler wrote error, warning and information messages to the

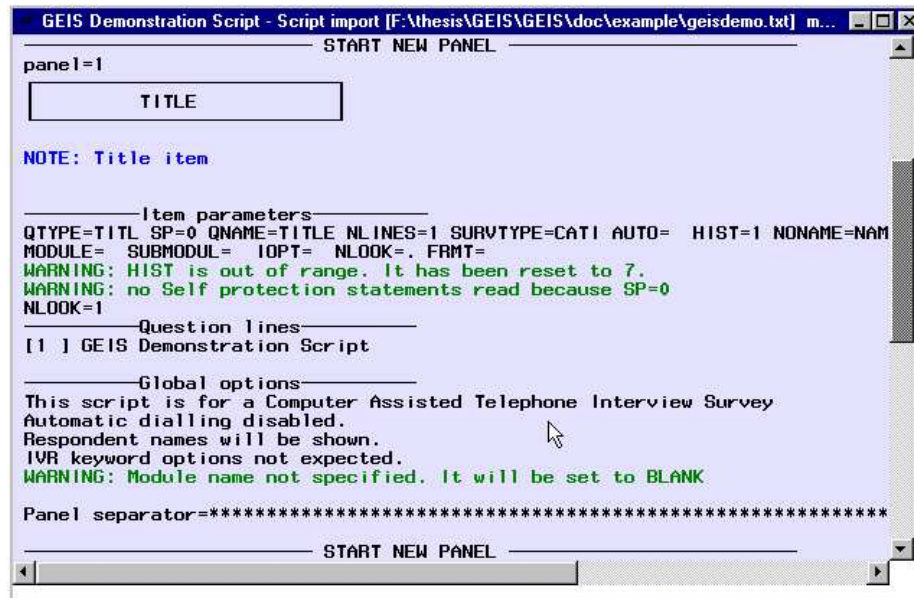


Figure 3.4: The GEIS script import results window.

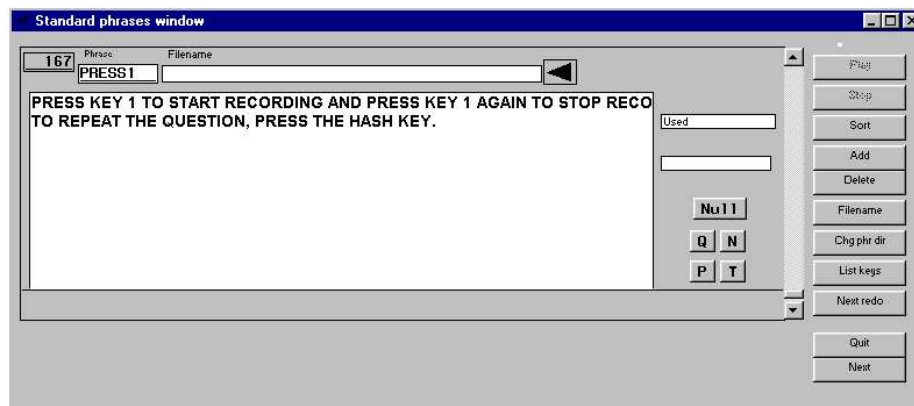


Figure 3.5: The GEIS phrase editor for IVR messages.

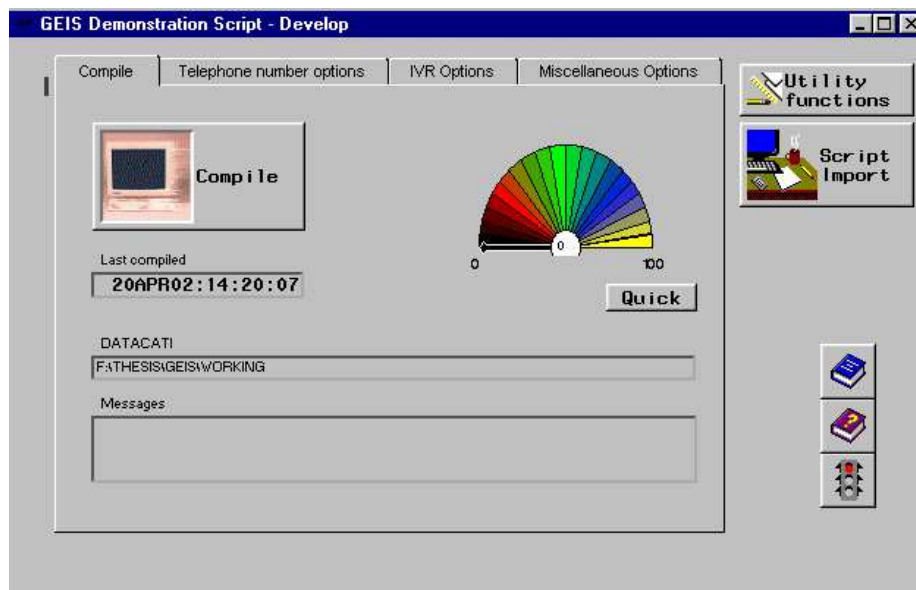


Figure 3.6: The GEIS compiler window.

GEIS log. After each compilation the GEIS log needed to be examined and if errors were found the script needed to be modified and recompiled. The compiler also deleted the project's data sets if errors occurred, preventing the user from proceeding with an invalid script.

Additional functions available from the compiler window included: (a) manual editing of the GEIS data sets; (b) viewing compiler information; (c) configuring the GEIS e-mail system; (d) processing the GEIS formats; (e) recovering interview information lost due to power failure; and (f) viewing automatically generated flow-charts showing the script's logic structure.

### Interview piloting

Interview scripts needed to be pilot-tested before the survey started. This is termed 'piloting' in this thesis. To do this, the full interviewing system was run from the control panel by selecting the '**Pilot**' button (Figure 3.2). The CATI

interviewers did not use the control panel; instead they ran the full interviewing system directly.

In CATI surveys the interviewing system consisted of three separate windows. Starting from the control panel, pressing the **‘PILOT’** button opened the interviewer log-in window (Figure 3.7). This window was the same window used by interviewers to log into the GEIS interviewing system. Once logged in, interviewers could then press the **‘START’** button which opened the respondent selection window (Figure 3.8). In the respondent selection window, the interviewer could select respondents that appeared in a pull-down list. These respondents were those that GEIS had determined were currently eligible for interviewing or which had been deliberately selected by the survey administrator. The pull-down list also allowed interviewers to search by name or telephone number for a particular respondent within the list.

Once a respondent had been selected an interview could commence. To start an interview, interviewers pressed either the **‘CALL ANSWERED’** or **‘CALL’** buttons\* which opened the respondent interviewing window (Figure 3.9).

For IVR surveys only the respondent interviewing window was used. Starting from the control panel, pressing the **‘PILOT’** button caused the respondent interviewing window to open. GEIS would then automatically proceed with selecting respondents, ringing numbers, and conducting interviews.

---

\*Whether the **‘CALL ANSWERED’** or **‘CALL’** buttons were displayed depended on the script.



Figure 3.7: The GEIS interviewer log-in window.

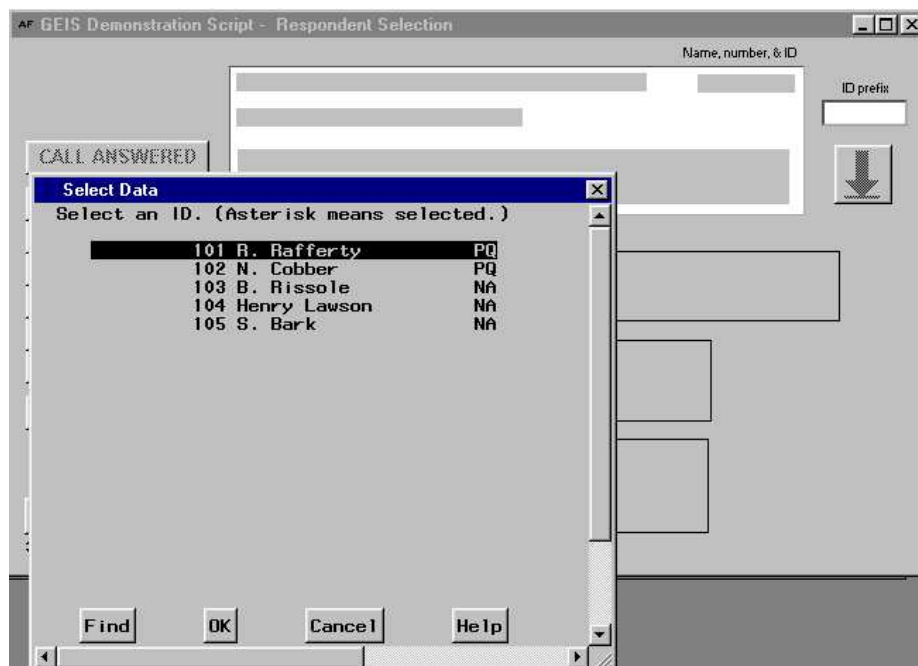


Figure 3.8: The GEIS respondent selection window showing pull-down list used to select a respondent.

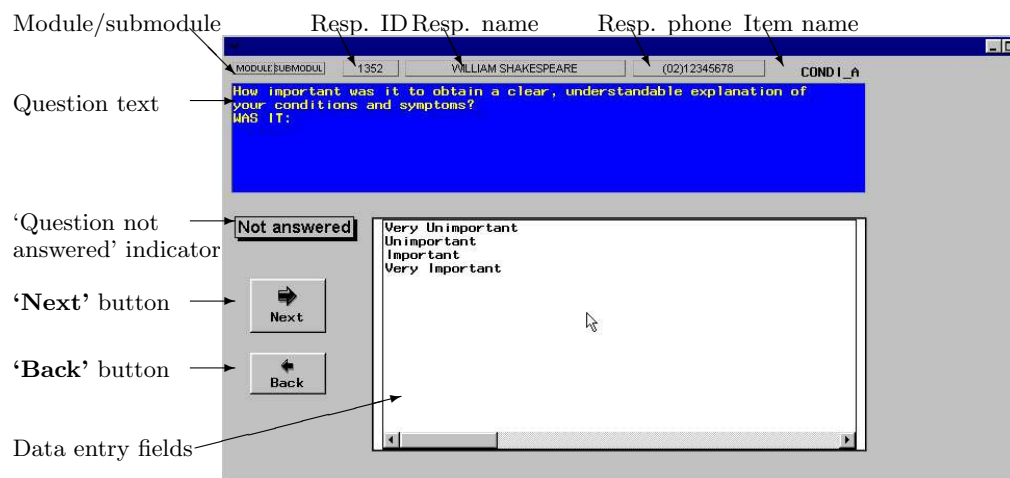


Figure 3.9: The GEIS interviewing window.

Figure 3.9 illustrates the interviewing window. To minimize interviewer-training and interview-variation the design of the interviewing window remained unchanged between projects. For example, the data entry controls remained in the same location on the screen for each question and each survey project.

For each question the interviewing window displayed the: (a) respondent ID code; (b) respondent name; (c) respondent telephone number; (d) item name; (e) module and submodule; (f) question text; (g) ‘question not answered’ indicator; (h) ‘**NEXT**’ and ‘**BACK**’ buttons; and (i) data entry fields.

The ‘**Next**’ button caused GEIS to display the next appropriate question and the ‘**Back**’ button caused GEIS to redisplay the last asked question. In the IVR mode the same functionality was present as in the CATI case, but movement back and forwards through the script was controlled by the respondent pressing keys on their touchphone. The \* key returned to the previous question, and the # key moved ahead to the next question or repeated the current one.

GEIS never displayed the next question until the current question had been answered. It did allow the previous question to be redisplayed so that the



	ID	STATUS	DURATION	LASTQ	NTIMES
1	101	PQ	2:17:32	QCHCE	1
2	102	PQ	0:00:08	QCHCE	1
3	103	NA	.	.	.
4	104	NA	.	.	.
5	105	NA	.	.	.

Figure 3.10: An example of a GEIS report window.

previously-entered data might be corrected. In this way the interviewer might return as far back as the beginning of the interview. This was also possible in the IVR mode. The means by which GEIS determined the next question to be displayed is discussed in Section 3.6.3.

### System reports

The reports window was opened from the control panel by selecting the ‘**Re-ports**’ button (Figure 3.2). GEIS project data sets were inspected from the reports window (see Figure 3.10) during and following a survey. The reports window was normally run to monitor the progress of data collection and interviewer performance. Once data were accumulated by CATI interviewers or the IVR system, they could be immediately analyzed using standard SAS programs.

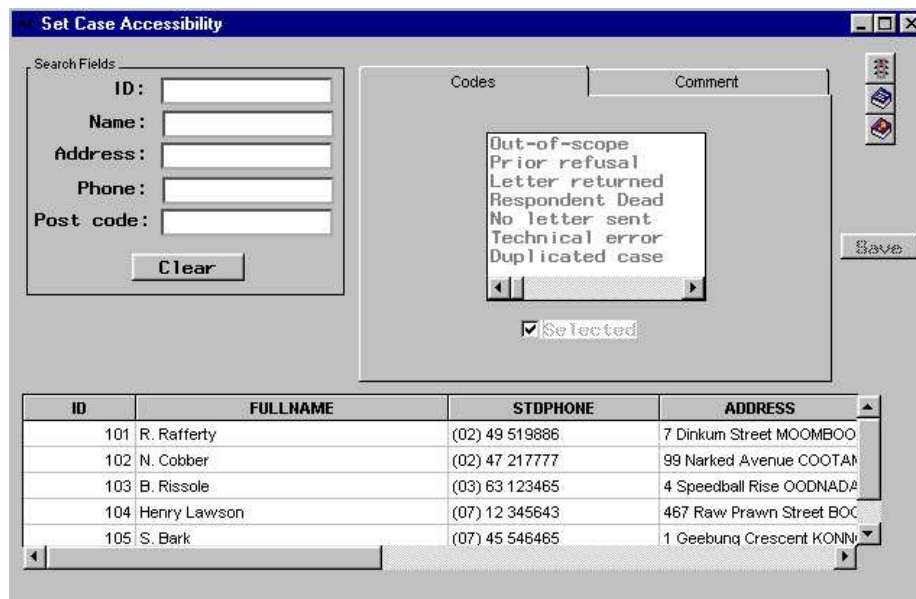


Figure 3.11: The GEIS respondent status code maintenance window.

### System maintenance

The maintenance window was opened from the control panel by selecting the ‘**Maintain**’ button (Figure 3.2). The maintenance window (see Figure 3.11) allowed users to edit interviewer details as well as modify respondent contact and control information. It did not allow respondent answers to be modified.

### Data set finalization

The output window was opened from the control panel by selecting the ‘**Output**’ button (Figure 3.2). The output window (see Figure 3.12) provided a backup facility that allowed all the files associated with a GEIS project to be saved together to another location. It could also perform other functions on the data sets to assist in later analysis, such as decoding answers to multiple-option questions.

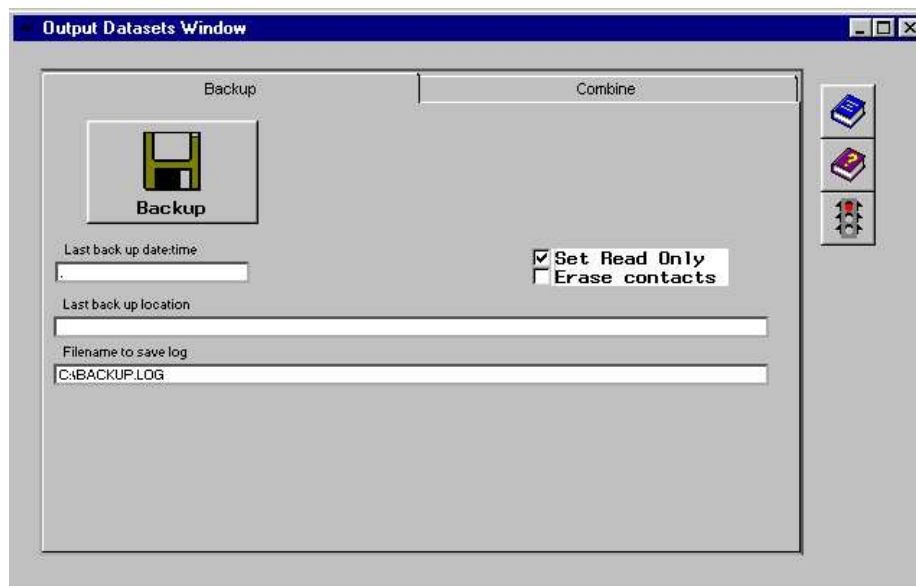


Figure 3.12: The GEIS output window.

### 3.4.2 Implementing IVR within GEIS

Additional hardware consisting of a voice card was required to introduce IVR functions within GEIS. Technical details for this are given in Appendix Section A.2. The card allowed GEIS to initiate or receive calls, play recorded voice messages stored in files, record messages spoken by respondents, and respond to touchphone key presses by the respondent. The voice recordings were stored in files that were specified in the script by the use of keywords.

As a convention in this thesis, the software is referred to as GEIS and the computer in which the voice card was installed is called Recorded Voice System (RVS).

When GEIS initiated a call to a respondent it needed to distinguish a person from an answering machine so it could respond correctly. The voice card provided two alternatives for detecting answering machines: acoustic analysis and salutation length determination. In acoustic analysis the voice card used

the quality of the sound to determine whether the line was answered by an answering machine. Initial experimentation indicated that this was not a reliable method. The preferred method was salutation length determination in which the voice card monitored how long the respondent spoke when answering a call. For example,

“Hello.”

is briefer than

“Hello, we’re not in at present.”

The first case is much shorter and would be assumed to be spoken by a person while the longer second case would be assumed to be a recording on an answering machine.

### 3.4.3 Selecting interviewees

There were several situations where selecting respondents for interviews needed to be carefully controlled. For example: (a) respondents who had completed an interview should not be re-contacted; (b) respondents who had arranged to complete the interview at a specific date and time should be called then, and not before or after; (c) numbers that were not answered needed to be called back later at intervals until contact was made; (d) numbers should not be called until any pre-contact letter, if used, had a chance to arrive at the household; (e) respondents who were known to be ineligible should not be called; and (f) respondents who had refused should not be called back.

Each survey might require a different method of selecting respondents for interview. To allow for this, *GEIS* used a series of variables that worked in com-

ination to control how respondents were selected for interview in a particular survey: `START`, `SELECTED`, `ELIGIBLE`, `STATUS`, `CBCKDATE` and `CBCKTIME`. The functions of these variables are described in Table 3.1. Only those respondents for which these variables contained appropriate values would appear in the CATI interviewer's pull-down list. The IVR system used an analogous method to select the respondents.

### 3.4.4 Status codes

The `STATUS` variable was used to store the status of interviews as they progressed through contact attempts to interview finalization. The `STATUS` variable used a two-letter coding system. The principal codes are shown in Table 3.2 and a complete list is shown in Appendix Table A.1 of Appendix Section A.3.

The status code for each case was initially set to code '`NA`' (No Attempt) and then typically progressed over a series of telephone calls through various codes, such as are shown in Table 3.3. The `STATUS` code could not be changed from a higher degree of completion to a lower one; e.g. it could not be changed from '`PQ`' to '`ET`', or from '`CB`' to '`NA`'.

Whenever an interview was stopped and recommenced at a later time, GEIS would restart the interview at the point left off. This typically occurred if the previous interview had been saved with `STATUS` code '`PQ`'.

After a fully-completed interview ended the status was normally set to code '`CQ`' (Completed Questionnaire) by a script `STAT` item (see Table 3.5 in Section 3.6.2). However, if after having reached the end of an interview, the interviewer then pressed the '**BACK**' button to return to an earlier item and changed an answer, the system would automatically reset the status back to

Table 3.1: GEIS variables used to select interviewees.

<i>Variable</i>	<i>Function</i>
START	This variable was set to the earliest date and time a respondent should be contacted. Respondents did not appear in the CATI interviewer's pull-down list and were not available to the IVR system until the START date and time had passed.
SELECTED <sup>a</sup>	If the interviews for certain cases were to be prevented indefinitely, this variable was set to zero. These respondents would then not appear in the CATI interviewer's pull-down list and were not available to the IVR system. As soon as the variable was reset to one the cases became available immediately for interviewing.
ELIGIBLE <sup>a</sup>	Some cases could be removed because of ineligibility by setting this variable to any positive value. These respondents would not appear in the CATI interviewer's pull-down list and were not available to the IVR system until the ELIGIBLE value was reset to zero.
STATUS	As interviews were completed or cases were eliminated from interviewing their condition was recorded using this variable. It is described in more detail in Section 3.4.4 on the preceding page.
CBCKDATE CBCKTIME	When a respondent made an appointment within an IVR call to be called back to complete an interview the date and time was stored in these variables. The IVR system then called them back at the due date and time.

<sup>a</sup> The **SELECTED** and **ELIGIBLE** variables performed the same function of selecting or de-selecting respondents for interview, but were used in different contexts. The **SELECTED** variable was used to select or deselect large numbers of respondents for interview whereas the **ELIGIBLE** variable was used to select or deselect single respondents for interview for specific reasons.

The **SELECTED** variable was principally set within SAS data step programs. An example was where respondents in one region were to be interviewed later in the survey than respondents in another region.

The **ELIGIBLE** variable was principally set from the GEIS control panel. An example was where a respondent had refused or an information letter had been returned. In these cases the **ELIGIBLE** variable would be set to an appropriate code.

Table 3.2: Principle STATUS codes<sup>a</sup>.

<i>Principle STATUS Code</i>	<i>Meaning</i>
'AM'	Answering machine.
'CB'	Call-back arranged.
'CQ'	Completed questionnaire.
'DO'	Dropped part-way due to refusal.
'DR'	Dropped before starting due to refusal.
'DT'	Disconnected tone.
'ER'	Error condition.
'ET'	Engaged tone.
'FM'	Fax machine.
'NA'	No attempt made to contact.
'OS'	Out of scope — ineligible.
'PQ'	Partly-completed interview.

<sup>a</sup> A complete list of status codes appears in Appendix Table A.1 of Appendix Section A.3.

Table 3.3: Example of STATUS codes in successive call attempts.

<i>Call Number</i>	<i>Call Outcome</i>	<i>STATUS Code</i>
0		'NA'
		↓
1	Call rang out without answering.	'RT'
		↓
2	Number was engaged.	'ET'
		↓
3	Call answered by answering machine.	'AM'
		↓
4	Respondent contacted and appointment made to call back.	'CB'
		↓
5	Part-interview completed.	'PQ'
		↓
6	Interview completed.	'CQ'

‘PQ’ (Partly-Completed). This status code then persisted until a script STAT item (Table 3.5) became active again and changed it again.

### 3.4.5 Data quality considerations

An advantage of an electronic interviewing system is that many checks can be placed on data entry to reduce inconsistencies and errors. The checks in GEIS can be categorized as: (a) range limit checks; (b) item non-response recording; and (c) script logic checks. These are described below.

#### Range limit checks

In GEIS, numeric quantities that could be entered were simple integers (i.e. ..., -2, -1, 0, 1, 2, ...), numbers with decimals (e.g. 4.7, 0.2663), dates (e.g. 310701, 31JUL2001), and times (e.g. 9PM, 12:45:00). Numeric quantities were subjected to range limit checks of two types: absolute limit checks and reasonableness limit checks. Both types were specified within the script. GEIS did not allow a number that fell outside absolute limits to be entered. A number that fell outside the reasonableness limits had to be confirmed before it was accepted. One or both of these limits, as well as the upper and lower values of each limit could be omitted.

#### Item non-response recording

GEIS would not display the next question until the current question had been answered. This ensured that answers to individual questions could not be omitted. However, respondents might drop out of an interview if they were required to answer all the questions, especially if some were sensitive. To avoid this, a



respondent's refusal to answer a question requiring a numeric response, or refusing to select one of a set of discrete options, was recorded using the code **.R**. The **.R** value was a special numeric value recognized by SAS. When these values were used in an item the script had to allow for them in its logic structure.

By default the SAS software omits the **.R** code when tabulating data. Also, item non-response rates could be calculated as the proportion of items with **.R** answers.

How the **.R** codes were entered by CATI interviewers and IVR respondents depended on the type of question. For example, for questions requiring a numeric response CATI interviewers entered "**.R**", whereas IVR respondents were instructed to press a specific touchphone key, such as key '0'. This is further discussed in Section 3.6.2 and Appendix Section A.7.4.

### Script logic checks

After a script was imported, GEIS compiled it into a format that it could use to generate the CATI or IVR interface. This process involved creating data sets, catalogues, and checking the script for errors. Compiler error and warning messages were written to the interviewing log. The compiler checks were very extensive and a summary is given in Appendix Section A.1.

### 3.4.6 Call scheduling

CATI interviewers would typically ring back non-contacted numbers at intervals until contact was made, a call limit was reached, or the project ended. Since the IVR system operated automatically it used a call scheduling algorithm shown in Figure 3.13. The algorithm was based on the optimal scheduling algorithms of

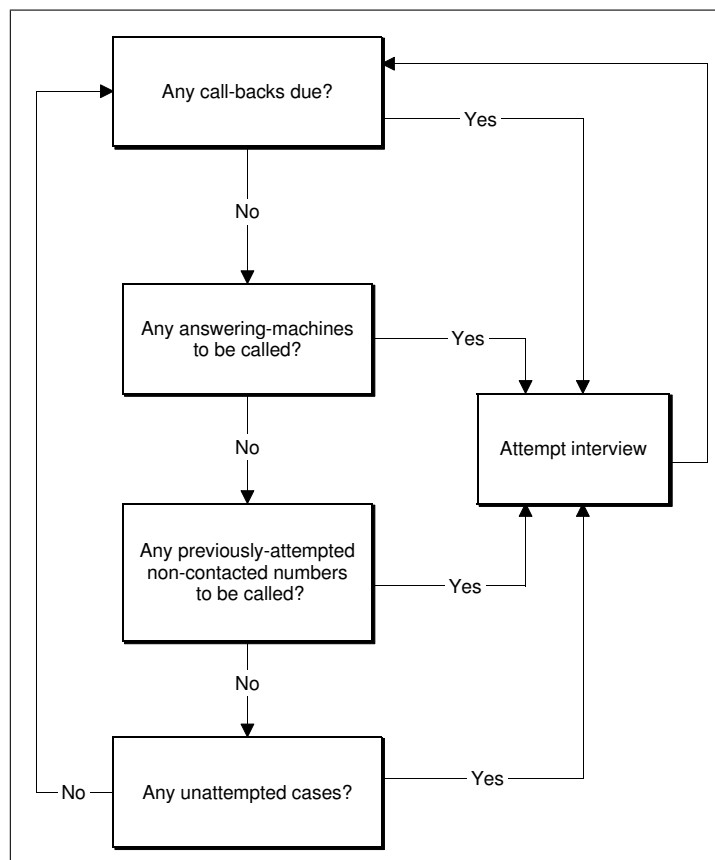


Figure 3.13: IVR call scheduling loop.

Kulka and Weeks<sup>188</sup> and Steel et al.<sup>52</sup>. The algorithm assigned cases to the four categories: (a) call-backs due, (b) answering machines, (c) previously-attempted non-contacts, and (d) unattempted cases. Except for call-backs that were called when they became due, calls were made at alternating short and long intervals. Thus, if no contact was made, a number would be rung thirty minutes later. If no contact was made on the second attempt, it would be re-rung eight hours later, or even later. If still no contact was made, the number was re-rung thirty minutes later, and so on. Further information on the algorithm can be found in Appendix Section A.4.

### 3.4.7 Call cost

Information was retained by GEIS to allow the cost of calls to be calculated. In Australia, the cost of telephone calls depended on the call duration, time of day, and distance between the caller's and called-party's telephone exchanges<sup>189</sup>.

The distance was taken to be the great circle distance from the local Wallsend exchange to the respondent's telephone exchange, as given by Equation 3.1<sup>190</sup>. The longitude and latitude of exchanges in Australia were available from Telstra<sup>189</sup>.

$$D = 1.852 \times 60 \times \arccos \{ \sin(L_1) \sin(L_2) + \cos(L_1) \cos(L_2) \cos(D_g) \} \quad (3.1)$$

where:  $L_1$  = latitude of the Wallsend exchange ( $-32.9^\circ$ ),

$L_2$  = latitude of the respondent's exchange ( $^\circ$ ),

$D_g$  = longitude of the Wallsend exchange ( $151.7^\circ$ ) minus the longitude of the respondent's exchange ( $^\circ$ ),

$D$  = computed distance (Km).

The distances were saved for all exchanges in Australia to a permanent SAS data set, DISTANCE (see Appendix Section A.5.2). The compiler obtained the distance to each respondent's location by searching the DISTANCE data set for the respondents' telephone number prefixes.

## 3.5 Data management in GEIS

### 3.5.1 Data storage in GEIS

There were several GEIS data sets, as shown in Table 3.4. When a script was imported, it was saved to the SCRIPT data set. During compilation, GEIS used

Table 3.4: The major GEIS data sets<sup>a</sup>.

<i>Name</i>	<i>Content</i>
CONFID	This data set contained the respondent contact details, including name, telephone number, and address.
SCRIPT	This data set was created when the script text file was imported. It contained all information needed to define the questionnaire, such as: (a) question texts, (b) allowable options for closed-form questions, (c) range-limits tests for numeric questions, and (d) the logical relationships between all items in the questionnaire.
ANSWERS	All answers given by respondents were stored in this data set.
CONTROL	This data set held the information used by GEIS to control the interviewing process.
INTRVS	The accumulated time each interviewer spent interviewing was stored in this data set.
ILOG	All system events and errors were stored in this data set.

<sup>a</sup> A complete listing of the GEIS data sets appears in Appendix Section A.5.

information in the CONFID and SCRIPT data sets to create all of the other data sets. The data sets are described in more detail in Appendix Section A.5.

### 3.5.2 Record locking

Only a single interviewer could access a particular respondent's data at one time. This eliminated the possibility that two interviewers might make simultaneous attempts to ring the same respondent. To do this, when the interviewer selected a respondent, the respondent's ID value was used to locate the appropriate record in all data sets and those records were then locked preventing them being modified by another user or program. To illustrate, in Figure 3.14, the records for ID 105 in four GEIS data sets have been locked.

CONFID		CONTROL		ANSWERS		PROTECT	
ID	SELECTED	ID	STATUS	ID	Q2	ID	Q2
101	1	101	DR	101	20	101	Y
102	1	102	DR	102	20	102	Y
103	1	103	DR	103	20	103	Y
104	1	104	DR	104	20	104	Y
105	1	105	DR	105	20	105	Y
106	1	106	DR	106	20	106	Y
107	1	107	DR	107	20	107	Y

Figure 3.14: Example of record-locking in multiple GEIS data sets.

### 3.5.3 Multiple data access

Several CATI interviewers typically conducted interviews simultaneously. Similarly, the IVR system might conduct up to four interviews at the same time. Multiple simultaneous access to the GEIS data sets in Table 3.4 was needed to do this. This was done by using the record-locking method described in Section 3.5.2 to lock individual respondent's records, and by means of a SAS share server<sup>191</sup>.

The share server was a program run on a computer connected to the same network to which the CATI workstations and IVR system were connected (see Appendix Figure A.1). The network allowed communication between the workstations, IVR system and share server. Using a share server meant that the data sets could also be accessed by other software when data were being collected. This allowed analysis programs to be run when interviewing was occurring on the CATI workstations and IVR system. The program written to control the share server and an example of a program used to access data on a share server is shown in Appendix Section A.6.

## 3.6 **GEIS** scripts

The structure (Section 3.6.1), item types (Section 3.6.2), logic statements (Section 3.6.3), and question texts (Section 3.6.4) of **GEIS** scripts are described below. More technical information about **GEIS** scripts is given in Appendix Section A.7.

### 3.6.1 Script structure

A **GEIS** script contained all the information needed to define the interview, including all:

- question texts;
- option texts and codes;
- range check information;
- logical question-branching information;
- interview termination status information;
- DO-loop information; and
- IVR information.

The **GEIS** script also provided a standard, non-proprietary, open method of project documentation. This meant that comparisons between survey projects could be achieved by publishing the scripts.

The script could be created in any text editor. The file was then imported into **GEIS** and saved to the **SCRIPT** data set. During the import process **GEIS** displayed a window (Figure 3.4) that repeated the script information to allow

CHCE	7	D_05	5	_MAKE_	LABEL
DEMOG	EDUCAT				
DOB>.					
What is the highest level of schooling you have completed:					
never attended school, primary school, secondary school,					
Technical, Trade, Vocational Certificate, or Diploma,					
University or College of Advanced Education degree,					
or some other kind of education?					
1	No education				
2	Primary school				
3	Secondary school				
4	Technical/ Trade/ Vocational Certificate or Diploma				
5	University or CAE degree				
6	Other Education				
.R	Refusal				
Education level					
*****					

Figure 3.15: Example of a GEIS script item.

the user to confirm the script had been read correctly. It also checked that the script conformed to the correct syntax and showed any warning and error messages as needed.

### 3.6.2 GEIS script item types

A GEIS script contained a series of items that corresponded to questions in a questionnaire or performed other functions. An example of a script item is shown in Figure 3.15. The example defines a question that asks about the respondent's education and which allows seven possible answers.

A script item's first and second lines were used to specify various options. In the example in Figure 3.15, the item's type is 'CHCE' and its name is 'D\_05'. A CHCE item type was used to specify a question with a restricted set of possible answers. The third line was used to specify an expression called the self-protection statement (Section 3.6.3). In the example this is 'DOB>.'. The next lines were used to specify the question text. In the example this is 'What is the highest

level of schooling you have completed: never attended...'. Following the question text were settings that depended on the item type. The question in the example only allows a restricted set of possible answers. These are given with their corresponding numeric codes. For example, '3 Secondary school'. One of the numeric codes is '.R', which was the code to be recorded if the respondent refused to answer (see Section 3.4.5). The penultimate line was used to specify the item's label assigned to the variable when the answer was saved to a SAS data set. In the example this is 'Education level'. The last line was a line of asterisks that visually divide the script items from each other.

The item type CHCE shown in Figure 3.15 indicated that the item was a question with a restricted series of possible answers. Table 3.5 shows other possible question types that could be specified in a GEIS script. Table 3.5 also lists some other item types that did not correspond to questions, but served additional purposes such as sending e-mails. Table 3.5 only shows the major item types; a complete list is given in Appendix Section A.7.4 and their syntax is given in Supplementary Materials Section 3.1.

GEIS configured the interviewing system according to the item type and survey mode. As an example, including a CHCE type (single-choice question) in the script caused GEIS to display a list box containing the possible answers of which only one could be selected. In the IVR case, GEIS only accepted answers that conformed to the item type. For example, for a CHCE type it would only accept a single touchphone keypress.

Each question item type could accept refusals, which is discussed in Appendix Section A.7.4.



Table 3.5: The major GEIS script item types<sup>a</sup>.

<i>Item</i>	<i>Item Function</i>
TITL	Specified the title of the project and was used to set global options.
CHCE	Used to specify a question that might be answered by exactly one of a limited set of possible answers.
INFO	Used to supply instructions to the interviewer or information to the respondent.
MAIL	Used to send an e-mail to a nominated recipient after the interview was completed.
MULT	Used to specify a question that might be answered by one or more of a limited set of possible answers.
NUM	Used to ask a question that required a numeric answer, such as a date of birth.
OPEN	Used to ask a question that required an open-ended answer.
STAT	Used to set the status of the case when the interview terminated: completed, refused, partly-complete, etc.
TABL	Used to ask a question that required multiple numeric or open-ended answers.

<sup>a</sup> A complete listing of GEIS script item types is given in Appendix Section A.7.4.

### 3.6.3 Self-protection statements

Questions within an interview appeared in a particular order. The selection of the question could have depended on the answers given beforehand, or there might have been an unvarying question order. In GEIS, a CATI interviewer might also have moved back several questions to a previously-entered answer and altered it. The determination of which question to display was done in GEIS by assigning a Self-Protection Statement (SPS) to each item.

A Self-Protection Statement (SPS) was a single logical statement that formed part of each item and which always evaluated either to true or false. The name, ‘Self-Protection Statement’, was selected because if the statement was false then the associated item was deemed to be protected and could not be displayed. Conversely, if the SPS was true then the associated item was deemed unprotected and could be displayed.

The example in Figure 3.16 shows the logical relationships between four questions. Each question has its own SPS. The only allowable answers to Q1 are “male”, coded as one, and “female” coded as two. The item Q2 in this example accepts any numeric response. The only allowable answers to Q3 are “yes”, coded as one, and “no” coded as two. The possible answers for question Q4 have been omitted.

The SPS for Q2 is “Q1 NE .”. In the SPS, the term “Q1” refers to the answer given to question Q1, the term “NE” means “not equal to”, and the term “.” means “not-answered”. Thus, the interpretation of the SPS is: Q2 is unprotected and displayable only if Q1 has an answer. So either answering “male” or “female” suffices for GEIS to display Q2.

The item Q3 has the SPS “(Q1 NE 2) and (Q2 NE.)”. This means that

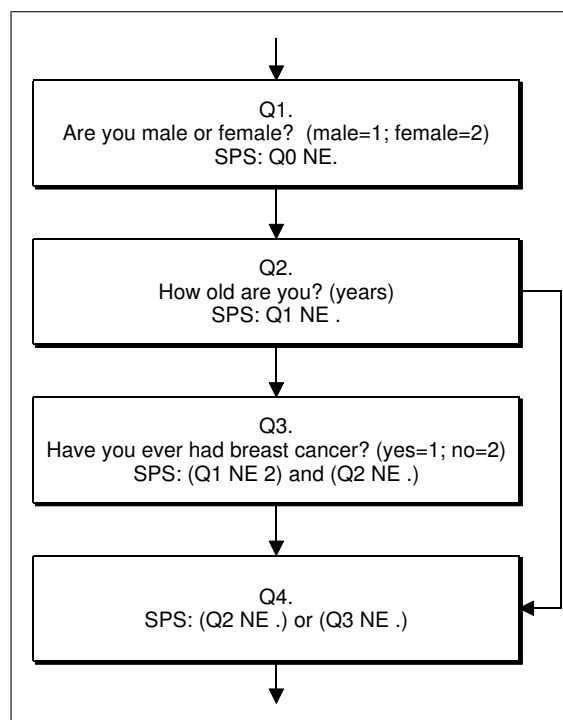


Figure 3.16: Example of the logical relationships between four questions, their self-protection statements, and option-code values.

Q3 may only be displayed if the answer to Q1 is “female” and Q2 is answered. If the answer to Q1 had been “male” Q3 would not be displayed. This illustrates that a SPS in GEIS could make use of the answer to any preceding question.

The last item Q4 has the SPS “(Q2 NE .) or (Q3 NE .)”. This means that Q4 may only be displayed if Q2 had been answered or Q3 had been answered. Overall, male respondents would be asked questions Q1, Q2, and Q4; female respondents would be asked questions Q1, Q2, Q3, and Q4.

The use of Self-Protection Statements (SPSs) permitted the complex logic structure of a questionnaire to be condensed to simple statements, one for each item in the questionnaire. They are described in more detail in Appendix Section A.7.3.

### 3.6.4 Answer-quoting in question texts

The question texts could be modified dynamically to fit each respondent’s situation. An example would be the question

“Is the telephone number ‘1234 5678’ the correct number to ring?”

The process by which GEIS did this was termed answer-quoting. Answer-quoting was the process of inserting the values of other variables into the question text. Often, these values were the answers to preceding questions.

To quote an answer to a previous item, the previous item’s name was inserted in the question text and surrounded by two carets. For example, “Last time you said ^Q^.”. If the quoted item was numeric then simple arithmetic operations could also be defined, such as ^Q5-3^, which subtracted 3 from the value of Q5 and then inserted the formatted value into the question text.

Answer-quoting could also be used for IVR interviews to dynamically modify the recorded messages. This is explained in Appendix Section A.7.2.

### 3.7 Conclusions

GEIS consisted of a series of windows that allowed scripts to be imported, compiled and run, projects to be maintained, and data sets finalized. Respondents were selected or deselected by means of a series of variables according to the allowable dates and times, and states of interviews. The system incorporated data quality checks including range limit checks, not allowing any eligible question to be skipped, and a comprehensive series of logic checks. Calls to respondents in the IVR mode were made according to the state of interviews, and date and time of previous attempts. Cost of calls was calculated using the distance to the respondent's telephone exchanges and the time of the day the call was made. Data were stored in SAS data sets and managed by a share server to control multiple access by interviewers. Scripts included all information needed to define the interviews, allowed for various types of questions to be asked, incorporated variable questions by means of answer-quoting, and permitted a complex logic structure to be specified.

The development of this software allowed research to be conducted with CATI and IVR using a common script syntax. The next stage in the study was to evaluate the reliability and acceptability of GEIS when making outbound or receiving inbound calls. This is described in the next chapter.

## Chapter 4

# Developing and testing the Interactive Voice Response system

### 4.1 Introduction

This chapter describes the development and testing of the Interactive Voice Response (IVR) interface.

It was necessary to develop and test Generalized Electronic Interviewing System (GEIS) and its capacity to conduct interviews using both IVR and Computer-Assisted Telephone Interview (CATI) as the planned main study involved both an IVR and a CATI component. Development and testing were done alternately, with changes being made as the software was tested.

Although both interview modes (IVR and CATI) required testing, it was considered that the CATI mode was sufficiently well developed. The Hunter Centre for Health Advancement (HCHA) was invited to use GEIS to support its CATI work. During the period 1999–2001 a total of 5,472 CATI interviews were completed using all item types to be used in the proposed main study. Any problems that arose with GEIS were solved and difficulties experienced by users

and interviewers corrected during this period. The CATI interface was found to be working correctly and had proved acceptable to users.

The IVR system could be used in two ways of immediate interest: (a) to make outbound calls and directly recruit respondents, or (b) to receive inbound calls from respondents who had already been invited to participate. Both uses of the IVR interface were novel and therefore needed testing. This was to be conducted initially using staff at the HCHA, and later with a small community pilot sample to examine acceptability and the response rate.

## 4.2 Aims

The aims of this chapter are to describe:

1. The testing of outbound dialing;
2. The testing of inbound dialing; and
3. The acceptability of the system with a community sample.

The three aims were to be addressed in three stages: Pretest One, Pretest Two, and Pilot One. As discussed in Pilot One (Section 4.5.6), it was later found necessary to conduct a second pilot study, called Pilot Two.

## 4.3 Pretest One: Testing the outbound mode

### 4.3.1 Introduction

In order to use IVR for household surveys it was necessary for the computer to make outbound calls and directly recruit respondents. The computer needed

to do this without any assistance by a person. In outbound dialing, GEIS had to dial the respondent's telephone number and then determine whether the call had been answered by a person or an answering machine. If the call was not answered it needed to be able to reschedule the call for another time. It also needed to be able to return calls at dates and times nominated by respondents.

During a call, respondents needed to be able to answer single-choice questions and numeric entry questions. However, as Chapter 1 has discussed, in both CATI and IVR, respondents listened to spoken questions without having the questions in written form, which placed demands on the respondent's memory. During a CATI, an interviewer can repeat questions when a respondent has difficulty in understanding a question. It was unclear whether respondents would be able to repeat questions in an IVR interview.

### 4.3.2 Pretest One aims

Aims for the pretest were to:

1. Assess GEIS's capacity to distinguish respondents from answering machines;
2. Assess if respondents could reschedule calls to more convenient times;
3. Assess if respondents could respond to single-choice questions and numeric entry questions;
4. Assess if respondents could repeat questions; and
5. Assess the method's acceptability to respondents.



### 4.3.3 Methods

#### Design

A survey of alcohol consumption of staff and students working at the HCHA was conducted using the outbound IVR method. The interviews were not intended to assess the alcohol consumption of the participants. Instead, the participants were invited to critically examine the interview experience.

#### Sample

The HCHA's electronic telephone listing was used to prepare a sampling frame. Records were deleted from the frame if: (a) an individual was not working at the HCHA; (b) the address was not that of the HCHA; or (c) there was no telephone number or name. A simple random sample of 20 individuals was then selected. This was a sufficient sample size to assess any practical difficulties. Two volunteers were then added, giving a total sample of 22. Adding volunteers was acceptable since the aim was to investigate the practicality of the method and not its acceptability to a general sample of people. The sampling program is shown in Appendix Section B.1.1.

#### Setting

This was an in-house pretest of staff and students from the HCHA and the Discipline of Behavioural Science in Relation to Medicine, at The University of Newcastle.

## Procedure

**Respondent recruitment and instruction** An e-mail was sent to the participants three working days before the first interview. Attached to the e-mail were an instruction sheet and an assessment form. (Figures B.1, B.2, and B.3 of Appendix Section B.1.2). The instruction sheet explained the purpose of the project and the role of the participants. It also described how to answer questions by pressing touchphone keys, repeat questions using the hash key (#), and return to the previous question using the star key (\*). The term ‘hash key’ may not have been familiar to respondents since it has also been called the pound key or octothorpe key<sup>192</sup>. To minimize confusion, the instruction sheet included a picture of a telephone keypad that clearly showed the hash and star keys (see Appendix Section B.1.2).

**Call scheduling** Initially, GEIS attempted to call each number for ten seconds (about four rings) before disconnecting. This was done to avoid triggering the HCHA’s voice mail system. However, since the voice mail system was in fact still triggered in several cases the calling time was later changed to five seconds. At intervals of one minute GEIS checked if any calls were due. Since it could be assumed that most respondents would be available during work hours, non-contacted numbers were retried every thirty minutes between the hours 1000–1700.

**Respondent identification** On answering a call, GEIS played an introductory message and then asked the respondent to press key ‘5’. If a key press was not detected within twenty seconds a message was played thanking the respondent and stating that they would not be able to take part and the call was then

terminated. This was done to allow for the possibility that some telephones were either not touchphones or were not set to tone dialing.

If a key press was detected a message was played that described the survey purpose. The message explained that GEIS could call back later if the current time was not convenient or the respondent was unavailable. By default, GEIS rang back the next day at the same time, but the respondent could change this to an alternative date and time up to thirty days in the future.

### **Apparatus**

A simplified version of the IVR hardware set up described in Appendix Section A.2 was used to minimize management problems for the pretest. Only a single line was considered necessary since the sample size was small.

### **Measures**

**Interview script** The interview script (Supplementary Materials Section 4.1.1) included items to handle the respondent identification procedures above. Also included were demographic items (date of birth, employment status, marital status, country of birth, aboriginality, age, and sex) and a five-item Alcohol Use Disorders Identification Test (AUDIT)<sup>193–195</sup>. The AUDIT and demographic items were used as examples of items from the proposed main survey.

For most single-choice questions, the number of answer options to items was restricted to five or less according to literature recommendations for telephone-based interfaces<sup>196</sup>. The demographic questions were used to test the practicality of single-choice questions and numeric entry questions and were implemented using the CHCE and NUM item types discussed in Section 3.6.2 of Chapter 3.

To minimize dropouts respondents could refuse to answer a question by pressing an indicated key. Which key was pressed depended on the question. For example:

“If you are male, please press key one.”

“If you are female, press key two.”

“If you do not wish to answer, press key three.”

Both age and date of birth were asked to provide a means of assessing the practicality of numeric data entry. The date of birth was entered as either a six or eight digit number in European format. For example, 23 October 1960 could have been entered as 291060 or 29101960. The age was entered in whole years. The age item was the first demographic item and date of birth was near the end of the demographic items.

Movement through the interview script was controlled by the hash (#) and star (\*) keys. If the current question had not been answered, pressing the hash key caused it to be repeated. However, if the current question had been answered pressing the hash key would cause the next question to be asked. Pressing the star key (\*) caused GEIS to move back to the previous question.

**Voice recordings** Voice recordings were made in 16-bit mono 11kHz format with an Optimus® omnidirectional 33-3104 microphone and Creative Sound Blaster® Vibra 128 sound card. Recordings were made using the candidate’s voice.

**Assessment form** After completing the interview, respondents completed a paper assessment questionnaire (Figure B.3 of Appendix Section B.1.2). The

questionnaire asked respondents if they realized initially that they were not talking to a person, if they had difficulties with data entry, whether system errors occurred, if they had to repeat questions, whether they arranged and received call backs, and how comfortable they were in using the system. The respondent was also asked for general comments on the interview.

**Interview control** Where an interview was completed the interview status code was set to ‘CQ’ (Completed Questionnaire). Any cases that were scheduled by the respondent to be completed at a later date were coded as ‘CB’ (Call Back). If the respondent hung up after starting the interview the status was set to ‘PQ’ (Partly-completed Questionnaire). For the partly completed interviews an automatic call back was set for the same time the next day. When partly completed interviews were recommenced GEIS played an explanatory message and then started from the point left off.

### **Analysis**

The number of calls required to make contact and to complete an interview was collated using the CONTROL data set. The number of cases where call backs were set by respondents was also enumerated. Answers and comments in the assessment form in respect to data entry and acceptability were coded after collection and frequency tabulations produced.

#### **4.3.4 Results**

Interviews were conducted between the hours 1044–1549 in August 1999. A total of 21 (95%) completed responses and follow-up assessments were obtained. A single remaining case was set as a call back but the person was not available when

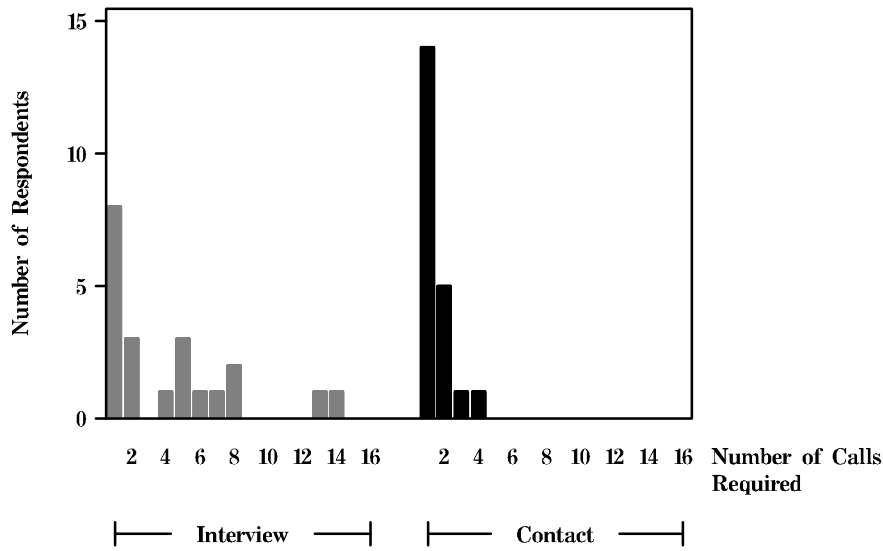


Figure 4.1: Pretest One frequency of the number of calls required to complete an interview and to contact a respondent.

the second call came due. In the follow-up assessment, 12 (63%) respondents reported that they realized immediately they were not talking to a person.

### Contacting respondents

All respondents were contacted and half completed the interview within three call attempts (Figure 4.1). A total of 75% completed the interview within six calls, but two cases took 13 and 14 calls. Only one respondent did not complete the interview due to non-availability. A total of four respondents arranged for GEIS to call them back.

### Call backs

A total of four (25%) respondents explicitly arranged for GEIS to call them back. In the follow-up assessment, none reported finding the process confusing, but one found the process too lengthy. None reported the call backs arriving at the

incorrect time, although not all respondents were available at the time it came.

### **Data entry**

Four respondents (25%) thought there should have been a greater interval between successive items. Six (32%) thought that keystrokes should be echoed back. Three (16%) respondents thought there were questions with too many options.

Two measures of the respondents' age were available: their age as entered in whole years, and their age as determined from their date of birth. There were three cases (14%) where the respondent's age as calculated from their entered date of birth did not agree with their entered age. From comments on assessment sheets two of these differences can be attributed to the respondents attempting to test GEIS's cross-validation capacity.

### **Repeating questions and acceptability**

Responses to the assessment questionnaire are shown in Table 4.1 and Table 4.2. The results are summarized in Table 4.3.

## **4.3.5 Discussion**

### **Contacting respondents**

The call scheduling system successfully contacted all but one respondent who was not available during the pretest. During this time the scheduling system continued to attempt to contact respondents until the whole interview had been completed. Thus, the pretest satisfactorily confirmed GEIS's capacity to contact respondents.

Table 4.1: Pretest One assessment results summary.

<i>Question</i>	<i>Yes</i>	<i>%</i>
When you initially answered the call did you realize you were not talking to a person?	12	63
Were there any questions with too many options so that it was difficult to remember them all?	3	16
Do you think there should have been a greater space between questions?	4	25
Did you have any problems in responding to questions by pressing the number keys?	1	5
Did you repeat any questions using the hash key?	6	32
If you repeated questions, did you find the process confusing?	1	17
Did you arrange a call-back to a later interview?	5	26
Did you find arranging call-backs confusing?	0	0
Was the process of setting call-backs too lengthy?	1	20
Did the call back arrive on the date/time you wanted?	5	100
When the call back came, did you find that answering the call was confusing?	1	20
Did you hang up after starting an interview or before pressing any keys?	0	0
Did the system do anything unexpected, such as fall silent or hang up suddenly?	3	17
Did you find that it would have been helpful if the system could have repeated each keystroke back to you?	6	32

Table 4.2: Pretest One agreement with feeling comfortable with the system.

<i>Agreement with feeling comfortable with the system<sup>a</sup></i>	<i>Yes</i>	<i>%</i>
Strongly agree	6	32
Agree	10	52
Neutral	3	16
Disagree	0	0
Strongly disagree	0	0

<sup>a</sup> Assessment question: ‘How do you feel about the statement: “I felt comfortable in using the system”?’.



Table 4.3: Pretest One summary of respondents' comments.

<i>Category</i>	<i>Comment</i>	<i>n</i>	<i>%<sup>a</sup></i>
<b>Technical error</b>	Problem with voice mail	3	14
	Identify respondent better	4	19
	Script error	2	10
		<b>8</b>	<b>38</b>
<b>Call back was convenient</b>		<b>3</b>	<b>14</b>
<b>Disliked/Impersonal system</b>	Impersonal start	1	5
	Impersonal interview	3	14
	Abrupt end	2	10
		<b>5</b>	<b>24</b>
<b>Liked the system</b>		<b>3</b>	<b>14</b>
<b>Poor voice</b>		<b>4</b>	<b>19</b>
<b>Liked the voice</b>		<b>4</b>	<b>19</b>
<b>Needs feedback</b>		<b>4</b>	<b>19</b>
<b>Does not need feedback</b>		<b>2</b>	<b>10</b>
<b>Too long/slow</b>	Interview too long	2	10
	Needs a progress indicator	2	10
	Interview too slow	1	5
		<b>5</b>	<b>24</b>
<b>Too fast</b>	Started too quick	2	10
	Spoken too fast	1	5
		<b>3</b>	<b>14</b>

<sup>a</sup> The categories are not mutually exclusive.

**Call backs**

The call back procedure was successful, if rather lengthy. In this process:

1. The respondent was offered a default date that could be amended forward or backward a day at a time; and
2. A default time was offered that could be amended forward or backward an hour at a time.

It was concluded that a better method would have been to offer to ring back the next day at the same time. This could be accepted with one key press. If rejected, the system could then proceed to offer to set a particular date and time.

**Data entry**

Although some respondents thought that keystrokes should be echoed back, others commented that this would slow down the interview too much. It was suggested by some that feedback be provided as data are entered for people with visual disabilities. For example, GEIS might repeat back to the respondent each touchphone key they pressed. A few thought this would slow the pace of the interview too much.

However, no particular difficulty was encountered with data entry, and feedback would also increase the complexity of script development since each unique option would need to be recorded. It was concluded that feedback was unnecessary, although it could be of use in some cases, such as studies of the elderly.

Some confusion was caused by an item asking about the number of alcoholic drinks consumed. The item had five possible keys to press corresponding to the

range of number of drinks (key ‘1’ for 1–2 drinks, key ‘3’ for 3–4 drinks, key ‘5’ for 5–6 drinks, key ‘7’ for 7–9 drinks, key ‘0’ for 10 and more drinks). This differed from all other items which used ascending digits, 1, 2, 3, 4, ..., for answer option 1, answer option 2, answer option 3, answer option 4, etc. It was concluded that ascending digits should have been used for all items.

Since only a single discrepancy was found between the calculated age and entered age it appeared that entry of complex numbers did not cause a significant problem. This meant that it would be feasible to include these types of questions in further surveys.

### **Repeating questions**

Six respondents repeated questions using the hash key and only one reported a problem. The single respondent had pressed the star key (\*) instead of the hash key (#). It was concluded that respondents could feasibly use the star and hash keys for repeating questions.

### **Acceptability**

Three of the technical errors that arose related to difficulties with the system leaving messages on voice mail. This was due to call backs that had been arranged by the respondent or automatically set by GEIS, but which were answered by the voice mail system. This suggested that when GEIS detected an answering machine it should hang-up without leaving a message and if a call back was due it should reschedule for some time later.

For some people, the introduction was judged as impersonal, while the ending was too abrupt. This was thought to relate to the timing of messages. Once the

respondent had pressed a key the system would immediately begin talking or hang up. If the respondent had been talking to a person there would have been a delay in their response, but GEIS responded instantly. The lack of a delay between questions may have caused some difficulty. It appears to have confused some respondents who, as a result, may have missed the first few words of each message. It was concluded that the system should insert a brief delay between obtaining a response from the respondent and proceeding to the next action.

Some respondents asked for feedback on how far they were through the survey or reported that the interview took too long. This may have been the result of the deliberately slow enunciation of the IVR messages and options. The suggestion that a message indicating how much of the interview remains to be completed may be a useful tool for increasing the friendliness of the system. It was also concluded that the messages should be spoken at a more natural pace and that there should be occasional messages to indicate progress through the interview.

#### 4.3.6 Conclusions of Pretest One

GEIS was found to have achieved the pretest's aims, but some modifications were needed:

1. GEIS satisfactorily distinguished respondents from answering machines and so no changes were made to this aspect;
2. Since the call scheduling system was slow to use, a single key-press function to ring back the next day was added;
3. Respondents successfully entered data using single-choice questions and

complex numeric entry questions so no changes were made to this aspect;

4. Respondents successfully used the star and hash keys for repeating questions so this was not modified; and
5. Although the method was acceptable to respondents, a few changes were made. It was decided to: (a) insert a brief delay between obtaining a response from the respondent and proceeding to the next action, (b) introduce reminder items to indicate progress through the interview, and (c) record messages at a more natural pace.

## 4.4 Pretest Two: Testing the inbound mode

### 4.4.1 Introduction

It was considered worthwhile developing an inbound mode for non-survey purposes. In this mode GEIS would receive calls rather than make them. It was thought that this would be a useful function for a number of areas. Examples of where it might be used are: (a) Demonstrating IVR interviews to interested parties who could be invited to ring into it at their convenience; (b) Providing a dial-up information service to give advice on a series of health topics; (c) Providing an alternative in an outbound survey for respondents to ring into when they are unable to complete an interview at the time the outbound call was made; (d) Remote monitoring of GEIS by having it play system report messages when called; and (e) Corporate in-house surveys where staff ring in to receive advice on, for example, occupational health and safety topics.

Inbound IVR systems have already been used in a number of published studies. Some examples include: (a) use as a research tool<sup>2,33,40</sup>; (b) provision

of health advice<sup>32,37,44</sup>; (c) daily monitoring of alcohol or drug use<sup>9,36,42,46</sup>; and (d) treatment or diagnosis of psychological conditions<sup>10,34,39,43,45,47,48</sup>.

The inbound mode would not be useful for household survey applications because the sample obtained would be self-selected. Since the probability of volunteers selecting themselves is unknown, any statistical analysis conducted on such data would be invalid<sup>197</sup>. However, the inbound mode has been used in official surveys of businesses<sup>198-200</sup>.

To test the inbound mode, an in-house pretest was conducted (Pretest Two). It was decided that the system would run continuously until a call was received, and, to avoid accidental dial-ins, respondents would have to enter a brief numeric security code before they could use the system. Once the respondent had successfully logged in, the interview would be identical to Pretest One. Therefore, only those aspects specific to inbound dialing needed to be tested.

#### 4.4.2 Pretest Two aims

This pretest was intended to test practical aspects of the IVR interface with respect to receiving inbound calls.

The specific aims were to assess the:

1. Capacity of respondents to dial into the system;
2. Feasibility of a numeric security code; and
3. Acceptability of the method to respondents.

### 4.4.3 Methods

#### Design

A survey of alcohol consumption of staff and students working at the HCHA was conducted using an inbound IVR method.

#### Sample

A list of all network users was generated using the JRB Utilities<sup>201</sup> command “LISTOBJ \* /f”. From this, a sample data set of 141 people was created using the program shown in Appendix Section B.2.1. This was more than in the sampling frame from the previous pretest (n=61) since the number of HCHA staff with network accounts was larger than the number of those with their own telephone.

#### Setting

All staff with an e-mail identity at the HCHA were considered eligible. This approach was used since the use of e-mail made contacting respondents efficient. It was not necessary that respondents had a personal telephone since they could ring in to the system from any telephone.

#### Procedure

**Respondent recruitment and instruction** An e-mail was sent to all identified network users three working days before the first interview. The e-mail invited the person who received it to ring in to the system. As in Pretest One, the e-mail text included instructions and an attached assessment form (see Figures B.4 and B.5 of Appendix Section B.2.2).

**Respondent identification** A security system was needed to ensure that only the respondents in the trial could use the IVR system. All respondents already had a unique network user name and so this was used as their security code. This avoided having to distribute security codes that could be lost, confused, or forgotten by respondents. It was considered likely that most people would be able to remember their user name. Additionally, this would allow a test of alphabetic data entry. The user names consisted of the person's first name followed by the first letter of their last name. Thus, 'Ross Corkrey' had a user name of `rossc`.

Entering user names is possible with touchphone keypads since most have keys labelled with letters as well as numbers. Most touchphones have consistent alphabetic key labeling, but this cannot be assumed to be generally true<sup>196</sup>, and many respondents may not be familiar with textual entry using a telephone keypad<sup>202</sup>. Most of the telephones in the HCHA had the keypad layout shown in Figure 4.2, but there were a small number with the layout shown in Figure 4.3. Fortunately, using either of these key layouts produced the same security code for all but one of the user names in the sample. For the single case where the security code differed between the two systems, the person was separately advised of the code to use. As an example of a security code, the user name `rossc` corresponded to a security code of `76772`. The user names were converted to security codes using the SAS program shown in Appendix Section B.2.1.

On answering a call, GEIS played an introductory message and then asked the respondent to enter their security code. Entry of incorrect codes, codes for completed interviews, or two simultaneous attempts using different telephone lines were blocked. If a correct code could not be entered within three attempts,





Figure 4.2: Keypad of an Interquartz facility phone: Model 9837.

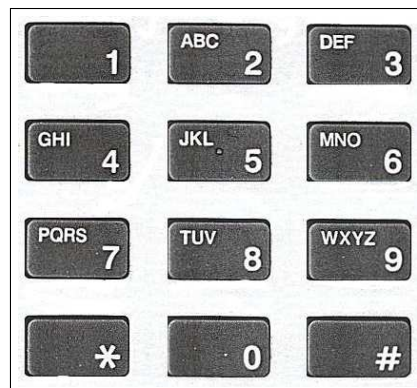


Figure 4.3: Keypad of an Interquartz facility phone: Model IQTEL300.

the respondent was advised that they would be unable to use the system. After a successful security code entry the script (Supplementary Materials Section 4.2.1) was run. All call attempts, whether successful or not, were logged to the ILOG data set (see Section 3.5.1 of Chapter 3), but only successful calls resulted in the entry codes being stored.

### **Apparatus**

**Equipment** The same computer hardware set up as in Pretest One was used, except that two lines were used to receive inbound calls. Other equipment remained the same.

**Software** The IVR interface described in Appendix Section A.2 was used. A SAS/SHARE server as discussed in Section 3.5.3 of Chapter 3 was used to coordinate the data transactions and monitor the system.

### **Measures**

**Interview script** The interview script is shown in Supplementary Materials Section 4.2.1. The interview questions were the same as in the Pretest One script except that there was no CALL, respondent identification, or call back items.

**Voice recordings** Voice recordings were made by a female volunteer who was not employed at the HCHA. The recordings were deliberately made with a slower enunciation than in Pretest One. This was done because of feedback from respondents that the Pretest One recordings had been spoken too quickly.

**Assessment form** After completing the interview, respondents were asked to complete and return a paper questionnaire (see Figure B.5 of Appendix Section B.2.2) that addressed difficulties with data entry, repeating questions, system errors, how comfortable they were in using the interface, and general comments.

**Interview control** Where an interview was completed the interview status code was set to ‘CQ’ (Completed Questionnaire). If the respondent hung up before the end of the interview the status was set to ‘PQ’ (Partly completed Questionnaire). For the partly completed questionnaires respondents could ring back at a later time to complete the interview, in which case GEIS played an explanatory message and then recommenced at the point previously left off.

### Analysis

The number of words in the voice recordings, duration of voice recordings obtained from GoldWave<sup>203</sup>, mean and median recording duration were tabulated. Equality of voice recording duration were compared between Pretest One and Pretest Two using the Wilcoxon test<sup>204</sup>. The number of calls received was collated using the CONTROL data set. The comments and answers in the assessment form in respect to data entry and acceptability were categorized and tabulated.

#### 4.4.4 Results

Calls were received between the hours 0830–2040 in September 1999.

As shown in Table 4.4, the voice recordings were spoken at a significantly slower rate than in Pretest One (Wilcoxon test,  $W=976$ ,  $P<0.01$ ,  $n=70$ ).

Table 4.4: Number and duration of IVR messages in Pretest One and Pretest Two.

<i>Project</i>	<i>Mode</i>	<i>Items<sup>a</sup></i>	<i>Words<sup>b</sup></i>	<i>Total<sup>c</sup></i>	<i>Mean<sup>d</sup></i>	<i>Median<sup>e</sup></i>
Pretest One	Outbound	34	1223	352.6	10.4	9.8
Pretest Two	Inbound	36	1147	649.8	18	17.9

<sup>a</sup> Number of script items with recorded messages.<sup>b</sup> Number of spoken words.<sup>c</sup> Total duration of all messages (seconds).<sup>d</sup> Mean message duration (seconds).<sup>e</sup> Median message duration (seconds).

### Dialling in to the system

A total of 21 (15%) responses, 19 (13%) completed interviews, and 14 (10%) follow-up assessments were obtained. The interviewing log recorded 17 failed attempts to enter a security code, some of which may have been those who later successfully logged in or other unidentified individuals.

### Accepting security codes

After multiple e-mails requesting all respondents who had not already used the system to do so, no feedback was received that any person was ultimately unable to ring in. One respondent had some initial difficulty entering a security code because their telephone lacked keys labelled with letters, but this was resolved quickly.

### Acceptability

The results of the assessment are shown in Table 4.5. Only one (7%) person reported a problem with too many options, more space between questions, difficulty repeating questions, or unexpected problems, suggesting that these aspects of the interface appeared to be working correctly. Only three (21%) respondents

Table 4.5: Pretest Two assessment results summary.

<i>Question</i>	<i>Yes</i>	<i>%</i>
Were there any questions with too many options so that it was difficult to remember them all?	1	7
Do you think there should have been a greater space between questions?	1	7
Did you have problems in responding to questions by pressing the number keys?	6	43
Did you repeat any questions using the hash key?	1	7
If you repeated questions, did you find the process confusing?	1	7
Did the system do anything unexpected, such as fall silent or hang up suddenly?	1	7
Did you think it would have been helpful if the system could have repeated each keystroke back to you?	3	21

Table 4.6: Pretest Two agreement with feeling comfortable with the system.

<i>Agreement with feeling comfortable with the system<sup>a</sup></i>	<i>Yes</i>	<i>%</i>
Strongly agree	1	8
Agree	8	67
Neutral	2	17
Disagree	1	8
Strongly disagree	0	0

<sup>a</sup> Assessment question: ‘How do you feel about the statement: “I felt comfortable in using the system”?’.

thought GEIS should repeat keystrokes. There were six (43%) respondents who had trouble in responding to questions. As shown in Table 4.6, a total of nine (83%) respondents agreed or strongly agreed with the statement that ‘I felt comfortable in using the system’. Where comments were made on assessment forms, they were categorized and tabulated as shown in Table 4.7. The principle comments were that the voice spoke too slowly and that they tended to enter a response as soon as the appropriate option came up.

Table 4.7: Pretest Two summary of respondents' comments.

<i>Category</i>	<i>Comment</i>	<i>n</i>	<i>%<sup>a</sup></i>
<b>Technical error</b>	Identify respondent better	1	8
	Script error	3	25
	No letters on telephone	2	17
		<b>5</b>	<b>42</b>
<b>Needs better identification of speaker</b>		<b>1</b>	<b>8</b>
<b>Poor voice</b>	Too slow	<b>8</b>	<b>67</b>
<b>Needs feedback</b>	Key press feedback is acceptable	4	33
	Needs more feedback	1	8
		<b>5</b>	<b>42</b>
<b>Too long/slow</b>	Interview too long	1	8
	Slow feedback	1	8
	Too many instructions	4	33
		<b>5</b>	<b>42</b>
<b>Data entry</b>	Pressed a key as soon as my option came up	6	50
	Numeric entry was easier than listening to response options	2	17
		<b>8</b>	<b>67</b>

<sup>a</sup> Categories not mutually exclusive.

#### 4.4.5 Discussion

##### **Dialling in to the system**

No respondents reported being unable to dial into GEIS. The failed attempts to enter a security code in the interviewing log may have been those who later successfully logged in.

##### **Accepting security codes**

One person reported difficulty entering a security code since their telephone lacked keys labelled with letters. This indicated that entry codes for inbound systems should be purely numeric.

##### **Acceptability**

It was noted by several respondents that the pace was too slow. This was probably because the introduction messages were enunciated too slowly. A slower enunciation had been used because it was thought that the Pretest One recordings were spoken too fast. For example, the mean length of the initial explanatory messages in Pretest One was 73.6 seconds long, and the same items in Pretest Two took 128.8 seconds. The slow enunciation throughout Pretest Two resulted in some frustration with the speed of the system, which explained why some respondents commented that they pressed a key as soon as they heard their option come up. It was concluded that following trials should be recorded at a brisker pace, but still slow enough to clearly hear all the options.

Some respondents had trouble in responding to questions because they had telephone keypads without lettered keys (for entering the security code), they were distracted and lost concentration, they expected to be told to press a hash

key, or they forgot to press the hash key. It was concluded that lettered keypads had not proved to be particularly successful and should not be used again, and that instructions for numeric entry questions needed to be clearer.

The majority of respondents felt comfortable with using the system, however only 16% of the possible respondents rang in to the system. Apart from the slowness, the interface appeared acceptable.

#### **4.4.6 Conclusions of Pretest Two**

The pretest's aims were met but some changes were required to the system:

1. Since respondents were able to dial into the system no changes were made to this aspect;
2. Respondents successfully entered a security code, but it was concluded that numeric codes were essential; and
3. The method was acceptable to respondents, but the message recording should be made at a faster pace.

### **4.5 Pilot One: A community pilot of the inbound and outbound methods**

#### **4.5.1 Introduction**

The pretests confirmed that respondents familiar with data entry methods were able to operate GEIS, set call backs, navigate the questionnaire, and enter complex numeric data by means of touchphone keypads. They also confirmed that GEIS could receive incoming calls and schedule outgoing calls. However, since



the pretests had been done with HCHA staff they gave no indication of the response rates that might be obtained with a community sample, nor its acceptability to respondents in that setting. A pilot was therefore carried out to establish whether GEIS could be used for a household survey.

### 4.5.2 Pilot One aims

The pilot was intended to test the ease and acceptability of the IVR interface within GEIS for outbound and inbound calls with a community household sample. The specific aims were to assess:

1. The system's capacity to identify the call outcome (respondents, disconnected lines, answering machines, and fax machines);
2. Whether household respondents could successfully respond to single-choice questions and numeric entry questions; and
3. The acceptability of the method to respondents in a general community household setting.

### 4.5.3 Methods

#### Design

A survey of 50 households was conducted using the IVR method operating in inbound and outbound modes. A follow-up manual telephone interview assessed the acceptability of the methods.

### Sample

A sampling frame was constructed from an electronic version of the telephone White Pages®<sup>205</sup> using the program shown in Appendix Section B.3.1. Mobile telephones were removed from the frame by the sampling program because the project was concerned with households rather than individuals. Additionally, data entry using a mobile telephone may have been more difficult than with a fixed telephone. The program also deleted entries with incomplete addresses to minimize the return of undelivered letters.

### Setting

Only households within Statistical Local Areas (SLAs) located 10 kilometres from Wallsend were selected in order to minimize call costs. A Statistical Local Area (SLA) is one of the spatial units defined within the Australian Standard Geographical Classification (ASGC) published by the Australian Bureau of Statistics (ABS)<sup>206</sup>. The ASGC is used for the collection and dissemination of geographically classified statistics. The ASGC has five hierarchical levels, comprising in ascending order: Census Collection Districts (CDs), SLAs, Statistical Subdivisions (SSDs), Statistical Divisions (SDs), and States and Territories (S/Ts). The CDs aggregate to form SLAs, SLAs aggregate to form SSDs, and so on, to collectively cover all of Australia without gaps or overlaps. While CDs are designed for the collection and dissemination of census data, SLAs are based on the boundaries of incorporated bodies of local government known as Local Government Areas (LGAs). A Local Government Area (LGA) is either geographically identical to a SLA or it may be split into two or more SLAs.

Distances from Wallsend were calculated using Equation 3.1 in Section 3.4.7

of Chapter 3 using latitude and longitude of place names obtained from the Australian Surveying & Land Information Group<sup>207</sup>. The program is shown in Appendix Section B.3.2. The households that were selected were from the SLAs of Lake Macquarie, Maitland, and Newcastle-Remainder. In 1996, in these SLAs there were 126,171 male and 134,884 female adults housed within 141,315 households<sup>208</sup>.

### **Procedure**

Approval was obtained for this study prior to commencement from the University of Newcastle Human Research Ethics Committee (HREC) and Hunter Area Research Ethics Committee (HAREC).

**Modes** The IVR method was used in two modes: outbound dialing in which GEIS contacted households and conducted an interview, and inbound dialing in which GEIS waited for respondents to dial in.

**Respondent instruction** Information letters were printed on letterhead stationery with the HCHA and The University of Newcastle listed as the sponsoring institutions. The letters included accompanying instruction sheets and were addressed to ‘The Household’ (see Figures B.6, B.7, and B.8 of Appendix Section B.3.3). They were sent ten working days before the first outbound call attempt. The instruction sheet indicated that pressing the star key would return to the previous question and pressing the hash key would repeat the current question.

**Call scheduling** Outbound calls were restricted to the hours 1000–1800. At intervals of one minute GEIS checked if there were any calls due using the call-

scheduling algorithm described in Section 3.4.6 in Chapter 3. The inbound method was started on the same day as the letter mail out in case some letters arrived very quickly.

**Respondent recruitment** The last birthday method was used to minimize within-household selection bias<sup>209</sup>. In this method, the person with the most recent birthday was invited to participate. This was to avoid bias introduced by self-selection. If the eligible respondent was unavailable or the current time was inconvenient GEIS offered to make an appointment to ring back.

Answering machines were distinguished from genuine individuals by the length of the salutation as described in Section 3.4.2 of Chapter 3. Salutations lasting longer than three seconds were presumed to be answering machines and in these cases a short message was played.

**Respondent identification** On answering a call, GEIS played an introductory message and then asked the person to press key ‘5’. If this was not detected within 20 seconds it played a message thanking the respondent and stating that he or she would not be able to take part, and the call then terminated. This was done to allow for the possibility that some telephones were either not touch-phones or were not set to tone dialing. If a key press was detected GEIS then described the survey purpose before asking if the respondent was over 18 years old and the member of his or her household to have had the most recent birthday. GEIS then explained it could call back later if the current time was not convenient or the eligible respondent was unavailable. By default, GEIS rang back the next day at the same time, but the respondent could change this to an alternative date and time ranging from the current date and time to thirty

days in the future.

### **Apparatus**

The hardware was set up as in the pretests except that one line was used for outbound calls and two lines were used to receive inbound calls.

### **Measures**

**Interview script** The interview scripts for the outbound and inbound modes are shown in Supplementary Materials Sections 4.3.1 and 4.3.2, respectively. The scripts were the same except for the introductory sections. The interview domains contained items on alcohol, marijuana, amphetamines, heroin, technology use, and demography.

**Voice recordings** The questions were recorded by a female staff member and were spoken at a brisker pace for Pilot One than in Pretest Two, but slower than Pretest One. This was done as a result of the feedback given by the pretest respondents.

**Assessment form** Attempts were made to re-contact all cases using a manual telephone interview. A number of questionnaires were prepared to deal with the possible outcomes of the initial interviews. Paper assessment sheets were constructed for those who completed the interview, failed to ring in to the inbound method, or who hung up on the system. These are shown in Supplementary Materials Section 4.3.3.

**Interview control** Where an interview was completed the interview status code was set to ‘CQ’ (Completed Questionnaire). Any cases that were scheduled

by the respondent to be completed at a later date were coded as ‘CB’ (Call Back). If the respondent hung up after starting the interview the status was set to ‘PQ’ (Partly-completed Questionnaire). For the partly completed interviews an automatic call back was set for the same time the next day. When partly completed interviews were recommenced GEIS played an explanatory message and then started from the point left off.

### Analysis

The number of words in the voice recordings, duration of voice recordings obtained from GoldWave software<sup>203</sup>, mean and median recording duration were tabulated. Equality of voice recording duration were compared between Pretest One, Pretest Two, and Pilot One using the Wilcoxon test<sup>204</sup>. The number of calls received was collated from the CONTROL data set. The comments and answers in the assessment form in respect to data entry and method acceptability were categorized and then tabulated. The type of telephone service as identified by the IVR method was tabulated against the type identified in the follow-up interview.

Exact methods were used to calculate confidence intervals for some variables because of small sample sizes. This was done with SAS/STAT software<sup>185</sup>, which uses the F distribution method described by Leemis and Trivedi<sup>210</sup>.

#### 4.5.4 Results

Call attempts were made by the outbound method between the hours 1038–1705. Both the outbound and inbound methods were conducted in November–December 1999.

Table 4.8: Number and duration of IVR messages in Pretest One, Pretest Two, and Pilot One.

<i>Project</i>	<i>Mode</i>	<i>Items</i> <sup>a</sup>	<i>Words</i> <sup>b</sup>	<i>Total</i> <sup>c</sup>	<i>Mean</i> <sup>d</sup>	<i>Median</i> <sup>e</sup>
Pretest One	Outbound	34	1223	352.6	10.4	9.8
Pretest Two	Inbound	36	1147	649.8	18.0	17.9
Pilot One	Outbound	88	3101	1197.8	15.1	12.4
Pilot One	Inbound	75	2707	1069.3	14.4	12.3

<sup>a</sup> Number of script items with recorded messages.<sup>b</sup> Number of spoken words.<sup>c</sup> Total duration of all messages (seconds).<sup>d</sup> Mean message duration (seconds).<sup>e</sup> Median message duration (seconds).

As shown in Table 4.8, the duration of the voice recordings for Pilot One were intermediate between those of Pretest One and Pretest Two.

### Call outcome

The outcomes of call attempts are shown in Table 4.9. Only two interviews were completed by the outbound method, which took 11 min 36 s and 17 min 21 s to complete. For the outbound method, the mean number of calls required to contact respondents was 2.2 and the mean number to complete interviews was 5.0.

Five individuals explicitly refused, three were classified as answering machines, while ten others that were recorded by the system as out-of-scope had either hung up or had the wrong type of telephone. The remaining numbers were classified by the system as disconnected or fax machines. Only two attempts were made to ring into the inbound method, and of these, only one resulted in a completed interview lasting 9 min 22 s. The incomplete interview was abandoned before any substantive questions were asked.

Therefore, the outbound method obtained the response rate of 9% (Exact

Table 4.9: Pilot One interview outcomes by mode.

<i>Interview Outcome</i>	<i>Outbound</i>		<i>Inbound</i>	
	<i>Frequency</i>	<i>%</i>	<i>Frequency</i>	<i>%</i>
Answering machine	3	13		
Completed	2	8	1	5
Partly-completed	0	0	1	5
Refused	5	21		
Disconnected	2	8		
Fax	2	8		
Out-of-scope	10	42		
Not attempted			17	90
Total	24	100	19	100
Letter returned	1		4	

95% CI: {1%, 28%}), and the inbound method obtained 5% (Exact 95% CI: {0%, 26%}).

The outcomes of the follow-up interviews were cross-classified with the initial interview status, as shown in Table 4.10.

In the case of the outbound method, the three cases initially identified as answering machines (Status code ‘AM’) were actually disconnected lines. The misclassification was the result of a Telstra automatic recorded message that is sometimes played on disconnected lines. For example,

“The number you have dialed is not connected...”

The two cases initially classified as disconnected lines were correct. One of the out-of-scope interviews and one of the refused interviews also resulted from disconnected lines. In these cases the system appeared to have misidentified the voices as real people. Another out-of-scope case appeared to have been entered by a non-English speaker. Apart from three cases that could not be contacted in the follow-up survey, the remaining cases were associated with numbers in active use. In the inbound method, one number was a disconnected line, one



Table 4.10: Pilot One IVR and follow-up interview outcomes.

<i>IVR<sup>a</sup></i>	<i>Follow-up Interview Outcome<sup>b</sup></i>									<i>Total</i>
	<i>AM</i>	<i>CB</i>	<i>CQ</i>	<i>DR</i>	<i>DT</i>	<i>FM</i>	<i>PQ</i>	<i>RT</i>	<i>NE</i>	
<i>Outbound</i>										
AM	0	0	0	0	3	0	0	0	0	3
CQ	0	0	1	0	0	0	0	1	0	2
DT	0	0	0	0	2	0	0	0	0	2
DR	0	2	0	0	1	0	1	1	0	5
FM	0	0	0	0	0	2	0	0	0	2
OS	0	1	3	2	1	0	1	1	1	10
<i>Total</i>	0	3	4	2	7	2	2	3	1	24
<i>Inbound</i>										
CQ	0	0	1	0	0	0	0	0	0	1
NA	1	1	4	5	1	1	1	3	1	18
<i>Total</i>	1	1	5	5	1	1	1	3	1	19

<sup>a</sup> Outcome of the IVR interview: AM=Answering machine, CQ=Completed interview, DT=Disconnected line, DR=Refused, FM=Fax machine, OS=Out-of-scope, NA=No interview attempted.

<sup>b</sup> GEIS Status code for the followup interview: AM=Answering machine, CB=Call-back, CQ=Completed, DR=Refused, DT=Disconnected, FM=Fax machine, PQ=Partly-completed Questionnaire, RT=Ring tone, NE=Non-English speaker.

was a fax machine, and the remainder were numbers in active use.

Of those followed up who did not successfully complete the outbound call, all reported that they did not remember the call. There were insufficient completed outbound calls to address questions of comfort in using the system, preference for a human interviewer, or wanting to do the interview in the future. Four cases assigned to the inbound mode, but who did not ring in, were asked why no calls had been made. None could report any reason for this.

### Data entry

There was no evidence of any difficulty in entering data as evidenced by the data collected during interviews. All respondents also recorded their occupations successfully using the recording feature and the recordings were clear and easily

Table 4.11: Pilot One acceptability by mode and IVR interview outcome.

<i>Qtn.</i>	<i>Mode</i>	<i>IVR Interview</i>	<i>SA</i> <sup>a</sup>	<i>A</i>	<i>N</i>	<i>D</i>	<i>SD</i>
			%	%	%	%	%
Comfort <sup>b</sup> (n=9)	Inbound	Not done	0	20	20	0	60
		Done	100	0	0	0	0
	Outbound	Not done	25	0	25	25	25
		Done	0	0	100	0	0
Preference <sup>c</sup> (n=9)	Inbound	Not done	40	40	20	0	0
		Done	0	100	0	0	0
	Outbound	Not done	25	25	0	50	0
		Done	0	0	100	0	0

<sup>a</sup> SA=Strongly agree, A=Agree, N=Neutral, D=Disagree, SD=Strongly disagree.

<sup>b</sup> "I felt/would feel comfortable using the system."

<sup>c</sup> "I would prefer/would have preferred a human interviewer."

understood.

### Acceptability

Those who did the interview were asked if they felt comfortable using the system and whether they would have preferred a human interviewer, and those who did not were asked if they would have felt comfortable using the system, and whether they would prefer a human interviewer (see Table 4.11). The range of responses was very wide. The sample size was too small to draw any conclusions, but it was of interest to note that the majority reported preferring a human interviewer. When asked if they had any general comments about the call, seven (58%) provided answers, as shown in Figure 4.4.

**Not contacted:**

“[Those with a] higher income would prefer to do a study on a computer because they don’t need to be led as much and prefer the end result and the lower income bracket need human interviewers and need to be led.”

“[I’ve] no tolerance for computers.”

**Contacted:**

“It was generally OK.”

“[It was] OK. [As] long as I won’t have to do it all the time.”

“Prefer humans”

“Straight to the point. Quite easy. Quite comfortable. Well put together.”

“[I’m] too old for computers.”

Figure 4.4: Pilot One general verbatim comments by IVR call outcome.

#### 4.5.5 Discussion

**Call outcome**

Apart from one case where someone had dialled in but did not proceed with the interview, all of those who attempted the interview completed it.

The proportion of usable touchphones was satisfactorily high (83%; exact 95% CI: {52%, 98%}). The misclassification of some disconnected lines as being out-of-scope or answering machines was a minor but not insignificant problem. These cases could potentially result in the call scheduling system continuing to call the cases identified as answering machines until the calling limit was reached, which would waste time and money, and may cause a nuisance. GEIS may also misclassify a respondent as an answering machine. Circumventing this problem could have been done by asking any respondent listening to press a key.

**Data entry**

There was no evidence that any of the respondents had difficulty entering data. This suggested that the design of the IVR mode in GEIS was satisfactory.

**Acceptability**

Due to the small number of completed cases it was difficult to determine the acceptability. However, despite the range of responses, there was no evidence that the system was notably unacceptable.

**4.5.6 Conclusions of Pilot One**

The aims of the pilot had been met:

1. The system determined the outcomes of the calls satisfactorily;
2. Respondents could enter their answers without difficulty; and
3. The method was acceptable to respondents.

However, the low response rate of the outbound method (9%) was a concern to the project team. The low response rate was consistent with previous work<sup>8,22,23</sup>, but, if valid, would not be sufficient for outbound community surveys. Therefore, an alternative method of interviewing was devised, which is described in the next section.

## 4.6 Pilot Two: A community pilot of the Hybrid method

### 4.6.1 Introduction

It was felt by the project team that the very low response rate obtained by the IVR outbound mode in Pilot One would not be sufficient to conduct a telephone household survey using this methodology. Therefore an alternative approach was needed to retain the advantages of IVR, such as economy and automation, but which might obtain a response rate similar to CATI.

To achieve this, an interviewing method was devised in which the telephone call was initiated by an interviewer, but in which the interview itself was handled by the IVR method. Since the interview method involved both a CATI and an IVR approach, it was called the Hybrid method.

The role of the interviewer in the Hybrid method was to contact respondents and persuade them to cooperate. Once this was achieved, it was hoped that most respondents would agree to the IVR interview. This was thought likely because high response rates can be achieved with CATI interviews<sup>52</sup>. The response rate of a CATI survey depends on a variety of factors including the topic of survey, interview duration, the organization conducting the survey, contact procedures, number of call attempts, and duration of the interviewing period<sup>52</sup>. Responses of around 70%<sup>211</sup> and as high as 80%<sup>52</sup>, may be expected. The response rate expected with IVR may be lower because respondents may find it easier to hang up on a call from a machine than a person<sup>8</sup>. However, it was thought this may not occur if the respondent had initially consented to the interview at the request of a human interviewer.

### 4.6.2 Pilot Two aims

In the Hybrid method the respondent would be transferred by the CATI interviewer to the Recorded Voice System (RVS) (The RVS is defined in Section 3.4.2 of Chapter 3) and the interview would then be conducted with the IVR method. These processes needed investigation.

The specific aims for the pilot were to assess:

1. The feasibility of call transfer within the Hybrid method;
2. Whether household respondents would answer IVR questions after the call was transferred; and
3. The acceptability of the method to respondents in a general community household setting.

### 4.6.3 Methods

The Hybrid method is described in detail under the heading ‘Hybrid method’ within this section (page 120) and illustrated in Figure 4.5. Briefly, in the Hybrid method, the interviewer initiated the telephone call, transferred the call to a computer running GEIS (the RVS) in IVR mode, which asked the alcohol and drug items, and then GEIS automatically transferred the call back to the interviewer.

#### Design

A cross-sectional survey of 26 New South Wales (NSW) households was conducted using the Hybrid method. The acceptability of the IVR methodology was assessed following the IVR interview by the interviewer using a CATI script.

Table 4.12: ARIA categories, scores, and descriptors.

<i>Category(Scores)</i>	<i>Descriptor</i>
Highly accessible ( $\leq 1.84$ )	Relatively unrestricted accessibility to a wide range of goods and services and opportunities for social interaction.
Accessible (1.85–3.51)	Some restrictions to accessibility of some goods, services and opportunities for social interaction.
Moderately accessible (3.52–5.80)	Significantly restricted accessibility of goods, services and opportunities for social interaction.
Remote (5.81–9.08)	Very restricted accessibility of goods, services and opportunities for social interaction.
Very remote (9.09–12.00)	Locationally disadvantaged, very little accessibility of goods, services and opportunities for social interaction.

### Sample

The sample was drawn from across NSW. Only those households classified by the Accessibility/Remoteness Index of Australia (ARIA) coding scheme<sup>212</sup> as being within highly accessible areas were included. The ARIA scheme is published by the Department of Health and Aged Care and the National Key Centre for Social Applications of Geographical Information Systems<sup>212</sup>. ARIA measures remoteness of localities in terms of access along a road network to service centres. Service centres are Australian Bureau of Statistics (ABS)-defined urban centres with a population of 5000 or more as at the 1996 Census. Localities that are more remote have less access to service centres; those that are less remote have greater access to service centres. More remote localities have higher ARIA scores than less remote localities. The ARIA scores can be categorized, as shown in Table 4.12.

Households classified as being in highly accessible areas would have a higher likelihood of having a touchphone rather than a rotary telephone. This was

desirable since an aim of the pilot was to test the practicality of the Hybrid method and touchphones were essential with this method.

As in Pilot One, the sample was drawn from an electronic version of the telephone White Pages<sup>205</sup> by the sampling program (see Appendix Section B.4.1). All mobile telephone numbers and entries with incomplete addresses were removed from the sampling frame by the program. The program then merged ARIA codes<sup>212</sup> with the frame, using the postcode as a key, and only retained those postcodes coded by the ARIA system as ‘highly accessible’.

The sample size was considered sufficient to provide an indication of any practical problems that may arise with the method. It was also desirable to test GEIS’s capacity to dial Subscriber Trunk Dialling (STD) numbers, and so the households were selected from across NSW rather than locally as was done in Pilot One.

### Setting

In 1996, there were in NSW 2,175,855 males and 2,284,144 females aged 18 and above housed within 2,387,825 households<sup>208</sup>.

### Procedure

Ethics approval was obtained for the Pilot study from the HREC and the HAREC.

**Hybrid method** The Hybrid method is illustrated in Figure 4.5. The interviewer initiated the telephone call, obtained the respondent’s consent to the interview, and then transferred the call to the RVS that was running GEIS in the inbound mode. GEIS asked the alcohol and drug questions within an IVR



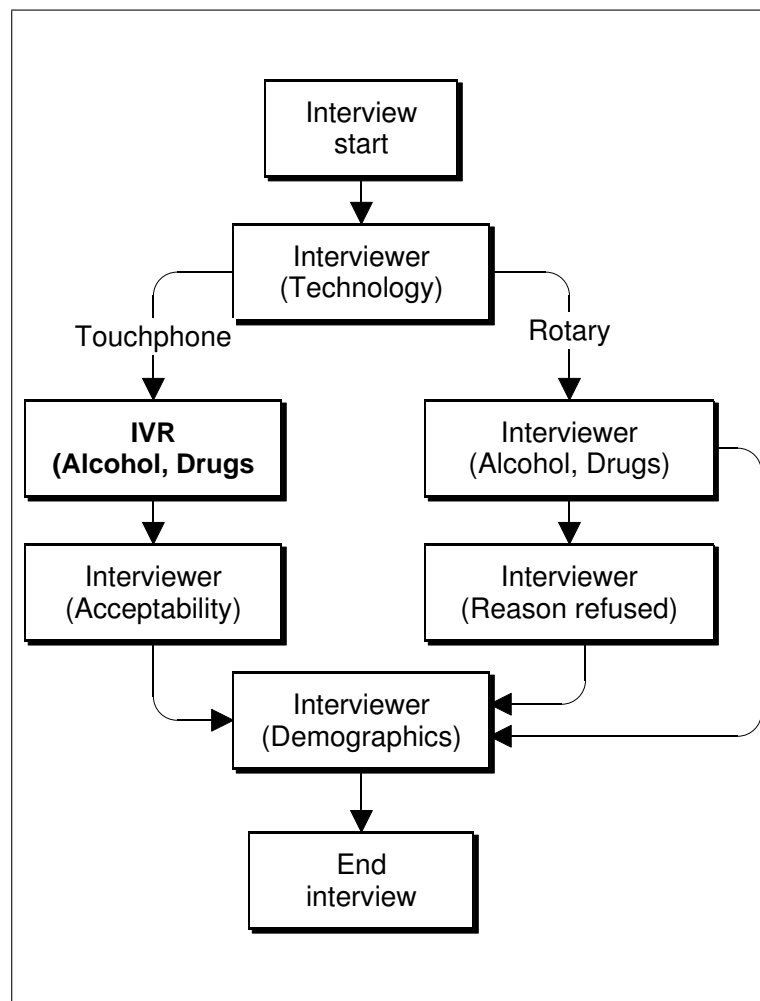


Figure 4.5: Pilot Two domains handled by the interviewer and GEIS running in IVR mode.

interview. At the conclusion of the IVR interview the respondent was automatically transferred back to the interviewer to enable the completion of the demographic items.

Before transferring the call to the RVS, the interviewer asked what sort of telephone the respondent was using. If it was a touchphone the interviewer asked the respondent to press a key. If the interviewer could hear a tone it could be safely assumed that the telephone was a correctly configured touchphone. In

cases where the respondent did not have a suitably configured touchphone, or where the respondent refused to do an IVR interview, the alcohol and drug items were asked by the interviewer.

**Respondent instruction** Information letters were printed on letterhead stationery with the HCHA and The University of Newcastle listed as the sponsoring institutions. Information letters and instruction sheets were posted ten days before the first outbound call attempt. The information letter (see Figure B.9 of Appendix Section B.4.2) was based on that used in Pilot One, but was shortened to eliminate superfluous information and simplify the language used. The instruction sheet (see Figure B.10 of Appendix Section B.4.2) was also based on that used in Pilot One, but was slightly redesigned to use simpler English.

**Respondent recruitment and identification** The Last Birthday method was used to minimize within-household selection bias<sup>209</sup>. If the eligible respondent was unavailable or the current time was inconvenient, the interviewer offered to make an appointment to ring back.

### **Apparatus**

The hardware set up described in Appendix Section A.2 was used except that only a single CATI interviewing station was used and a single telephone line was used to transfer calls to the RVS. Both the interviewer's handset and the RVS were connected to a Private Automatic Branch Exchange (PABX) (Ericsson, model MD110). The interviewer used a telephone handset (Interquartz, model IQ450) with programmable keys. One of the keys was programmed with the keystrokes that would put the call on hold, dial the extension of the RVS, and

then hang up.

The Hybrid method was implemented using two communicating systems: a CATI system on one computer and the RVS with GEIS running in inbound mode on another computer. Call transfers between the two systems were coordinated using the AUTOCALL data set (see Appendix Section A.5.14). The technical aspects of the use of the AUTOCALL data set within the Hybrid method are given in Appendix Section B.4.3.

### Measures

**Interview script** The script contained the same questions as Pilot One but divided into two files. The first file contained the interviewer's CATI script and the second file contained the IVR script, as shown in Supplementary Materials Sections 4.4.1 and 4.4.2, respectively. The CATI interview contained the technology items, touchphone test items, call transfer items, post-IVR interview acceptability assessment items and demographic items. It also contained the drug and alcohol items for those who did not do the IVR interview. The IVR script contained introductory items, alcohol and drug items, and items to transfer the call back to the interviewer.

**Voice recordings** The same voice recordings that had been used in Pilot One were reused where the questions were the same. However, a small number of new recordings that were specific to the Hybrid method were made by the same speaker as in Pilot One. These were principally introductory and terminating messages.

**Assessment form** Following the IVR interview, the call was transferred back to the interviewer. To identify any difficulties that may have occurred during the IVR interview, the interviewer asked the question

“Did you have any problems?”

Depending on comments made by the respondent, the interviewer recorded whether the IVR interview appeared to have been at least partly done, or not at all. This was done to identify cases where technical failures may have prevented the IVR interview occurring at all. If at least part of the IVR interview had been done, the interviewer asked a series of acceptability questions:

“Thinking about the recorded voice interview, do you think it was hard to do?”

“Do you think the interview was stressful?”

“Did you strongly like it, like it, neither like nor dislike it, dislike it, strongly dislike it?”

“When you were asked about drugs and alcohol, would you have preferred a human interviewer rather than a recorded voice?”

**Interview control** A single interviewer was used. The interviewer rang each case in turn, but there was insufficient time for multiple calls. This was considered acceptable since the aims of the pilot concerned technical difficulties and not response rates.

## Analysis

The number of words in the voice recordings, duration of voice recordings obtained from GoldWave software<sup>203</sup>, mean recording duration and median record-

Table 4.13: Pilot Two CATI interview outcomes.

<i>Interview outcome</i>	<i>Frequency</i>	<i>%</i>
Answering machine	1	4
Completed	9	38
Refused	5	21
Disconnected line	3	13
Partly-completed	3	13
Ring tone	3	13
Total	24	100
Letter returned	2	

ing duration were tabulated. Equality of voice recording duration were compared between Pilot One and Pilot Two using the Wilcoxon test<sup>204</sup>. The outcomes of the CATI interviews and the IVR interviews were collated using the respective CATI and IVR CONTROL data sets. The response rate was calculated as the proportion of completed interviews over all eligible cases. The assessment items from the CATI interview and demographic items were tabulated.

#### 4.6.4 Results

Calls were made between the hours 1711–2002 in April 2000. Of the 24 attempted interviews, 12 (50%) resulted in complete or partly completed interviews and five (21%) were refused (Table 4.13). Of the 12 complete or partly completed CATI interviews, three (13%) partly completed the CATI questions, but all twelve completed the IVR questions. One of these three resulted from an interviewer error terminating the call, one from a script error, and one from the respondent suspending the interview. The response rate was 39% (95% CI: {20%, 62%}).

The duration of IVR messages in Pilot Two (see Table 4.14) did not differ

Table 4.14: Number and duration of IVR messages in Pretest One, Pretest Two, Pilot One, and Pilot Two.

<i>Project</i>	<i>Mode</i>	<i>Items<sup>a</sup></i>	<i>Words<sup>b</sup></i>	<i>Total<sup>c</sup></i>	<i>Mean<sup>d</sup></i>	<i>Median<sup>e</sup></i>
Pretest One	Outbound	34	1223	352.6	10.4	9.8
Pretest Two	Inbound	36	1147	649.8	18.0	17.9
Pilot One	Inbound	75	2707	1069.3	14.4	12.3
Pilot One	Outbound	88	3101	1197.8	15.1	12.4
Pilot Two	Hybrid	37	1175	437.7	12.7	12.3

<sup>a</sup> Number of script items with recorded messages.

<sup>b</sup> Number of spoken words.

<sup>c</sup> Total duration of all messages (seconds).

<sup>d</sup> Mean message duration (seconds).

<sup>e</sup> Median message duration (seconds).

significantly from the duration of IVR messages in the outbound arm of Pilot One (Wilcoxon test,  $W=1939$ ,  $P=0.41$ ,  $n=Missing$ ).

### Call transfer

Since all twelve respondents who agreed to an interview had properly configured touchphones they were all asked to attempt the IVR interview. All twelve completed the IVR interview but three were subsequently interrupted after transfer back to the CATI interviewer.

### IVR interview completion

All relevant alcohol and drug items (between 9 and 22 items depending on the logic path taken) in the IVR interview were completed for each respondent in a mean time of 3.1 minutes (95% CI: {2.0, 4.2}). For the completed interviews, all relevant items (between 17 and 38 items) were completed. The mean time for the CATI interview including the IVR interview was 12.5 minutes (95% CI: {10.3, 14.7}).

Table 4.15: Pilot Two IVR interview acceptability questions.

<i>Qtn.</i>	<i>Level</i>	<i>n</i>	<i>%</i>
Hard to do? <sup>a</sup>	<i>Yes</i>	1	10
	<i>No</i>	9	90
Enjoyable? <sup>b</sup>	<i>Yes</i>	4	40
	<i>Don't know</i>	4	40
	<i>No</i>	2	20
Stressful? <sup>c</sup>	<i>Yes</i>	1	10
	<i>No</i>	9	90
Like it? <sup>d</sup>	<i>Strongly like it</i>	1	10
	<i>Like it</i>	2	20
	<i>Neither like nor dislike it</i>	7	70
	<i>Dislike it</i>	0	0
	<i>Strongly dislike it</i>	0	0
Prefer interviewer? <sup>e</sup>	<i>Yes</i>	6	60
	<i>Don't know</i>	1	10
	<i>No</i>	3	30

<sup>a</sup> Do you think the recorded voice interview was hard to do?<sup>b</sup> Do you think the interview was enjoyable?<sup>c</sup> Do you think the interview was stressful?<sup>d</sup> Did you: strongly like it, like it, neither like nor dislike it, dislike it, strongly dislike it?<sup>e</sup> When you were asked about drugs and alcohol, would you have preferred a human interviewer rather than a recorded voice?

### Acceptability

One of the partly completed CATI interviews completed the acceptability questions giving a total of ten responses. The results of the acceptability questions are shown in Table 4.15. The majority of respondents thought the IVR interview was not hard to do and not stressful, while 40% thought it was enjoyable, and 30% liked it. As in Pretest One, most respondents reported that they would have preferred a human interviewer to ask them the drug and alcohol questions.

There were nine responses to the demographic questions. The respondents' demographic characteristics are shown in Table 4.16. Although the sample was too small to draw any statistical conclusions, it appeared that there were no

Table 4.16: Pilot Two demographic characteristics of follow-up cases.

<i>Variable</i>	<i>Level</i>	<i>n</i>	<i>%</i>
Sex	<i>Male</i>	5	56
	<i>Female</i>	4	44
Age	<i>&lt;30</i>	3	33
	<i>30–49</i>	2	22
	<i>50–69</i>	1	11
	<i>≥70</i>	3	33
Education	<i>Primary</i>	1	11
	<i>Secondary</i>	5	56
	<i>Higher</i>	3	33
Marital status	<i>Married/de facto</i>	2	22
	<i>Divorced/separated</i>	1	11
	<i>Widowed</i>	4	44
	<i>Never married</i>	2	22
Country of birth	<i>Australia</i>	8	89
	<i>Outside Australia</i>	1	11
Employment status	<i>Employed</i>	3	33
	<i>Not employed</i>	6	67
Personal income	<i>&lt;\$12,000</i>	1	11
	<i>\$12,000–20,000</i>	3	33
	<i>\$20,000–30,000</i>	1	11
	<i>\$30,000–40,000</i>	1	11
	<i>\$40,000–50,000</i>	1	11
	<i>≥\$50,000</i>	2	22

obvious demographic anomalies.

#### 4.6.5 Discussion

##### Call transfer

Although the response rates to the Hybrid interview appeared low, with multiple calls the response rate could be expected to increase. All respondents agreed to the IVR interview, suggesting that the drop out rate at this point was small. Further, all respondents who completed the IVR interview were successfully returned to the interviewer.

This suggested that respondents might generally dislike surveys, but not IVR



interviews in particular. If there was no human contact at the beginning they may find it easier to terminate an IVR interview than a CATI interview. The respondents were told that they would be rung back if accidentally disconnected, which might also have contributed to the IVR interview completion rate.

### **IVR interview completion**

Despite the complexities involved in the Hybrid method, no technical difficulties were encountered. The interviewer did not report any problems with explaining the process to the respondents or in the call transfer process. The respondents also completed their IVR interviews without difficulty.

### **Acceptability**

From the results of the acceptability questions it was concluded that the IVR interview was broadly acceptable.

## **4.6.6 Conclusions of Pilot Two**

The aims of the pilot had been met:

1. Call transfer had been accomplished without difficulty, indicating that the technical challenges had been met;
2. Since all items in the IVR were completed by those who were transferred, it could be concluded that respondents could complete interviews of this type without difficulty; and
3. It was concluded that the Hybrid method was acceptable to respondents.

## 4.7 Conclusions

The capacity of GEIS to conduct outbound and inbound dialing with a community sample was tested. When outbound calls were made, GEIS was able to differentiate respondents from answering machines, disconnected lines, and fax machines reasonably well. Respondents were able to arrange for call-backs to be made if the call was received at an inconvenient time.

When used as an inbound system the GEIS responded to incoming calls and accepted security codes, although it was concluded that these should be numeric and short.

Respondents were able to provide answers to single-choice questions and numeric entry questions using the interface, and successfully used the hash key to repeat questions. The acceptability of the methodology to respondents in Pilot One and Pilot Two was mixed, as is to be expected with such small sample sizes. However, it was clear that the IVR and Hybrid methods were not unacceptable.

Of greater concern was the low response rate for the IVR outbound method when used for a community survey. The response rate was too low for household surveys, but the Hybrid method obtained a more reasonable response.

Overall, the system needed to be tested with a larger household sample that would compare CATI to IVR and the new method, Hybrid. It was thought that this would resolve the remaining uncertainties, such as the response rate, and also determine whether the technology could be used as a practical tool in survey and public health research. This is explored in the next chapter.

## Chapter 5

# A comparison of four computer-based telephone interviewing methods\*

### 5.1 Introduction

A sensitive question was previously defined in Section 2.4 in Chapter 2 as one that raises concerns about disapproval or other consequences such as legal sanctions for reporting truthfully. Such questions are subject to a social desirability bias, which is the respondent's tendency to over-report the frequency of performing a socially desirable behaviour, or under-report the frequency of performing a socially undesirable one<sup>65,66</sup>. In particular, questions about alcohol and illicit drug use are likely to be sensitive and associated with a strong social desirability response bias<sup>67</sup>.

While alcohol and drug use are sensitive behaviours, it was concluded in Chapter 2 that they are, nevertheless, most easily gauged through self-report. It was noted that Interactive Voice Response (IVR) was a new and promising

---

\*A paper based on material in this chapter has been accepted for publication in *Behavior Research Methods, Instruments, & Computers*

technique for obtaining self-report. However, issues such as acceptability, cost-effectiveness and feasibility in the general community setting remained to be assessed.

## 5.2 Aims

Terms and concepts used in this section are defined in the Methods (Section 5.3).

The aims of this chapter are to:

1. Compare the interview methods Hybrid I, Hybrid II, IVR and Computer-Assisted Telephone Interview (CATI) on the following dimensions:
  - Outcome rate;
  - Sample demographic profile;
  - Interview duration;
  - Calling effort;
  - Item non-response rate;
  - Cost;
  - Self-report rate of alcohol and drugs; and
  - Acceptability.
2. Compare the interview methods with the National Drug Strategy Household Survey (NDSHS)<sup>213</sup> for males and females in the Australian population with respect to:
  - Alcohol drinking status, quantity and frequency of consumption; and
  - Drug (marijuana, amphetamines, heroin) consumption, frequency of use, and age profile.

## 5.3 Methods

### 5.3.1 Design

An Australia-wide telephone survey of households was conducted in 2000 using four different telephone interview methods, CATI, IVR, Hybrid I and Hybrid II. These methods are described in detail below (Section 5.3.4). A follow-up CATI survey was conducted within the following two weeks of each interview to assess the acceptability of each interview method.

### 5.3.2 Setting

In 1996 the Australian population numbered 17,892,423 housed within 7,175,237 households and was distributed across six states and two territories<sup>208</sup>. In 1998, the fixed telephone coverage was high, ranging from 96.0% to 97.6% between states and territories, except for the Northern Territory which had a relatively low coverage of 91.4%. The proportion of households paying bills or transferring funds by telephone ranged from 27.4% to 44.0% between states and territories, indicating that Australians were increasingly familiar with IVR technology<sup>173</sup>.

### 5.3.3 Sample

A study sample size was determined to allow a difference in the proportions estimated by any two of the methods of 10%, such as the proportions of respondents contacted or the proportions of respondents who completed interviews. Equation 5.1 was used to calculate the sample size needed to detect this difference<sup>214</sup>. Equal sample sizes per method and sampling from an infinite population was assumed. To compare two of the methods, the statistical power was set to

$1 - \beta = 0.8$ , the significance to  $\alpha = 0.05$ , and the nominal proportion assumed for both methods to 50%. The sample size required was calculated to be 387 per method. This sample size was inflated to 690 per method to allow for a nonresponse of 30% and letter return rate of 15%. This was then rounded to 700 per method, except for IVR which was increased to 1000 to allow for the possibility of a low response rate.

$$n = \frac{\left( Z_{\alpha/2} \sqrt{(2) \overline{P}\overline{Q}} - Z_{1-\beta} \sqrt{P_1 Q_1 + P_2 Q_2} \right)^2}{(P_2 - P_1)^2} \quad (5.1)$$

where:  $n$  denotes the required sample size for each method,

$P_i$  denotes the nominal proportion in method  $i$ ,

$$Q_i = 1 - P_i,$$

$$\overline{P} = (P_1 + P_2)/2,$$

$$\overline{Q} = 1 - \overline{P},$$

$Z_{\alpha/2}$  denotes the value cutting off the proportion  $\alpha/2$  in the upper tail of the standard normal distribution curve, and

$Z_{1-\beta}$  denotes the value cutting off the proportion  $1 - \beta$  in the upper tail of the standard normal distribution curve.

A total of 3100 households with fixed telephone connections (i.e. not mobile telephones) were selected using simple random sampling from an electronic version of the telephone White Pages covering all states and territories of Australia<sup>215</sup>. The sampling program is shown in Appendix Section C.2. Cases that were eliminated from the sample, due to return of letters or respondents indicating in advance that they would refuse an interview, were not replaced.

### 5.3.4 Procedure

Approval was obtained for this study from the University of Newcastle Human Research Ethics Committee (HREC) and Hunter Area Research Ethics Committee (HAREC).

#### Interviewer training

Five trained and experienced telephone interviewers attended a five hour briefing session covering survey aims, interviewing standards, script familiarization, software, and practice interviews.

#### Interview methods

The Generalized Electronic Interviewing System (GEIS) software described in Chapter 3 was developed to allow both CATI and IVR methods to be used for conducting interviews within telephone surveys. However, the potentially low response rate identified in Chapter 4 for household surveys using outbound IVR calls resulted in GEIS being further extended to support the Hybrid method. The development and testing of the Hybrid method is described in Chapter 4.

The Hybrid method, as described in Section 4.6 of Chapter 4, involved transferring the respondent to the Recorded Voice System (RVS) where the more sensitive questions were asked (The RVS is defined in Section 3.4.2 of Chapter 3). Afterwards, the respondent was transferred back to the interviewer to complete the remainder of the interview. Although the Hybrid method obtained good response rates there was a concern the method would not be economical. This was because the interviewer needed to wait until the interview was completed on the RVS.

Therefore, a variant method was developed in which the interview would not be transferred back to the interviewer. In this method, the interviewer could contact another respondent whilst the first one was handled by the RVS. In the current chapter the original method is called Hybrid I, and the new method is called Hybrid II.

In this study, four computer-based telephone interview methods were used, called herein: CATI, IVR, Hybrid I, and Hybrid II. These are illustrated in Figure 5.1, which shows the processes involved in each interview method. The same questionnaire was presented within each method, but, as indicated in Figure 5.1, specific domains were asked by either interviewer or by GEIS running in IVR mode on the RVS.

Respondents in the Hybrid methods with unsuitable telephones were automatically re-assigned to the CATI method. To minimize break-offs, interviewers always emphasized that they would immediately ring back if the respondent was inadvertently disconnected from the RVS.

### **Respondent instruction**

Information letters printed on letterhead stationery and addressed to ‘The Household’ were posted using DL-sized envelopes<sup>†</sup> with the institution details (The University of Newcastle, Australia and Hunter Centre for Health Advancement, NSW Health) plainly marked, one week before the first call. Institutional letterhead was used since previous work had shown that university sponsorship can increase response rates in surveys<sup>167</sup>. The letters (shown in Appendix Section C.3) stated that the interview would be voluntary and involve questions

---

<sup>†</sup>Envelope size 220 x 110mm.



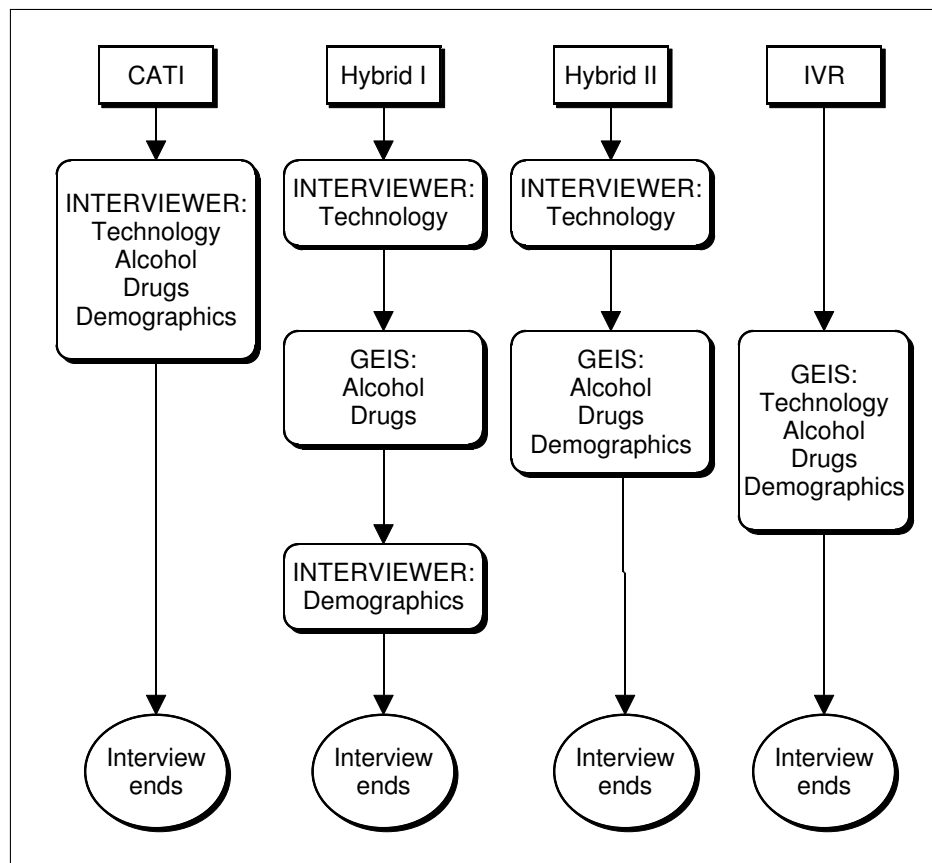


Figure 5.1: Assignment of interviewer or RVS and domains (technology, alcohol, drugs, demography) for each of the methods CATI, Hybrid I, Hybrid II, and IVR.

about alcohol and drug use. The Hybrid and IVR letters clearly explained that a recorded voice would be used. The letters made use of the term ‘recorded voice’ to avoid such words as ‘technology’ or ‘computer’. This was done to encourage those that disliked computers to participate<sup>216</sup>.

No remuneration was offered although previous research has shown that small financial incentives can increase response rates<sup>167</sup>. Incentives were not offered because it was thought that they might have had a differential effect on respondents from varying demographic backgrounds<sup>217</sup>. This effect would then have been confounded with other differences between interview methods.

Simple instruction sheets showing a picture of a typical touchphone keypad were included with the letters to the Hybrid and IVR households (shown in Appendix Figures C.3 and C.5). The instructions indicated that pressing the hash key would repeat a question and pressing the star key would return to the previous question.

### **Respondent recruitment**

Assignment of respondents to each method was unknown to interviewers until an interview had begun. The last birthday method<sup>209</sup> was used to select the respondent. This method attempts to reduce the bias introduced by self-selection by inviting the person with the most recent birthday to participate.

If the eligible person was unavailable interviewers made an appointment to ring back. Interviewers recorded businesses as out-of-scope and did not proceed with the interview.

In the IVR method, GEIS distinguished answering machines from genuine individuals by the length of the salutation. For example:

“Hello?”

is shorter than

“Hello, we’re not in at present...”

The first salutation would be assumed to be a person while the second would be assumed to be an answering machine. If connected, respondents were asked to press key ‘5’ and the system hung up if no key press was detected after three repetitions. GEIS then asked if the telephone number belonged to a business or a residence. Business numbers were recorded as out-of-scope. For households, GEIS asked to speak to the eligible person. If the eligible person was unavailable, it offered to call back at a date and time convenient to the household.

### **Call scheduling**

Higher contact rates in surveys can be obtained by scheduling calls at appropriate times with efficient rescheduling for non-contacts<sup>188,218</sup>. Initial calls were made on weekday afternoons or evenings. CATI Interviewers and GEIS running on the RVS rang back non-contacted numbers at regular intervals. Interviewers made at least seven call attempts and usually called back non-contacts within a few hours, or if there was still no contact, at least a day later. In the IVR method, non-contacted numbers were re-attempted at alternating thirty minute and eighteen hour intervals. There was no calling limit for the IVR method within the limits of the study period.

### 5.3.5 Apparatus

#### Equipment

The interviewing stations and the RVS were Pentium II computers. A four line Dialogic D/41H voice card was installed in the RVS. Interviewers used a mouse or keyboard to enter responses. GEIS running on the RVS played sound files and respondents answered questions by pressing keys on their telephones. Seven telephone lines were used, three for the interviewers and four for the RVS. In the Hybrid methods, interviewers transferred calls to vacant RVS lines. Voice recordings were made in 16-bit mono 11kHz format with an Optimus 33-3104 omnidirectional microphone and Creative Sound Blaster Vibra 128 sound card. Voice recordings were by a single female staff member selected using a voice assessment method based on Oksenberg et al.<sup>53</sup>.

#### Software

All methods were implemented using GEIS as described in Chapter 3. In all methods, answers to questions were provided by selecting one of a set of options, entering a number or date, or entering an open-ended response. Answers to open-ended questions were entered verbatim by interviewers, while GEIS running on the RVS allowed respondents to record a short spoken sentence. For numeric answers, absolute and reasonable limits prevented range errors. Invalid responses triggered an appropriate message or display. After three repetitions of an item, a non-response caused GEIS running on the RVS to hang up and record a refusal. In all interview methods, respondents could refuse to answer a particular item, and if they wished they could return to earlier questions and modify their answers.

### 5.3.6 Measures

#### Scripts

The interview scripts for CATI and Hybrid I are shown in Supplementary Materials Sections 5.1.1 and 5.1.2, respectively. The Hybrid II scripts are shown in Supplementary Materials Sections 5.2.1 and 5.2.2, and the IVR script is shown in Supplementary Materials Section 5.3. The follow-up CATI scripts are shown in Supplementary Materials Sections 5.4.1 and 5.4.2.

#### National Drug Strategy Household Survey

The National Drug Strategy Household Survey (NDSHS)<sup>213,219</sup> was a well known biennial survey conducted by the Commonwealth Department of Health and Family Services to collect information from Australian households on awareness, attitudes, and behaviour relating to alcohol and drug use. The survey used a multi-stage stratified sample design and face-to-face interviews to collect a total of 10,030 interviews. The survey achieved a contact rate of 69% and response rates for three waves of 55%, 61% and 60%. The more sensitive questions were asked using confidential Self-Administered Questionnaires (SAQs). The completed SAQs were sealed in envelopes by the respondents prior to handing them to the interviewer. As explained below, the NDSHS<sup>213</sup> was used in this study to provide a ‘gold standard’.

#### Alcohol Use Disorders Identification Test

The Alcohol Use Disorders Identification Test (AUDIT) (see Appendix Section C.1) was designed for use in primary care settings to identify persons whose alcohol consumption has become hazardous or harmful to their health<sup>194</sup>. The

AUDIT usage guidelines state that it should be used intact and when embedded within a larger document should be preceded by an introductory text<sup>194</sup>. However, in this study the AUDIT questions were inter-mixed with questions from the NDSHS. It was possible that embedding the AUDIT might affect its validity and psychometric properties. However, the AUDIT appears from previous work to be robust against question order changes, modified wording, and question-splitting<sup>220</sup>. Additionally, the AUDIT scores were compared between interviewing methods and not to other studies.

### **Domains**

Questions were grouped into four domains (technology, alcohol, drugs, demography).

The technology domain attempted to assess individuals' acceptance of technology by including items about ownership or use of electronic devices, use of the Internet, previous experience with IVR, and whether the respondent would normally leave messages on answering machines. There was also an item asking if the respondent liked technology. This was scored with a five point Likert scale<sup>158</sup>. The domain items were equivalent to a subset of those used by Rosen and Weil<sup>221</sup> to assess technology use and attitudes among university students.

As discussed in Chapter 2, it was expected that the self-report of the sensitive questions relating to alcohol and drugs would be higher in Hybrid I, Hybrid II, and IVR than CATI. The alcohol domain included consumption questions to be compatible with the National Drug Strategy Household Survey (NDSHS) questionnaire<sup>219</sup> and the five-item Alcohol Use Disorders Identification Test (AUDIT)<sup>193,195</sup>. The drugs domain covered amphetamines, marijuana, and

heroin, which are the more commonly used drugs in Australia<sup>213</sup>. The drug questions were selected to be compatible with the NDSHS<sup>219</sup> questionnaire and covered the age at first consuming, use in the previous twelve months, and frequency of use.

If the self-report rates for Hybrid I, Hybrid II, and IVR were found to be higher than for CATI, this would not have provided an indication of their greater accuracy. A reliable external comparison was needed in order to gauge the external validity<sup>217</sup> of the methods. Intrusive methods such as biochemical testing (see Section 2.5.2 of Chapter 2) were inappropriate since the survey was a general household study rather than a clinical study. Instead, the NDSHS<sup>213</sup> was used to provide a ‘gold standard’ to which the interviewing methods were compared.

The alcohol and drug items were asked in a fixed order (alcohol, marijuana, amphetamines, heroin) so that the least sensitive questions came before the more sensitive questions. This particular order was used to reduce possible break-offs and because earlier work had shown that order effects were of minor importance<sup>66,222</sup>.

It was possible that sample composition might differ between the telephone interview methods. For example, the potentially low response rate for the IVR method could produce a biased sample<sup>62</sup>. Therefore, the sample demographic composition of each of Hybrid I, Hybrid II, and IVR needed to be compared to that of CATI. The demography domain included age, education, marital status, sex, country of birth, and employment status. These questions were selected to be compatible with the Australian Bureau of Statistics (ABS) Census data<sup>208</sup>.

### Follow-up

A random 30% selection of respondents who had completed or partly-completed interviews was selected by GEIS for a follow-up CATI. A single interviewer who was not involved in the main interviews did all the follow-up interviews. In the follow-up interview, respondents rated the previous interview for ease, enjoyableness, stressfulness, and likeableness using a five point Likert scale<sup>158</sup> and standard questions from Bonevski et al.<sup>177</sup>. CATI and Hybrid respondents were asked if they would have preferred a recorded voice to ask about alcohol and drug use, while IVR respondents were asked if they would have preferred an interviewer. All were asked if they thought people would be more honest with an interviewer than a recorded voice.

### 5.3.7 Analysis

Analyses were conducted with SAS/STAT software<sup>185</sup>. Since simple random sampling was used, no adjustment was required to correct for unequal sampling probabilities. Where sample sizes were too low for some statistical tests, significances were estimated using Monte Carlo methods<sup>223</sup>.

### Weighting

While comparisons between the telephone interviewing methods were made using unweighted data, comparisons to ABS or NDSHS<sup>213</sup> data were done with weighted data. The weights<sup>224</sup> were calculated to match the age and sex profile of the 1996 Census<sup>208</sup> using Equation 5.2. The strata in Equation 5.2 consisted of the age groups 15–19, 20–24, ... 50–54, and  $\geq 65$ , separately for males and females. When an analysis only concerned either males or females then the



weights were rescaled to sum to the sample sizes for males or females, respectively.

$$W_h = \frac{N_h/N}{n_h/n} \quad (5.2)$$

where:  $N_h$  denotes the eligible population in stratum  $h$ ,

$N$  denotes the total eligible population of Australia,

$n_h$  denotes the sample size in stratum  $h$ , and

$n$  denotes the total sample size.

### **Sample demographic profile**

Sample demographic statistics (proportions by sex, age, education, marital status, country of birth, and employment status) were compared between methods using contingency table analysis and unweighted counts. The demographic data were compared to the ABS 1996 Census<sup>208</sup> using chi-square goodness-of-fit tests<sup>225,226</sup>.

Ages were compared to quartiles of the Australian population aged eighteen years or older: 18–30, 31–42, 43–57,  $\geq 58$ <sup>208</sup>. If the sample was unbiased with respect to age then each quartile was expected to contain 25% of the total sample. Quartiles were used to describe the sampling age structure. Although a larger number of divisions could have been used, the use of quartiles avoided constructing groups with small sample counts. Over or under-representation in each quartile was readily apparent from the deviations from the expected 25%.

**Outcome rate**

The outcome rates of interest were the contact, cooperation, response, and refusal rates. The contact rate was the proportion of households that each method could successfully contact. In particular, it was thought possible that the contact rate for IVR may be less than the other methods.

The cooperation rate was the proportion of *contacted* households that completed interviews. This rate was expected to give an indication of respondents' amenableness to an interview method once they had been introduced to it.

The response rate was the proportion of all households that completed interviews. For the Hybrid and IVR methods to be usable they needed to obtain comparable response rates to CATI.

The refusal rate was the proportion of households that refused to complete interviews. The refusal rate measured how resistant potential respondents were to a particular method.

Contact, cooperation, response, and refusal rates were calculated according to the American Association for Public Opinion Research (AAPOR) coding system<sup>227</sup>. This is a standard outcome coding system used in survey research. Contact, cooperation, response, and refusal rates were calculated using Equations 5.3, 5.4, 5.5, and 5.6. respectively, after converting GEIS status codes to AAPOR codes. Rates were compared between interview methods using contingency table analyses<sup>226</sup>.

$$\text{Contact rate} = \frac{(I + P) + R + O}{(I + P) + (R + NC + O) + (UH + UO)} \quad (5.3)$$

$$\text{Cooperation rate} = \frac{I}{(I + P) + (R + O)} \quad (5.4)$$

$$\text{Response rate} = \frac{I}{(I + P) + (R + NC + O) + (UH + UO)} \quad (5.5)$$

$$\text{Refusal rate} = \frac{R}{(I + P) + (R + NC + O) + (UH + UO)} \quad (5.6)$$

where:  $I$  = Completed interview (STATUS code: 'CQ'),

$P$  = Partial interview, completed at least one domain (STATUS code: 'PQ'),

$R$  = Refusal and break-off (STATUS codes: 'D3', 'DR', 'DO'),

$NC$  = Non-contact (STATUS codes: 'CB', 'AM'),

$O$  = Other (STATUS codes: 'RD', 'RS'),

$UH$  = Unknown if household (STATUS codes: 'NA', 'ET', 'RT'), and

$UO$  = Unknown eligibility (STATUS codes: 'RV', 'RM').

### Interview duration

It was unclear whether respondents would find Hybrid I, Hybrid II, and IVR to be slower or faster to complete than CATI. A particularly slow method would have been undesirable.

In the Hybrid methods there were two or three intervals in which respondents interacted with a CATI interviewer and GEIS running on the RVS. These were summed to obtain the total interview duration. For example, in Hybrid I a respondent would speak to an interviewer twice and interact with GEIS running on the RVS once. Equivalence of total duration between the Hybrid I and Hybrid II methods was tested using the Kruskal-Wallis test<sup>204</sup>.

### Calling effort

It was unclear whether less or more effort would be required to contact respondents and complete interviews with Hybrid I, Hybrid II, and IVR than CATI.

For example, it was possible that the Hybrid methods may have involved making many more calls than CATI because of respondent resistance to the new methods, in which case a survey using these methods would take longer to complete. The number of calls required to achieve stable survey estimates was used as a measure of calling effort, instead of the response rate (c.f. Steel et al.<sup>52</sup>). This was because the response rates could have stabilized at a low level while the survey estimates remained unstable. The number of calls required to obtain stable estimates was determined to be the point at which the mean values of survey variables came within  $\pm 2.5\%$  of the mean value obtained after fifteen calls.

The number of calls to contact respondents and to complete interviews was compared using Poisson Regression with adjustment for over-dispersion<sup>225</sup>. The method of Signorini<sup>228</sup> was used to calculate the power of the test using the results from Chapter 4. Since the method of Signorini<sup>228</sup> assumed a comparison of two groups the significance level was reduced by Bonferroni adjustments to account for all pairwise comparisons<sup>229</sup>. In Pilot 1, the mean number of calls to contact respondents and complete interviews was 2.2 and 5.0, respectively (Section 4.5.4 of Chapter 4). By assuming an overall significance of 5%, power 80%, and 387 respondents per method, the detectable difference in rate in another method was calculated. Assuming 2.2 calls to contact respondents and 5.0 calls to complete interviews in one method, the detectable differences in the calling rates in the other method were 10% and 2%, respectively.

**Item non-response rate**

It was thought that the item non-response rates for Hybrid I, Hybrid II, and IVR might have been lower than CATI because CATI interviewers might encourage respondents to answer all questions. In this study, the item non-response rate was defined as the number of refused items as a proportion of the eligible questions in a domain. Item non-response rates were compared between methods by contingency table analysis<sup>226</sup>. Only fully completed interviews were used in the calculation to avoid confounding the item non-response rate with the rate of break-offs.

**Cost**

The cost of each method was to be compared. It was expected that the IVR would be cheaper than CATI, but the cost of the other methods was uncertain.

Cost was calculated by summing salary cost and accumulated telephone call charges. Salary cost was calculated by multiplying the interview duration plus non-contact attempt duration by the interviewer's salary rate. The mean interview cost was calculated using the cost of obtaining completed interviews only, and the total survey cost per completed interview was calculated using all costs. Equivalence of cost between interviewing methods was tested using Kruskal-Wallis tests<sup>204</sup>.

**Self-report rate of alcohol and drugs**

Self-reported alcohol and drug use was compared between all methods, and between CATI and the other methods, by contingency table analyses<sup>226</sup>. Median age of first consumption was compared using the Mann-Whitney test<sup>230</sup>. The

proportion of hazardous drinkers, indicated by an AUDIT score greater than five<sup>193</sup>, was compared between methods using contingency table analyses<sup>226</sup>.

### Acceptability

All methods had to be acceptable to respondents for ethical reasons. Acceptability was scored using a Likert scale<sup>158</sup> for ease of use (5-point), enjoyableness (4-point), stressfulness (4-point), and likeableness (5-point). Preference for a human interviewer was investigated by asking a question with ‘*prefer interviewer*’ or ‘*prefer recorded voice*’ as answer options. The respondent’s belief that people would be more or less honest with an IVR system was investigated by asking a question with ‘*person*’ or ‘*recorded voice*’ as options. Levels of Likert scales were not collapsed during analyses.

Equality of acceptability scores (ease of use, enjoyableness, stressfulness, and likeableness) were compared without collapsing levels using Kruskal-Wallis tests<sup>204</sup>. When significant, Wilcoxon tests<sup>204</sup> with Bonferroni adjustments<sup>229</sup> were used to compare CATI scores to the other methods. Median scores were presented in the table after reversing the scales for ease of use and enjoyableness. After reversals, the lowest possible score (1) corresponded to ‘*very easy*’, ‘*very enjoyable*’, ‘*not stressful*’, and ‘*like it a lot*’. The highest possible score (4 or 5) corresponded to ‘*very hard*’, ‘*not enjoyable*’, ‘*very stressful*’, and ‘*dislike it a lot*’.

After ordering the methods according to the increasing degree of automation (CATI, Hybrid I, Hybrid II, IVR), Cochran-Armitage Trend tests with modified ridit scores<sup>231</sup> were used to test for a trend in the proportion of respondents preferring an interviewer, or agreeing with the statement that people are more

likely to be honest about their alcohol use to a person than a recorded voice. This is a test for trends in binomial proportions.

Wilcoxon tests<sup>204</sup> were used to assess the relationship between the acceptability scores (ease of use, enjoyableness, stressfulness, and likeableness) and questions with two response options ('yes', 'no'): technology questions (ownership of mobile telephones, usage of answering machines, previous experience with IVR, home access to the Internet), and attitude questions (preference for an interviewer, or agreeing with the statement that people are more likely to be honest about their alcohol use to a person than a recorded voice). Spearman's correlation<sup>225</sup> was used to assess the relationship between the acceptability scores (ease of use, enjoyableness, stressfulness, and likeableness) and a question with five response options that asked if the respondent liked technology. Scores were not collapsed during computations.

### **Comparison to the National Drug Strategy Household Survey**

Comparisons to the NDSHS<sup>213</sup> results were done using weighted data. This was because the NDSHS results were also weighted to match ABS estimates<sup>219</sup>. To ensure comparability with the NDSHS results, only individuals aged twenty or older were retained in analyses.

To obtain a more powerful test of differences between the interview methods and the NDSHS data, the data from the two Hybrid methods were combined. Before combining data from the two Hybrid methods, tests for homogeneity<sup>225</sup> were conducted for each parameter of comparison, and the data were not combined if the methods differed.

**Alcohol drinking status, quantity, and frequency** The NDSHS profile for drinking status (*regular, occasional, ex-drinker, never drank*), drinking consumption (*1–2, 3–4, 5–6,  $\geq 7$  drinks per day*), drinking frequency (*every day, 4–6 days/week, 2–3 days/week, 1 day/week, less often*) were compared by sex and method to the weighted sample data using chi-square goodness-of-fit tests, or Z-tests if there was only one level<sup>225,226</sup>.

**Drug consumption, quantity, and frequency** The NDSHS results for marijuana, amphetamines, and heroin were compared by sex and method to the sample weighted data using a chi-square goodness-of-fit test, or Z-tests if there was only one level for consumption status (*lifetime use, use in the last year*), frequency of use (*at least monthly, less often*), and age distribution of users (*20–39,  $\geq 40$* )<sup>225,226</sup>.

## 5.4 Results

The study was conducted between May–October 2000.

### 5.4.1 Comparisons between interview methods

#### Outcome rate

Table 5.1 presents the outcome rates. All methods had similar contact rates (84–86%), but IVR obtained a response rate (12%) significantly less than CATI (61%). Compared to CATI, Hybrid II had a significantly lower cooperation rate (66% vs 73%) and higher refusal rate (27% vs 21%). Re-assignment of respondents with unsuitable telephones resulted in an inflated CATI method sample and deflated Hybrid method samples.



Table 5.1: Outcome rates and sample sizes by telephone interview method.

	<i>CATI</i>	<i>Hybrid I</i>	<i>Hybrid II</i>	<i>IVR</i>
<i>Rate</i>	%	%	%	%
Contact rate	84	84	85	86
Cooperation rate	73	67	66	14
Response rate	61	56	56	12
Refusal rate	21	25	27	74
<i>Sample</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>
Sample size	661	706	697	806
Respondents	316	322	319	90
After reassignment <sup>b</sup>	390	272	295	90

<sup>a</sup> P values for chi-square comparing rates for each method to CATI.

<sup>b</sup> Number of respondents after reassignment due to telephones not being suitable for use with Hybrid or IVR methods.

### Sample demographic profile

As shown in Table 5.2, the unweighted sample distributions for all demographic characteristics of Hybrid I, Hybrid II, and IVR did not differ significantly from CATI. However, as shown in Table 5.3, compared to the ABS Census data<sup>208</sup>, CATI marginally over-represented Australian-born persons; Hybrid I marginally over-represented the second population age quartile (31–42), and the *married/de facto* category; and Hybrid II marginally over-represented Australian-born persons, those in the labour force, those with university qualifications, and males. IVR did not deviate on any variable, possibly because of insufficient sample size.

### Interview duration

Figure 5.2 shows the interview duration for fully completed interviews for each method. For the Hybrid methods, the durations for the interviewer and GEIS running on the RVS are also shown. The central bar of each box plot indicates the median interview duration, the central box indicates the 25th and 75th

Table 5.2: Unweighted respondent demographic composition for each nominal<sup>a</sup> telephone interview method.

<i>Variable</i>		<i>CATI</i>	<i>Hybrid I</i>	<i>Hybrid II</i>	<i>IVR</i>
		%	%	%	%
<b>Sex</b>					
	Male	37	34	42	33
	P <sup>b</sup>		0.43	0.26	0.41
<b>Age<sup>c</sup></b>					
	<i>Q</i> <sub>1</sub>	20	16	20	23
	<i>Q</i> <sub>2</sub>	24	32	31	21
	<i>Q</i> <sub>3</sub>	26	26	27	36
	<i>Q</i> <sub>4</sub>	30	26	23	20
	P		0.15	0.17	0.14
<b>Education</b>					
	<i>School or none</i>	57	55	50	55
	<i>Vocational</i>	22	24	23	18
	<i>University</i>	21	21	27	27
	P		0.74	0.13	0.45
<b>Marital status</b>					
	<i>Married/de facto</i>	61	64	66	68
	<i>Divorced/Separated</i>	13	12	11	11
	<i>Widowed</i>	8	10	5	2
	<i>Never married</i>	19	15	17	19
	P		0.45	0.38	0.25
<b>Country of birth</b>					
	<i>Australia</i>	79	75	79	83
	P		0.23	0.99	0.38
<b>Employment</b>					
	<i>In labour force</i>	58	55	64	63
	P		0.50	0.12	0.42
<b>n<sup>d</sup></b>		315	321	277	88

<sup>a</sup> The nominal telephone interview method was the one to which respondents were randomly assigned during sample selection.

<sup>b</sup> Two-sided P values, chi-square tests comparing Hybrid I, Hybrid II, and IVR to CATI.

<sup>c</sup> Quartiles of the Australian population aged 18 years or older. *Q*<sub>1</sub>: 18–30 years old, *Q*<sub>2</sub>: 31–42 years old, *Q*<sub>3</sub>: 43–57 years old, *Q*<sub>4</sub>: ≥58 years old.

<sup>d</sup> Number of respondents reporting demographic information.

Table 5.3: Weighted respondent demographic composition by nominal<sup>a</sup> telephone interview method compared to the Australian population.

<i>Variable</i>		<i>CATI</i>	<i>Hybrid I</i>	<i>Hybrid II</i>	<i>IVR</i>	<i>ABS</i>
		%	%	%	%	%
<b>Sex</b>						
	<i>Male</i>	48	45	55	43	49
	P <sup>b</sup>	0.10	0.23	0.03	0.29	
<b>Age<sup>c</sup></b>						
	<i>Q<sub>1</sub></i>	25	20	26	21	26
	<i>Q<sub>2</sub></i>	24	30	29	21	25
	<i>Q<sub>3</sub></i>	23	24	25	37	25
	<i>Q<sub>4</sub></i>	28	26	20	21	24
	P	0.36	0.03	0.39	0.12	
<b>Education</b>						
	<i>School or none</i>	57	52	46	53	54
	<i>Vocational</i>	22	26	26	22	25
	<i>University</i>	21	22	28	25	21
	P	0.41	0.69	0.01	0.64	
<b>Marital status</b>						
	<i>Married/de facto</i>	58	62	65	69	58
	<i>Divorced/Separated</i>	13	11	10	11	11
	<i>Widowed</i>	7	10	4	2	7
	<i>Never married</i>	22	17	21	18	24
	P	0.65	0.03	0.12	0.17	
<b>Country of birth</b>						
	<i>Australia</i>	79	74	79	80	73
	P	0.02	0.92	0.03	0.22	
<b>Employment</b>						
	<i>In labour force</i>	60	59	72	67	62
	P	0.54	0.32	<0.01	0.33	
<b>n<sup>d</sup></b>		315	321	277	88	

<sup>a</sup> The nominal telephone interview method was the one to which respondents were randomly assigned during sample selection.

<sup>b</sup> Two-sided P values, chi-square goodness-of-fit tests comparing each method to ABS Census data<sup>208</sup>.

<sup>c</sup> Quartiles of the Australian population aged 18 years or older. *Q<sub>1</sub>*: 18–30 years old, *Q<sub>2</sub>*: 31–42 years old, *Q<sub>3</sub>*: 43–57 years old, *Q<sub>4</sub>*: ≥58 years old.

<sup>d</sup> Number of respondents reporting demographic information.

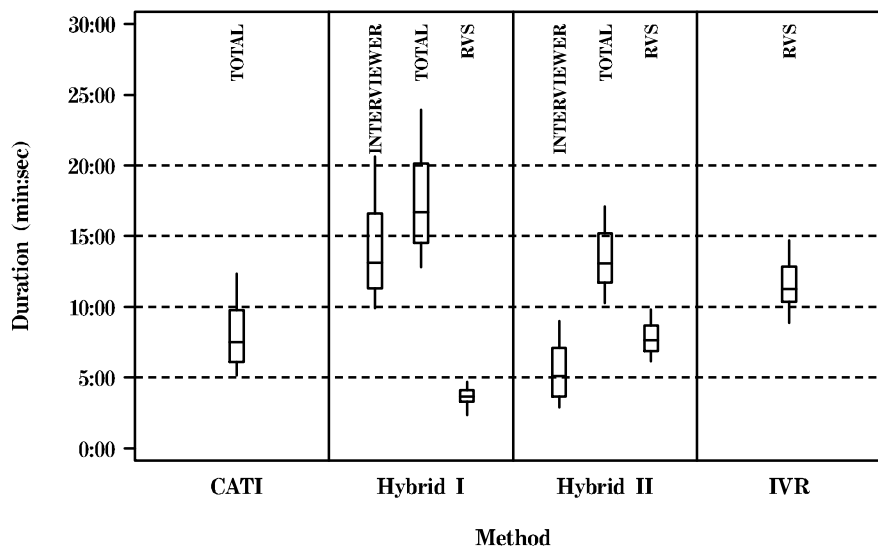


Figure 5.2: Interview durations shown as box plots for the whole interview, the part conducted by the interviewer, and the part conducted by the RVS.

percentiles, and the whiskers indicate the 10th and 90th percentiles.

The median total duration differed significantly for all methods ( $\chi^2=616$ ,  $df=3$ ,  $P<0.01$ ,  $n=1006$ ), with CATI being the briefest method and Hybrid I the longest. The median interviewer duration also differed significantly ( $\chi^2=502$ ,  $df=2$ ,  $P<0.01$ ,  $n=917$ ) between methods, with Hybrid II being briefest and Hybrid I the longest.

### Calling effort

As shown in Table 5.4, the number of calls to complete an interview did not differ significantly between methods, but the number of calls to contact respondents was significantly higher for IVR.

To investigate how many calls were needed to obtain stable estimates, the values of important variables were plotted against the number of calls to complete an interview. The variables with corresponding figures were: mean age

Table 5.4: Mean number of calls by method.

<i>Status</i>	<i>CATI</i>	<i>Hybrid I</i>	<i>Hybrid II</i>	<i>IVR</i>	<i>P</i> <sup>a</sup>
To contact respondents	2.1	2.1	1.9	2.6	<0.01
To complete interviews	2.9	2.8	2.9	2.7	0.86

<sup>a</sup> Poisson regression.

(Figure 5.3), percentage of females (Figure 5.4), AUDIT score (Figure 5.5), percentage consuming alcohol (Figure 5.6), percentage consuming marijuana (Figure 5.7), percentage consuming amphetamines (Figure 5.8), percentage consuming heroin (Figure 5.9). By inspection of the graphs in the figures it may be seen that within seven calls the value of each variable stabilizes to within  $\pm 2.5\%$  of the mean value obtained after fifteen calls. No obvious differences between methods were noted.

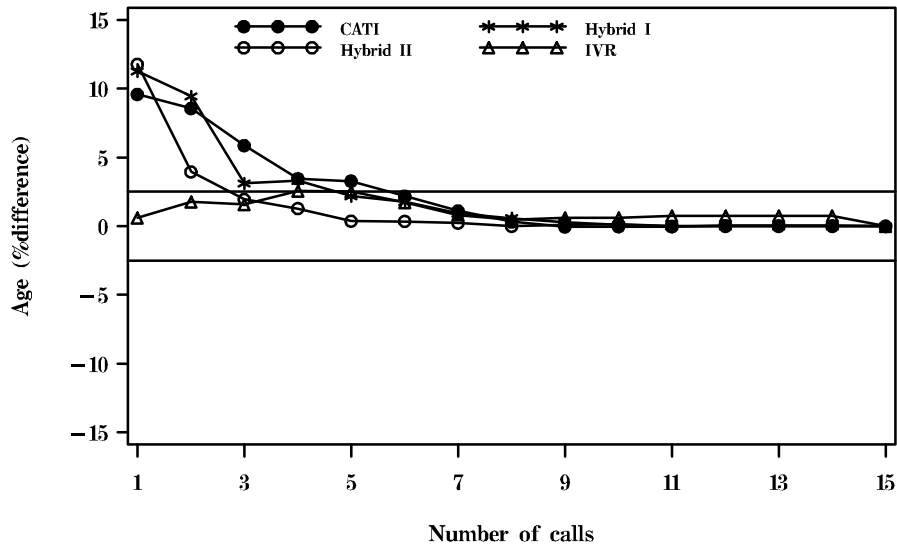


Figure 5.3: Percentage change in age by number of calls to complete interviews

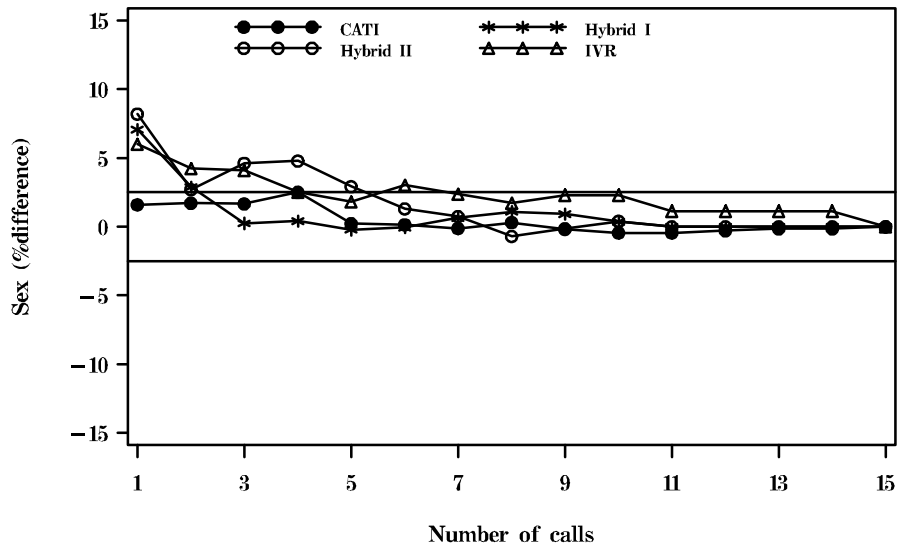


Figure 5.4: Percentage change in proportion of female respondents by number of calls to complete interviews with  $\pm 2.5\%$  limits.

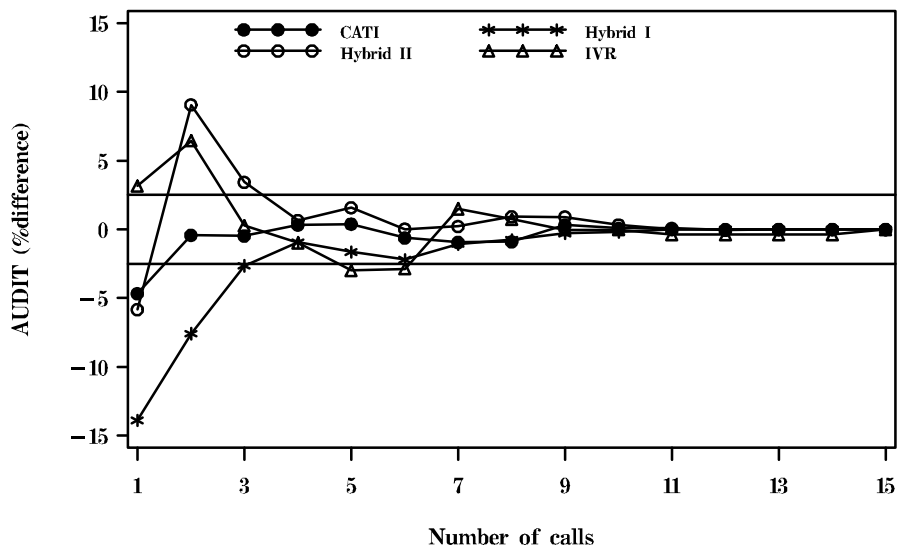


Figure 5.5: Percentage change in AUDIT score by number of calls to complete interviews with  $\pm 2.5\%$  limits.

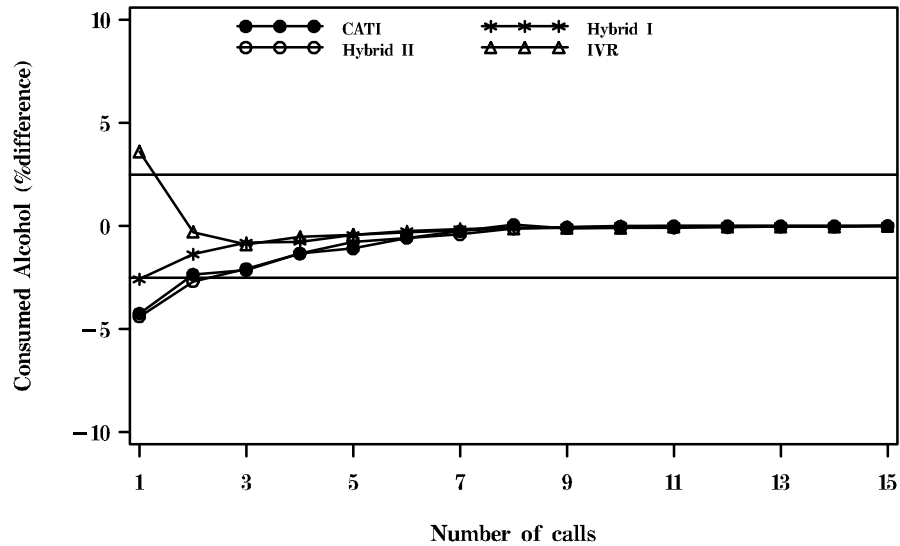


Figure 5.6: Percentage change in proportion consuming alcohol by number of calls to complete interviews with  $\pm 2.5\%$  limits.

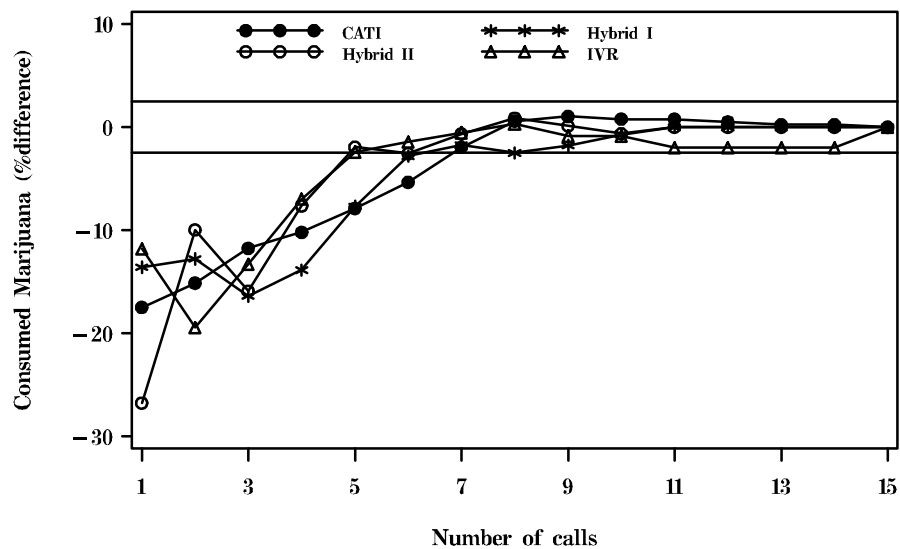


Figure 5.7: Percentage change in proportion consuming marijuana by number of calls to complete interviews with  $\pm 2.5\%$  limits.

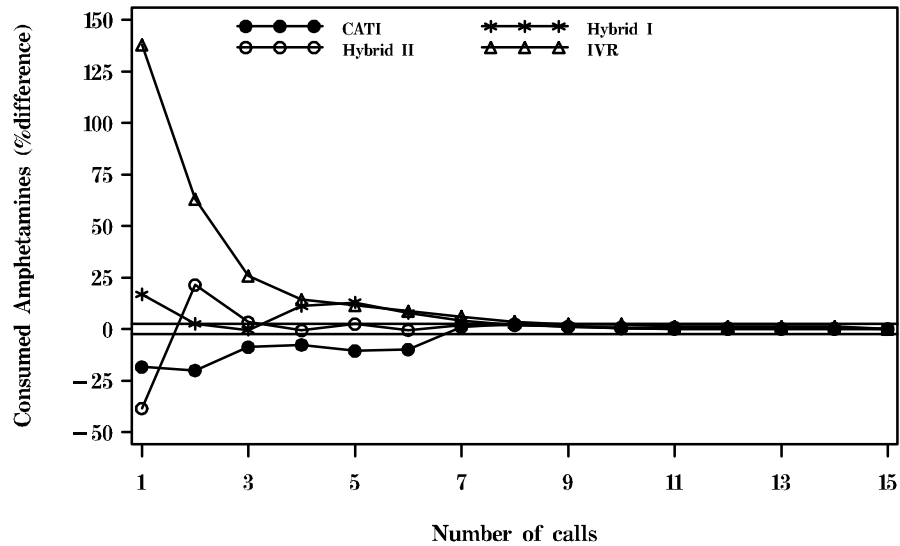


Figure 5.8: Percentage change in proportion consuming amphetamines by number of calls to complete interviews with  $\pm 2.5\%$  limits.

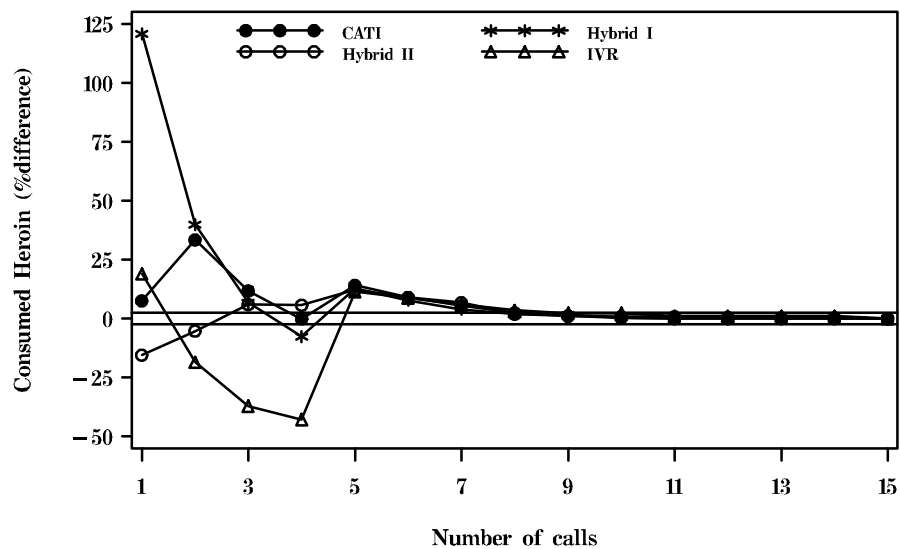


Figure 5.9: Percentage change in proportion consuming heroin by number of calls to complete interviews with  $\pm 2.5\%$  limits.



Table 5.5: Item non-response rate by method.

<i>Domain</i>	<i>CATI</i>	<i>Hybrid I</i>	<i>Hybrid II</i>	<i>IVR</i>	<i>P<sup>a</sup></i>
Technology	0.3	0.4	0.4	0	0.60
Alcohol	0.3	1.8	0.5	0	<0.01
Drugs	1	8	1.9	1.1	<0.01
Demography	1.6	1.3	11.3	7.4	<0.01

<sup>a</sup> P values for chi-square tests comparing all methods.

Table 5.6: Mean cost by method.

<i>Cost type</i>	<i>CATI</i>	<i>Hybrid I</i>	<i>Hybrid II</i>	<i>IVR</i>	<i>P<sup>a</sup></i>
Mean <sup>b</sup>	4.97	9.46	3.64	2.27	<0.01
Total <sup>c</sup>	6.03	11.44	5.27	7.92	<0.01

<sup>a</sup> Analysis of variance.

<sup>b</sup> Mean cost per completed interview.

<sup>c</sup> Total survey cost per completed interview.

### Item non-response rate

As shown in Table 5.5, the item non-response rates differed significantly between methods for the alcohol, drugs and demography domains. In comparison to CATI, Hybrid I had higher item non-response rates for the drugs domain, while Hybrid II and IVR had higher item non-response rates for the demography domain.

### Cost

As shown in Table 5.6, the mean completed interview cost and total survey cost per completed interview both differed significantly between methods. IVR obtained the minimum interview cost, Hybrid I had the largest interview cost and survey cost per completed interview, and Hybrid II had the minimum survey cost per completed interview.

Table 5.7: Alcohol, marijuana, amphetamines and heroin consumption by method.

<i>Topic</i>	<i>CATI</i>	<i>Hybrid I</i>	<i>Hybrid II</i>	<i>IVR</i>	<i>P1</i> <sup>a</sup>	<i>P2</i> <sup>b</sup>	<i>T</i> <sup>c</sup>
	%	%	%	%			
<b>Alcohol</b>							
Full glass	91	97	94	97	0.02	<0.01	1.
Age <sup>d</sup>	17	17	17	17	0.14	0.94	2.
Hazard <sup>e</sup>	21	29	32	30	0.03	<0.01	1.
<b>Marijuana</b>							
Ever tried	29	35	38	37	0.12	0.02	1
Age <sup>d</sup>	18	18	17	20	0.05	0.52	2.
12 months <sup>f</sup>	8	7	9	12	0.46	0.52	1.
Monthly <sup>g</sup>	2	2	6	2	0.07	0.96	1.
<b>Amphetamines</b>							
Ever tried	6	8	8	5	0.58	0.45	1.
Age <sup>d</sup>	18	19	19	16	0.29	0.50	2.
12 months <sup>f</sup>	1	1	3	4	0.03	0.08	1.
Monthly <sup>g</sup>	0	0.4	0.7	2	0.05	0.08	1.
<b>Heroin</b>							
Ever tried	1	2	3	2	0.42	0.27	1.
Age <sup>d</sup>	21	22	22	28	0.99	1.00	2.
12 months <sup>f</sup>	0.3	0	1	2	0.04	0.20	1.
Monthly <sup>g</sup>	0	0	0	1	0.13	0.28	1.
<b>n</b> <sup>h</sup>	390	272	295	90			

<sup>a</sup> Two-sided P values; any differences between methods.<sup>b</sup> Two-sided P values; comparing CATI with all other methods.<sup>c</sup> Statistical tests: 1=chi-square, 2=Wilcoxon-Mann-Whitney test.<sup>d</sup> Median age first tried.<sup>e</sup> Hazardous drinking.<sup>f</sup> Used in last twelve months.<sup>g</sup> Used at least Monthly.<sup>h</sup> Number of respondents reporting alcohol and drug information.

### Self-report rate of alcohol and drugs

As shown in Table 5.7, compared to the CATI method the non-CATI methods obtained a significantly higher proportion of respondents who reported consuming alcohol or marijuana, and consuming hazardous levels of alcohol. The proportion consuming amphetamines in the previous twelve months differed significantly, with Hybrid II and IVR being the highest. The ages of respondents first trying alcohol or a drug did not vary significantly between the methods.

Table 5.8: Median acceptability scores<sup>a</sup> (ease, enjoyable, stressful, likeable) for the interview methods.

<i>Measure</i>	<i>CATI</i>	<i>Hybrid I</i>	<i>Hybrid II</i>	<i>IVR</i>	<i>P-Value</i> <sup>b</sup>
Ease	1	1	2	1	<0.01
Enjoyable	2	3	3	2	<0.01
Stressful	2	2	2	2	0.15
Likeable	3	2	2	3	0.56

<sup>a</sup> These measures are explained in the text.

<sup>b</sup> P-values for Kruskal-Wallis tests.

### Acceptability

The acceptability of each method is summarized in Tables 5.8 and 5.9. Table 5.8 shows the median score obtained for each acceptability measure. As explained in Section 5.3.7, the lowest possible score (1) corresponded to ‘very easy’, ‘very enjoyable’, ‘not stressful’, and ‘like it a lot’, and the highest possible score (4 or 5) corresponded to ‘very hard’, ‘not enjoyable’, ‘very stressful’, and ‘dislike it a lot’.

The methods were not distinguished by stressfulness or likeableness, but differences were found for ease and enjoyableness. After Bonferroni corrections, Hybrid II was rated significantly less easy than CATI (Wilcoxon test,  $W=5490$ ,  $P<0.01$ ,  $n=157$ ), and Hybrid I and Hybrid II were rated less enjoyable than CATI (Wilcoxon test,  $W=15444$ ,  $P<0.01$ ,  $n=368$ ; Wilcoxon test,  $W=5538$ ,  $P<0.01$ ,  $n=157$ ).

As shown in Table 5.9, with increasing automation (CATI, Hybrid I, Hybrid II, IVR), there were significant decreasing trends for preference for an interviewer, and in believing people would be more likely to be honest with a interviewer.

Spearman’s rank correlations<sup>225</sup> were calculated for acceptability measures

Table 5.9: Preference and honesty ratings of interview methods.

<i>Measure</i>	<i>CATI</i>	<i>Hybrid I</i>	<i>Hybrid II</i>	<i>IVR</i>	<i>P-Value</i> <sup>a</sup>
	%	%	%	%	
Preference (n=457) <sup>b</sup>	94	47	43	33	<0.01
Honesty (n=457) <sup>c</sup>	57	26	18	5	<0.01

<sup>a</sup> Cochran-Armitage trend test with modified ridit scores.

<sup>b</sup> Preference for an interviewer.

<sup>c</sup> Percentage agreeing with the statement: "People are more likely to be honest with their alcohol use to a person than a recorded voice".

and the acceptability score for liking technology. Wilcoxon tests were used to assess whether acceptability scores were associated with ownership of mobile telephones, previous use of an IVR system, use of answering machines, or home access to the Internet.

Only a few correlations were significant as determined by Monte Carlo probability estimates<sup>223</sup>. Those liking technology thought the RVS interview was easier ( $r_s=0.14$ ,  $P<0.01$ ,  $n=468$ ) and less stressful ( $r_s=0.10$ ,  $P=0.04$ ,  $n=468$ ).

Those who had previously used an IVR system found the RVS interview more enjoyable (Wilcoxon test,  $W=15556$ ,  $P<0.01$ ,  $n=468$ ) and liked it (Wilcoxon test,  $W=16171$ ,  $P=0.02$ ,  $n=468$ ) than those who had not experienced one before. Those with Internet access thought the RVS interview was easier (Wilcoxon test,  $W=40799$ ,  $P<0.01$ ,  $n=468$ ) and less stressful (Wilcoxon test,  $W=16171$ ,  $P=0.02$ ,  $n=468$ ) than those with no access.

#### 5.4.2 Comparison of the results to the National Drug Strategy Household Survey

The consumption of alcohol, marijuana, amphetamines, and heroin for each method were compared to the NDSHS<sup>213</sup> results. To maximize the power of

the comparisons, the two Hybrid methods were combined after testing for homogeneity. The results are presented below for alcohol and each drug in turn. The data were weighted by age and sex according to the ABS Census data<sup>208</sup>.

### Alcohol

Since the distributions of drinking status, quantity, and frequency did not differ between Hybrid I and Hybrid II for males ( $\chi^2=3.0$ ,  $df=3$ ,  $P=0.40$ ,  $n=200$ ;  $\chi^2=2.6$ ,  $df=3$ ,  $P=0.47$ ,  $n=170$ ;  $\chi^2=8.5$ ,  $df=4$ ,  $P=0.08$ ,  $n=170$ ) or females ( $\chi^2=2.6$ ,  $df=3$ ,  $P=0.45$ ,  $n=324$ ;  $\chi^2=5.1$ ,  $df=3$ ,  $P=0.16$ ,  $n=251$ ;  $\chi^2=9.0$ ,  $df=4$ ,  $P=0.06$ ,  $n=251$ ) the data from the two methods were combined in further analyses.

As shown in Table 5.10, the Hybrid and IVR methods did not differ significantly from the NDSHS results with respect to drinking status. However, the CATI method differed significantly from the NDSHS for both males and females. CATI female respondents under-reported being ex-drinkers and over-reported never having drunk, while male respondents over-reported being regular drinkers.

As shown in Table 5.11 compared to the NDSHS, the Hybrid and IVR methods did not differ significantly from the NDSHS results with respect to quantity of alcohol consumed. However, CATI female respondents over-reported drinking between one and two drinks a day and under-reported drinking at least five drinks a day, while male respondents under-reported having more than seven drinks a day.

As shown in Table 5.12 the Hybrid and IVR methods did not differ significantly from the NDSHS results with respect to frequency of drinking alcohol. However, CATI female respondents under-reported drinking on four to six days

Table 5.10: Drinking status by method for males and females, compared to the NDSHS results.

<i>Method</i>	<i>Sex</i>	<i>Drinking Status</i>	<i>%<sup>a</sup></i>	<i>NDSHS (%)</i>	<i>P<sup>b</sup></i>
CATI	Female (n=233)	<i>Regular</i>	39	39	<0.01
		<i>Occasional</i>	33	39	
		<i>Ex-drinker</i>	7	11	
		<i>Never drank</i>	21	12	
	Male (n=133)	<i>Regular</i>	72	59	0.03
		<i>Occasional</i>	19	25	
		<i>Ex-drinker</i>	5	9	
		<i>Never drank</i>	4	7	
Hybrid <sup>c</sup>	Female (n=301)	<i>Regular</i>	39	39	0.62
		<i>Occasional</i>	40	39	
		<i>Ex-drinker</i>	11	11	
		<i>Never drank</i>	10	12	
	Male (n=186)	<i>Regular</i>	59	59	0.63
		<i>Occasional</i>	28	25	
		<i>Ex-drinker</i>	7	9	
		<i>Never drank</i>	6	7	
IVR	Female (n=53)	<i>Regular</i>	39	39	0.37
		<i>Occasional</i>	38	39	
		<i>Ex-drinker</i>	17	11	
		<i>Never drank</i>	6	12	
	Male (n=24)	<i>Regular</i>	54	59	0.52
		<i>Occasional</i>	20	25	
		<i>Ex-drinker</i>	16	9	
		<i>Never drank</i>	10	7	

<sup>a</sup> Weighted percentage.<sup>b</sup> Goodness-of-fit P Values.<sup>c</sup> Combined Hybrid I and Hybrid II methods.

Table 5.11: Alcohol quantity by method for males and females, compared to the NDSHS results.

<i>Method</i>	<i>Sex</i>	<i>Drinks per day.</i>	<i>%<sup>a</sup></i>	<i>NDSHS (%)</i>	<i>P<sup>b</sup></i>
CATI	Female (n=169)	1-2	74	62	<0.01
		3-4	21	22	
		5-6	3	9	
		≥7	2	7	
	Male (n=120)	1-2	45	36	<0.01
		3-4	27	31	
		5-6	21	15	
		≥7	7	18	
Hybrid <sup>c</sup>	Female (n=237)	1-2	66	62	0.53
		3-4	20	22	
		5-6	9	9	
		≥7	5	7	
	Male (n=159)	1-2	36	36	0.07
		3-4	37	31	
		5-6	16	15	
		≥7	11	18	
IVR	Female (n=41)	1-2	74	62	0.46
		3-4	14	22	
		5-6	6	9	
		≥7	6	7	
	Male (n=17)	1-2	43	36	0.88
		3-4	26	31	
		5-6	17	15	
		≥7	14	18	

<sup>a</sup> Weighted percentage.<sup>b</sup> Goodness-of-fit P Values.<sup>c</sup> Combined Hybrid I and Hybrid II methods.

a week, while male respondents under-reported drinking on four to six days a week and less often than daily, but over-reported daily drinking.

### **Marijuana**

No significant difference in distributions was found for males between Hybrid I and Hybrid II for lifetime marijuana consumption ( $\chi^2=2.9$ ,  $df=1$ ,  $P=0.09$ ,  $n=198$ ), consumption in the last year ( $\chi^2=0.09$ ,  $df=1$ ,  $P=0.66$ ,  $n=198$ ), frequency of use ( $\chi^2=3.7$ ,  $df=1$ ,  $P=0.06$ ,  $n=23$ ), and age distribution ( $\chi^2=2.7$ ,  $df=1$ ,  $P=0.10$ ,  $n=187$ ). Similarly, no difference in distributions was found for females between Hybrid I and Hybrid II for lifetime marijuana consumption ( $\chi^2=0.48$ ,  $df=1$ ,  $P=0.49$ ,  $n=323$ ), consumption in the last year ( $\chi^2=1.4$ ,  $df=1$ ,  $P=0.24$ ,  $n=322$ ), frequency of use ( $\chi^2=0.15$ ,  $df=1$ ,  $P=0.70$ ,  $n=22$ ), and age distribution ( $\chi^2=0.67$ ,  $df=1$ ,  $P=0.42$ ,  $n=304$ ). Therefore, the data from the two methods were combined for the above variables.

As shown in Table 5.13, using the Hybrid method the lifetime use for both males and females did not differ significantly from the NDSHS, but use in the last year was less than the NDSHS for both males and females. The female lifetime and use in the last year was under-reported using CATI. Male self-report rates did not differ significantly from the NDSHS, but the under-reporting of use in the last year was close-to-significant. None of the IVR self-report rates differed significantly from the NDSHS.

Analyses were not attempted for frequency because of small sample sizes. The age distributions for those consuming marijuana (Table 5.14) did not differ significantly from the NDSHS for any method except for females in the IVR method.



Table 5.12: Frequency of alcohol consumption by method for males and females, compared to the NDSHS results.

<i>Method</i>	<i>Sex</i>	<i>Frequency</i>	<i>%<sup>a</sup></i>	<i>NDSHS (%)</i>	<i>P<sup>b</sup></i>
CATI	Female (n=186)	<i>Every day</i>	11	7	<0.01
		<i>4-6 days/week</i>	4	10	
		<i>2-3 days/week</i>	17	16	
		<i>1 day/ week</i>	18	17	
		<i>Less often</i>	50	50	
	Male (n=197)	<i>Every day</i>	21	11	<0.01
		<i>4-6 days/week</i>	5	13	
		<i>2-3 days/week</i>	24	20	
		<i>1 day/ week</i>	25	17	
		<i>Less often</i>	25	40	
Hybrid <sup>c</sup>	Female (n=271)	<i>Every day</i>	5	7	0.11
		<i>4-6 days/week</i>	10	10	
		<i>2-3 days/week</i>	18	16	
		<i>1 day/ week</i>	10	17	
		<i>Less often</i>	57	50	
	Male (n=174)	<i>Every day</i>	16	11	0.37
		<i>4-6 days/week</i>	13	13	
		<i>2-3 days/week</i>	17	20	
		<i>1 day/ week</i>	17	17	
		<i>Less often</i>	37	40	
IVR	Female (n=50)	<i>Every day</i>	12	7	0.10
		<i>4-6 days/week</i>	7	10	
		<i>2-3 days/week</i>	18	16	
		<i>1 day/ week</i>	5	17	
		<i>Less often</i>	58	50	
	Male (n=21)	<i>Every day</i>	13	11	0.54
		<i>4-6 days/week</i>	0	13	
		<i>2-3 days/week</i>	32	20	
		<i>1 day/ week</i>	15	17	
		<i>Less often</i>	40	40	

<sup>a</sup> Weighted percentage.<sup>b</sup> Goodness-of-fit P Values.<sup>c</sup> Combined Hybrid I and Hybrid II methods.

Table 5.13: Consuming marijuana by method for males and females, compared to the NDSHS results.

<i>Method</i>	<i>Sex</i>	<i>Use</i>	<i>%<sup>a</sup></i>	<i>NDSHS (%)</i>	<i>P<sup>b</sup></i>
CATI	Male (n=133)	<i>Lifetime</i>	42	44	0.66
		<i>Last year</i>	15	21	0.06
	Female (n=233)	<i>Lifetime</i>	24	35	<0.01
		<i>Last year</i>	5	15	<0.01
Hybrid <sup>c</sup>	Male (n=184)	<i>Lifetime</i>	45	44	0.85
		<i>Last year</i>	13	21	<0.01
	Female (n=300)	<i>Lifetime</i>	35	35	0.82
		<i>Last year</i>	7	15	<0.01
IVR	Male (n=24)	<i>Lifetime</i>	58	44	0.16
		<i>Last year</i>	14	21	0.39
	Female (n=52)	<i>Lifetime</i>	27	35	0.24
		<i>Last year</i>	11	15	0.43

<sup>a</sup> Weighted percentage.<sup>b</sup> P values for Z-tests.<sup>c</sup> Combined Hybrid I and Hybrid II methods.

Table 5.14: Consuming marijuana by method by age and sex, compared to the NDSHS results.

<i>Method</i>	<i>Sex</i>	<i>Age</i>	<i>%<sup>a</sup></i>	<i>NDSHS (%)</i>	<i>P<sup>b</sup></i>
CATI	Female (n=56)	<i>20-39</i>	66	72	0.37
		<i>≥40</i>	33	28	
	Male (n=50)	<i>20-39</i>	59	65	0.36
		<i>≥40</i>	41	35	
Hybrid <sup>c</sup>	Female (n=100)	<i>20-39</i>	66	72	0.23
		<i>≥40</i>	33	28	
	Male (n=75)	<i>20-39</i>	65	65	0.89
		<i>≥40</i>	35	35	
IVR	Female (n=13)	<i>20-39</i>	46	72	0.03
		<i>≥40</i>	54	28	
	Male (n=14)	<i>20-39</i>	51	65	0.78
		<i>≥40</i>	49	35	

<sup>a</sup> Weighted percentage.<sup>b</sup> Goodness-of-fit P Values.<sup>c</sup> Combined Hybrid I and Hybrid II methods.

**Amphetamines**

No significant difference in distributions was found for males between Hybrid I and Hybrid II for lifetime amphetamines consumption (Fisher's exact test,  $\chi^2=0.13$ ,  $P=0.72$ ,  $n=198$ ), and recent amphetamines consumption (Fisher's exact test,  $\chi^2=2.3$ ,  $P=0.13$ ,  $n=198$ ). Similarly, no difference in distributions was found for females between Hybrid I and Hybrid II for lifetime amphetamines consumption (Fisher's exact test,  $\chi^2=0.19$ ,  $P=0.66$ ,  $n=321$ ), and recent amphetamines consumption (Fisher's exact test,  $\chi^2=0.75$ ,  $P=0.39$ ,  $n=321$ ). Therefore, the data from the two methods were combined for the above variables.

As shown in Table 5.15, using the Hybrid method the lifetime use for both males and females did not differ significantly from the NDSHS. The lifetime use and use in the last year for CATI method females was significantly less than the NDSHS results. Other comparisons with the NDSHS were not significant. Analyses were not attempted for quantity or frequency because of small sample sizes.

**Heroin**

No significant difference in distributions was found for males between Hybrid I and Hybrid II for lifetime heroin consumption (Fisher's exact test,  $\chi^2=0.84$ ,  $P=0.36$ ,  $n=199$ ), and recent heroin consumption (Fisher's exact test,  $\chi^2=1.8$ ,  $P=0.18$ ,  $n=199$ ). Similarly, no difference in distributions was found for females between Hybrid I and Hybrid II for lifetime heroin consumption (Fisher's exact test,  $\chi^2=1.2$ ,  $P=0.26$ ,  $n=323$ ), and recent heroin consumption (Fisher's exact test,  $\chi^2=2.9$ ,  $P=0.09$ ,  $n=323$ ). Therefore, the data from the two methods were combined for the above variables.

Table 5.15: Consuming amphetamines by method for males and females, compared to the NDSHS results.

<i>Method</i>	<i>Sex</i>	<i>Use</i>	<i>%<sup>a</sup></i>	<i>NDSHS (%)</i>	<i>P<sup>b</sup></i>
CATI	Female (n=233)	<i>Lifetime</i>	3	7	0.02
		<i>Last year</i>	0	3	0.01
	Male (n=133)	<i>Lifetime</i>	9	11	0.53
		<i>Last year</i>	3	5	0.20
Hybrid <sup>c</sup>	Female (n=298)	<i>Lifetime</i>	6	7	0.88
		<i>Last year</i>	1	3	0.11
	Male (n=184)	<i>Lifetime</i>	12	11	0.61
		<i>Last year</i>	4	5	0.70
IVR	Female (n=52)	<i>Lifetime</i>	2	7	0.07
		<i>Last year</i>	2	3	0.95
	Male (n=24)	<i>Lifetime</i>	6	11	0.44
		<i>Last year</i>	6	5	0.80

<sup>a</sup> Weighted percentage.<sup>b</sup> Goodness-of-fit P Values.<sup>c</sup> Combined Hybrid I and Hybrid II methods.

As shown in Table 5.16, no significant differences between methods from the NDSHS results were found for lifetime or recent use. Analyses were not attempted for quantity or frequency because of small sample sizes.

## 5.5 Discussion

### 5.5.1 Comparisons between interview methods

The Hybrid I and Hybrid II methods performed comparably to CATI on the criteria: outcome rate, sample demographic profile, and calling effort. However, differences were found for interview duration, item non-response rate, cost, self-report rates of alcohol and drugs, and acceptability. These are discussed below.

Table 5.16: Consuming heroin by method for males and females, compared to the NDSHS results.

<i>Method</i>	<i>Sex</i>	<i>Status</i>	<i>%<sup>a</sup></i>	<i>NDSHS (%)</i>	<i>P<sup>b</sup></i>
CATI	Female (n=233)	<i>Lifetime</i>	0.3	1.4	0.15
		<i>Last year</i>	0.3	0.5	0.65
	Male (n=133)	<i>Lifetime</i>	2.4	3.0	0.70
		<i>Last year</i>	0.0	0.5	0.42
Hybrid <sup>c</sup>	Female (n=300)	<i>Lifetime</i>	1.3	1.4	0.90
		<i>Last year</i>	1.1	0.5	0.22
	Male (n=185)	<i>Lifetime</i>	5.0	3.0	0.11
		<i>Last year</i>	0.8	1.1	0.68
IVR	Female (n=52)	<i>Lifetime</i>	2.4	1.4	0.56
		<i>Last year</i>	2.4	0.5	0.06
	Male (n=24)	<i>Lifetime</i>	4.0	3.0	0.79
		<i>Last year</i>	4.0	1.1	0.18

<sup>a</sup> Weighted percentage.<sup>b</sup> Goodness-of-fit P Values.<sup>c</sup> Combined Hybrid I and Hybrid II methods.

### Outcome rate

All methods contacted similar proportions of their respective samples, but the IVR respondents were less cooperative, with fewer responding, and more refusing. A CATI response rate of 61% was obtained, which although moderate, may occur with sensitive telephone surveys<sup>232</sup>. The response rate of all methods may have been reduced by addressing letters to ‘The Household’ making them appear to be junk mail and thrown out without reading, or delivered late by the postal service. Also, the Hybrid and IVR letters may have dissuaded respondents by stating that a recorded voice would be used. The low IVR response rate, which is consistent with earlier work<sup>8,22,23</sup>, suggested that IVR requires motivated respondents, such as in institutional surveys.

**Sample demographic profile**

The Hybrid I, Hybrid II and IVR methods did not differ significantly from CATI in any demographic characteristic, but a few differences were found between CATI, Hybrid I, and Hybrid II and the ABS Census<sup>208</sup>. It appeared that relative sampling biases between methods were small compared to the bias introduced by telephone sampling itself.

**Interview duration**

The overall interview duration for fully completed CATI interviews was the shortest of all the methods. Hybrid I had the longest overall duration since it involved the respondent being transferred twice and additional explanatory messages were played by GEIS running on the RVS after the call transfer. For the methods that used an interviewer, the duration of questions asked by an interviewer was least for Hybrid II since most questions were asked by GEIS running on the RVS. For both interviewer and respondent, the Hybrid I was slower than the other methods. This study only used a brief interview. If a longer Hybrid I interview had been used the time spent in call transfer would have been a smaller proportion of the overall duration.

**Calling effort**

All methods required an equivalent number of calls to complete interviews, but IVR obtained a greater number of contact calls. This was because the IVR method did not have an upper calling limit and could continue indefinitely if needed. For all methods the mean values of major variables appeared to stabilize after about seven calls, which is comparable to recommended limits for CATI<sup>218</sup>.

**Item non-response rate**

Hybrid II had a higher item non-response rate for the demography domain. Although the demography domain was the last domain asked in the interviews the higher rate was not due to drop-outs because only completed interviews were used to calculate the item non-response rates. In addition, the item non-response rate for the previous domain (drugs) was relatively small and comparable to the other methods. The higher item non-response rate could not be explained by respondent fatigue since they still had to listen and then select an option in order to indicate a refusal. This suggested that some respondents may have been reluctant to provide demographic information that was apparently unrelated to the survey topic: alcohol and drugs. In addition, the lower demography domain item non-response rates for the Hybrid I and CATI methods may have resulted from respondents providing socially desirable answers to interviewers instead of refusing.

Hybrid I had a higher item non-response rate for the alcohol and drugs domains than Hybrid II. This may have been because the Hybrid I respondents knew they would be transferred back to an interviewer making the method appear less anonymous than Hybrid II. CATI had a lower item non-response rate for the alcohol and drugs domains than the Hybrid methods. Since the CATI method also obtained lower self-report rates for alcohol and drugs, it may be speculated that the lower item non-response for CATI was due to respondents providing socially acceptable answers rather than declining to answer.

**Cost**

Hybrid I had the largest interview cost and survey cost per completed interview because interviewers had to wait until the interview running on the RVS completed. IVR obtained the lowest mean interview cost since it lacked a salary component. Hybrid II obtained the lowest total survey cost per completed interview, reflecting the method's efficiency: interviewers concentrated on contacting and persuading respondents to participate rather than conducting interviews.

**Self-report rate of alcohol and drugs**

One of the limitations of this study was that the small sample size obtained by IVR meant that many of the comparisons had insufficient power to detect differences from the other methods. Comparisons to the NDSHS involving amphetamines and heroin were also limited by small sample sizes. However, sample sizes of the Hybrid and CATI methods were sufficient to detect differences in proportions of approximately 11%, close to the design specification of 10%.

The higher reporting rates for consuming alcohol and marijuana for the non-CATI methods supported the hypothesis of higher reporting rates for sensitive behaviours with more anonymous methods. This result was in line with other work<sup>3,233–235</sup>. Although the data had been weighted by age and sex to match ABS census data, some of the differences between the methods might have been attributable to differences in other variables, such as education, marital status, and employment. However, the demographic profiles of the methods did not differ significantly, which suggested that any biases introduced by demographic differences would have been small compared to those resulting from telephone sampling itself.



**Acceptability**

Although all methods were rated as acceptable, the Hybrid methods were least enjoyable and Hybrid II was the least easy. However, the marked decreasing trend in preference for an interviewer with increasing automation (CATI to IVR) suggested that attitudes to recorded voice interviews might be related to exposure. This was supported by the parallel decreasing trend in believing people would be more likely to be honest with an interviewer. These results suggested that use of automated methods was acceptable and became more so with exposure. It also accorded with the findings of other studies<sup>216,221,236,237</sup>. Katz et al.<sup>236</sup> found that individuals were more likely to like voice mail systems if they had experience of them. Rosen et al.<sup>216</sup> reported that university students who had done courses that used computers not involving programming had lower anxiety scores when using computers and more positive attitudes toward computers. Rosen and Weil<sup>221</sup> found that the reluctance to use technological devices by university students was less in those who had received computer training. In a survey of 1012 householders about attitudes to computers, McQuarrie and Iwamoto<sup>237</sup> found that exposure to computers, and not necessarily their use, was sufficient to cause a positive shift in attitude towards them.

Few strong effects were found between attitudes to technology and acceptability. This is in agreement with other work<sup>238</sup> that reports that technophobia is not more common in older or younger age groups, nor in males than females. However, it appears that the Hybrid and IVR interview methods may be slightly more popular with persons who use technology or had used IVR before.

### **5.5.2 Comparison of the results to the National Drug Strategy Household Survey**

As discussed in Section 5.3.6, the National Drug Strategy Household Survey (NDSHS) was used as an external comparison. The response rates for the CATI and Hybrid methods (56%–61%) were very comparable to the response rates achieved by the NDSHS study (55%–61%)<sup>219</sup>. Comparisons were made between the NDSHS results and the results of this study after weighting by age and sex. It may be concluded that differences between the two studies cannot be accounted for by response rate differences or sample composition differences relating to age and sex.

#### **Alcohol drinking status, quantity, and frequency**

For alcohol, the combined Hybrid methods results generally agreed with the NDSHS (drinking status, frequency, and quantity), whereas CATI deviated in the direction expected for sensitive questions. Self-report rates for more sensitive topics were lower when the less confidential method (CATI) was used. The IVR results also agreed with the NDSHS, but the sample size in this case was too small to be informative.

#### **Drug consumption, quantity, and frequency**

With the exception of marijuana consumption, the Hybrid methods produced similar reports to the NDSHS while CATI rates were under-reported for both marijuana and amphetamines. Other comparisons with the NDSHS, such as frequency of marijuana consumption and age distribution of marijuana users, were limited by the small sample size of identified users. Such comparisons would

require a survey sample size of a similar magnitude to that of the NDSHS.

The general conclusion to be drawn is that the Hybrid method can obtain a comparable result to a more expensive face-to-face survey, and within the sampling limitations of telephone surveys, would be a preferred option when sensitive questions are to be asked.

## 5.6 Conclusions

By assuming that the NDSHS was a gold standard, Hybrid I and Hybrid II provided more accurate telephone survey data on sensitive topics than CATI, while Hybrid II also did so at a lesser cost. While IVR had a low response rate, it would probably be most useful in business surveys or with motivated respondents.

Although IVR and the Hybrid methods all used a recorded voice to ask questions, only the Hybrid methods had acceptable response rates. This suggested that the initial human contact was important in persuading respondents to cooperate. Human contact also helped to keep the respondent on the line, as was evidenced by the lower cooperation rate for Hybrid II than Hybrid I. Once the RVS interview began, GEIS could ask questions in a more consistent fashion than an interviewer, but to keep the cost below that of CATI the interviewer must not wait for the call to transfer back before ringing another number. Cost can be minimized by reducing the questions the interviewer asks but without eliminating the human contact. This would allow interviewers to concentrate on recruiting respondents.

In conclusion, the IVR method is not preferred for household surveys due

to its low response rate. Instead, the Hybrid method is preferred to CATI for assessing sensitive behaviours since it obtained similar demographic sample profiles and response rates, but higher self-report rates on sensitive issues, and when applied to alcohol and drug self-report, produced results comparable to those obtained by anonymous self-completed NDSHS questionnaires.

## Chapter 6

# Interactive Voice Response as a brief intervention method: A pilot study

### 6.1 Introduction

The previous chapters provided evidence that Interactive Voice Response (IVR) could be used to ask sensitive questions about health behaviours and obtain higher self-report. It was thought that IVR might also allow health promotion advice to be provided to members of a community at relatively minimal cost. Previous work in this area had only used IVR to provide advice to clinic patients<sup>30,31</sup> or to assess risky behaviours using inbound calls<sup>34,44</sup>. There had been no previous use of IVR as a population health intervention tool.

In order to evaluate the potential of IVR for use in population health interventions it was necessary to identify a health promotion area in which it might be useful. Cervical screening was selected for the following reasons: (a) cervical cancer carried a significant burden of illness, (b) cervical screening was a potentially modifiable behaviour, and (c) the outcome of the intervention could be determined by use of readily available statistics. However, before an IVR

intervention could be attempted the IVR method needed to be piloted to assess participation rates and acceptability.

This chapter discusses the rationale for the cervical screening intervention and describes the pilot study. The next chapter describes the main study.

### 6.1.1 Cervical cancer burden of illness

Cervical cancer is the eleventh most common cancer affecting Australian women<sup>239</sup> and the thirteenth most common form of cancer death in New South Wales (NSW) women<sup>240</sup>. In American women, the cumulative lifetime risk of invasive cervical cancer is 3.4%<sup>241</sup>. Cervical cancer is preceded by asymptomatic changes in cells lining the cervix known as cervical intraepithelial neoplasia<sup>240</sup>. It is mainly caused by infection with the human papilloma virus, of which there are at least twenty types, acquired during sexual intercourse<sup>242</sup>.

In New South Wales (NSW), screening is performed by means of routine Papanicolaou smear tests (Pap tests)<sup>243,244</sup> carried out by general practitioners, gynaecological specialists, and nurses<sup>240</sup>. The National Cervical Screening Program<sup>245</sup> recommends that a Pap test should be conducted every two years for women aged 18–69 who have an intact uterus and who have no symptoms or previous history of cervical cancer, or after first sexual intercourse, ceasing at 70 years of age after at least two normal Pap tests. Women who have had their cervix removed, such as by hysterectomy, are no longer considered at risk of cervical cancer. Cervical screening using Pap tests can be cost-effective when used every two years<sup>241</sup> as long as intensive recruitment is used<sup>246</sup>.

The age-standardized rate of cervical cancer in NSW is 10.0 per 10<sup>5</sup> women-years, and the age-specific incidence increases rapidly with age to 20.5 cases

per  $10^5$  at 35–39 years of age, then gradually increases to 23.1 per  $10^5$  for women aged 70 years or older<sup>240</sup>. Recent studies suggest that many at risk women do not have regular Pap tests<sup>247–251</sup>. In 2000, the screening rate for NSW women was only 61.2% for those aged 18–49 and 60.0% for those aged 50–69<sup>252</sup>. However, the Pap test is effective in preventing cervical cancer in 90% of cases<sup>253</sup>.

### 6.1.2 Brief interventions

Given the pivotal role of regular Pap tests in reducing incidence and mortality from cervical cancer, it is essential that programs aimed at increasing screening rates continue. These programs need to be organized to overcome barriers to screening that exist at the patient, practitioner and structural levels<sup>254</sup>. For a wide range of health areas, brief advice has been shown to be as efficacious in reducing risky health-related behaviours as more intensive interventions<sup>255</sup>, as detailed in Table 6.1. The table was constructed by searching the Medline<sup>15</sup> database for articles published between mid-1998 and mid-2001. These results supported the prospect of an IVR brief intervention for cervical screening.

### 6.1.3 Strategies to increase cervical screening

Letters<sup>278–280</sup> and advice<sup>281–283</sup> from general practitioners have been shown to increase cervical screening rates relative to control women. Although a key role can be played by general practitioners in promoting cervical screening, previous research has shown that they may not consistently provide opportunistic screening advice during consultations<sup>284,285</sup>, and that women may not recall the advice provided by their doctors<sup>282,286,287</sup>. Therefore, other complementary methods

Table 6.1: Brief intervention studies: 1998–2001.

<i>Study</i>	<i>Health Area</i>
Cordoba et al.(1998) <sup>256</sup>	Non-dependent drinkers
Tomson et al.(1998) <sup>257</sup>	Hazardous drinking
Bacha and Rispoli(1999) <sup>258</sup>	Myofunctional therapy
Chang et al.(1999) <sup>259</sup>	Alcohol consumption during pregnancy
Evans et al.(1999) <sup>260</sup>	Parasuicide
Gentilello et al.(1999) <sup>261</sup>	Hazardous drinking
Irvine et al.(1999) <sup>262</sup>	Passive smoking
Monti et al.(1999) <sup>263</sup>	Hazardous drinking
Mullins et al.(1999) <sup>264</sup>	Smoking
Thome and Alder(1999) <sup>265</sup>	Fatigue
Aalto et al.(2000) <sup>266</sup>	Heavy drinking in women
Borsari and Carey(2000) <sup>267</sup>	College Student binge drinking
Freeborn et al.(2000) <sup>268</sup>	Hazardous drinking
Glasgow et al.(2000) <sup>269</sup>	Smoking
Lang et al.(2000) <sup>270</sup>	Cannabis
Manwell et al.(2000) <sup>271</sup>	Alcohol consumption in women
Moore et al.(2000) <sup>272</sup>	Back pain
Paton-Simpson et al.(2000) <sup>273</sup>	Hazardous drinking
Richmond et al.(2000) <sup>274</sup>	Hazardous drinking
Black et al.(2001) <sup>275</sup>	Infant feeding
Blondell et al.(2001) <sup>276</sup>	Hazardous drinking
Peleg et al.(2001) <sup>277</sup>	Adolescent drinking



of encouraging screening are needed.

Previous studies have tested the use of mail, computer, and telephone interventions to prompt women to screen<sup>254,278–280,288</sup>, and the use of radio broadcasts<sup>289</sup>. Television campaigns have also been shown to increase cervical screening rates<sup>254</sup>. The use of computers has been shown to be acceptable to women when delivering brief advice about cervical screening interventions<sup>288</sup>.

Brief advice, particularly where this advice is tailored to individual screening status and addresses known barriers to screening, has also been shown to be effective in increasing cervical screening rates<sup>254,278–280,285,288</sup>.

#### 6.1.4 Use of IVR in brief interventions

It was thought that IVR technology would have promise as a brief advice tool for health promotion, and cervical screening in particular. There were several reasons to believe this. It was thought that IVR may be able to overcome potential barriers, handle sensitive issues, provide access to information for those persons difficult to reach at home, avoid literacy and language problems, and be economical to run. These particular issues are discussed below.

**Barriers** Many potential barriers to having Pap tests have been identified, including embarrassment, lack of knowledge, discomfort, and preference for female doctors<sup>290</sup>. It was thought that information about barriers such as these could be systematically addressed by an IVR system by targeting specific messages to individual women.

**Sensitivity** It was known that some women considered Pap tests and cervical screening to be sensitive and private topics<sup>290</sup>. The previous work in this

thesis had found that survey respondents were more likely to report sensitive behaviours in an IVR survey than to a human-mediated interview. It was thought that IVR could enhance a woman's privacy, since the call was taken from home and only they could hear the questions.

**Accessing women** IVR could be used to contact anyone with a touchphone.

It could also contact individuals on a regular systematic basis. Once contacted, it could deliver a message targeted at that particular individual. Women who frequently moved could still be accessed by an IVR system that regularly sweeps through telephone number ranges.

**Automatic reminders** Women could arrange for an IVR system to contact them again automatically some time later. An IVR system could also be linked to a Pap test database and issue reminders in the same manner as postal reminders are currently produced by the NSW Pap Test Register (PTR).

**Literacy and language problems** IVR could address the literacy difficulties of paper based approaches and handle non-English speakers<sup>5-7,31,291-294</sup>.

**Cost** The cost of contacting each woman using IVR consists only of development costs and call charges. This means that IVR could be an extremely economical method of providing an individual intervention for each woman.

There had been no previous studies that had used IVR to directly contact a household sample and administer brief advice on cervical screening or any other health topic. Therefore, a trial was planned to evaluate the use of IVR in delivering a brief intervention for health promotion.

As a first step, a pilot was needed to examine the likely contact rate, whether women would respond to questions in order to determine their health status, the degree to which women would use IVR technology for obtaining health information, its acceptability, and its cost.

## 6.2 Aims

The aims of this chapter are to:

1. Describe the participation rates, (the proportion of those contacted who listen to the IVR message);
2. Compare the demographic profile of participants to the Australian Bureau of Statistics (ABS) Census profile;
3. Describe the IVR self-report cervical screening rate;
4. Assess how screened and unscreened women used the IVR system;
5. Compare the IVR and Computer-Assisted Telephone Interview (CATI) follow-up self-report screening rates;
6. Describe the acceptability of the IVR brief intervention; and
7. Compare the cost of the IVR brief intervention to a standard mail reminder.

## 6.3 Methods

### 6.3.1 Design

An IVR system was used to contact a random selection of 5000 households in the Cessnock Statistical Local Area (SLA) in September 2000 and deliver an IVR brief intervention promoting cervical screening. A CATI follow-up was used to assess the IVR brief intervention.

### 6.3.2 Setting

The Cessnock Statistical Local Area (SLA) (Statistical Local Areas (SLAs) are defined in Section 4.5.3 of Chapter 4) is located within the Hunter region, which is considered herein to be the area encompassed by the SLAs: Cessnock, Dungog, Lake Macquarie, Maitland, Merriwa, Murrurundi, Muswellbrook, Newcastle, Scone, and Singleton. Cessnock SLA has a population of 44,362, which is the forty-third largest of the 189 SLAs in NSW and fifth largest in the Hunter region<sup>208</sup>.

The population of Cessnock is 90% Australian-born, with a small proportion born overseas: United Kingdom, 3%; New Zealand, 1%; Germany, 0.5%; Netherlands, 0.2%; other countries or unknown, 5.3%<sup>208</sup>. Only 1.6% of the population aged five or more speak a language other than English at home<sup>208</sup>.

Cessnock SLA has 22,421 female residents of which 13,858 are aged 18–69<sup>208</sup>. Cessnock SLA has an Accessibility/Remoteness Index of Australia (ARIA) (The ARIA is defined in Section 4.6.3 of Chapter 4) score of 0.9945, which measures access to resources along a road network<sup>212</sup>. This is a low score indicating that Cessnock has “relatively unrestricted accessibility to a wide range of goods and

services and opportunities for social interaction” (see Table 4.12 of Chapter 4). Only 30% of SLAs in NSW have a lower ARIA score than this indicating greater access to resources.

Despite the adequate access to resources Cessnock SLA has consistently maintained a low cervical screening rate, and is the lowest-screening SLA in the Hunter region<sup>252</sup>.

### 6.3.3 Sample

A total of 5000 households with fixed telephone connections were selected from an electronic version of the White Pages<sup>215</sup> using simple random sampling. The sampling program is shown in Appendix D.1.1. The ABS Regional Localities Index<sup>295</sup> was used to ensure the street addresses of selected households were within the Cessnock SLA. Following the completion of the IVR brief advice calls, the program shown in Appendix D.1.2 was used to generate a sample that consisted of those who had listened to the IVR call and an equal number of those who had hung up on the IVR call. This sample was used for the follow-up CATI.

### 6.3.4 Procedure

Approval was obtained for this study from the University of Newcastle Human Research Ethics Committee (HREC) and Hunter Area Research Ethics Committee (HAREC).

**Participant instruction**

Information letters printed on letterhead stationery and addressed to ‘The Household’ were posted using DL-sized\* envelopes with the institution details (Hunter Centre for Health Advancement, Hunter Health, and the The University of Newcastle, Australia) plainly marked, one week before the first call. The letters (Appendix Section D.2) were in English since only a very small proportion of the Cessnock population speak another language at home<sup>208</sup>. The letters explained the project aims, stated that each household would receive an IVR call about cervical screening, and that participation would be voluntary. No remuneration was offered.

**Participant recruitment**

The Generalized Electronic Interviewing System (GEIS) running on the Recorded Voice System (RVS) (The RVS is defined in Section 3.4.2 of Chapter 3) initiated calls to households. When a call was answered, GEIS played an introductory message and then asked the person to press key ‘5’, but any key sufficed. If no key press was detected after three repetitions the system thanked the person and the call terminated. Otherwise, GEIS described the survey purpose before asking if there were any women aged 18–69 years old in the household. Since there may have been more than one eligible woman in the household, it asked to speak to the woman with the most recent birthday<sup>209</sup>. If the current time was not convenient or the woman was unavailable GEIS explained it could call back the next day at the same time or another at the woman’s convenience. Self-identified businesses were recorded as out of scope. GEIS distinguished an-

---

\*Envelope size 220 x 110mm.

swering machines from genuine individuals by the length of the salutation

“Hello?”

“Hello. We’re not in at present. . .”

### **Cervical screening barriers**

To identify known barriers to cervical screening that would be addressed by the IVR brief intervention, the Medline<sup>15</sup> database was searched for studies between 1998–2001 using the search terms ‘cervical screening barriers’. Principal cervical screening barriers identified in these studies were:

- Cost<sup>296–301</sup>;
- Cultural/Fatalism<sup>297,302–306</sup>;
- Being disabled<sup>307</sup>;
- Discomfort/Fear of pain<sup>305,306,308–311</sup>;
- Embarrassment<sup>296,296,299,304–306,308,310–314</sup>;
- Fear of finding cancer<sup>311,315</sup>;
- Lack of knowledge of cancer<sup>296,300,314,316</sup>;
- Lack of knowledge of Pap tests<sup>298,299,301–304,309,311,312,312,314</sup>;
- Language<sup>303,312</sup>;
- Limited access to medical services<sup>299,302,303,308,309,312,312–314</sup>; and
- Lack of time<sup>311</sup>.

Not all of these barriers were relevant to the current study. The barriers *cultural factors/fatalism* and *language* were not relevant because the Hunter area had a relatively homogeneous Australian-born English-speaking population<sup>208</sup>. The barrier *disabled* was also not considered because the proportion of women in NSW aged 15–70 years old with moderate to severe disabilities and living in private dwellings was 2.7%, which was too small for a general study<sup>317</sup>. The remaining barriers were considered relevant to this study: the barriers *cost*, *discomfort/fear of pain*, *embarrassment*, *fear of finding cancer*, and *lack of time* were addressed by means of targeted messages, while *lack of knowledge of cancer*, *lack of knowledge of test*, and *limited access to medical services*, were handled by optional information messages.

### IVR intervention

The IVR intervention was applied as part of a community based intensive awareness campaign in Cessnock between July–October 2000. The implementation of the IVR intervention was designed in consultation with the Hunter Cervical Screening Coordinator (HCSC) who managed the Cessnock community awareness campaign. Apart from IVR, the community awareness campaign<sup>318</sup> involved: (a) contacting key leaders from the SLA to encourage them to promote the campaign; (b) informing general practitioners in the area of the campaign; (c) deploying an extra women’s health nurse; (d) distributing pamphlets and posters about cervical screening at community access points; (e) inserting information leaflets in council rates notices; (f) inserting articles in school newsletters; and (g) announcements on local radio and print media.

The IVR intervention consisted of calls made by GEIS to selected households.



GEIS conducted the following sequence of operations:

1. Explained the nature of the call;
2. Identified if women aged 18–69 years old were present;
3. Selected one eligible woman at random using the last birthday method<sup>209</sup>;
4. Determined her screening status;
5. Delivered a message that either congratulated her on being correctly screened, a message of encouragement if she was under-screened, or another message appropriate to her status;
6. Offered to deliver additional messages designed to counter common barriers to screening (*Cost, Discomfort/Fear of pain, Embarrassment, Fear of finding cancer, and Lack of time*);
7. Offered to provide additional information on cervical screening and cancer (*Lack of knowledge of cancer, Lack of knowledge of test, and Limited access to medical services*);
8. Offered to read out contact numbers and an Internet address where she could obtain more information;
9. Offered to arrange for the HCSC to ring her back if she still had questions; and
10. Offered to record a question the woman may wish answered by the HCSC when she was called back.

**Call scheduling**

To ensure high contact rates a similar scheduling system was used for the IVR brief advice call to that for the IVR method described in Chapter 5. Initial calls were made on weekday afternoons or evenings. Non-contacted numbers were re-attempted at alternating thirty minute and eighteen hour intervals.

**Follow-up CATI**

Following the community-based intervention, a CATI was conducted to assess the acceptability of the campaign components. The CATI was scheduled to occur one month after the end of the campaign and two months after the IVR calls. All households that had listened to the call and an equal number of randomly selected households which had hung up on the call, were contacted by the CATI interviewer. If the woman had not listened to the IVR call the interviewer offered to arrange for GEIS to ring her back.

**6.3.5 Apparatus****Equipment**

The equipment used was the same as in Chapter 5. Voice recordings were spoken in English by a female staff member.

**Software**

Both the IVR system and CATI follow-up were implemented using GEIS (see Chapter 3). Data were accumulated on a central computer using SAS/SHARE software<sup>191</sup>. With both the initial IVR call and the follow-up CATI, answers to questions were provided by selecting one of a set of options, entering a number

or date, or entering an open-ended response. Answers to open-ended questions were entered verbatim by the CATI interviewers, while GEIS allowed women to record a short spoken question. The recorded question was stored as a file on the computer's hard disk and also sent attached to an e-mail to the HCSC. For numeric answers, absolute and reasonable limits prevented range errors. Invalid responses triggered an appropriate message or display. After three repetitions of an item, a non-response caused GEIS to hang up and record a refusal. Women could refuse to answer a particular item, and if they wished they could return to earlier questions and modify their answers.

### 6.3.6 Measures

#### Scripts

Two scripts were used, one to handle the IVR call (Supplementary Materials Section 6.1), and the other for the CATI follow-up survey (Supplementary Materials Section 6.2). The domains within the IVR script are discussed in the next section. Some technical information on the script items is presented in Appendix Section D.3.

The IVR script handled all aspects of the cervical screening messages, including contacting the household, identifying an eligible woman, delivering the messages, and providing information, and offering to arrange for a call-back. Towards the end of the IVR script, the woman was asked if she wanted someone to ring her back. If she chose this option, she could also record a question beforehand. The expert was the Hunter Cervical Screening Coordinator (HCSC) responsible for cervical screening promotion in the Hunter region.

The CATI follow-up script assessed awareness of the cervical screening cam-

paign, knowledge of cervical cancer and screening, reasons for listening or not listening to the IVR call, and the woman's screening status and demographic characteristics. If the woman had not listened to the call she was offered another opportunity to try it. If she had listened to it, the acceptability of the IVR call was assessed. She was also asked for suggestions for additional IVR health promotion ideas.

### **Domains**

The IVR script (Supplementary Materials Section 6.1) contained domains concerned with (a) Pap status, (b) barrier messages, (c) demography, and (d) information messages. These are discussed below.

**Pap status domain** In the Pap status domain, the woman's cervical screening status was determined using the method illustrated in Figure 6.1. A message appropriate to the woman's screening status was played as shown in Table 6.2.

**Barrier messages domain** For under-screened women, a series of messages were offered that addressed previously identified barriers to screening as shown in Table 6.3. Women could listen to one or more messages, or could skip ahead.

**Demography domain** The questions of the demography domain (education, marital status, country of birth, aboriginality, employment status, and income) were selected to be compatible with the ABS Census data<sup>208</sup>.

**Information messages domain** In the information messages domain, messages covering additional barriers to screening were offered. Women could listen to one or more information messages or could skip ahead. Contact telephone

numbers were offered for the Australian Cancer Society, Royal Newcastle Hospital sexual health clinic, New South Wales Cancer Council, and FPA Health. Lastly, the system offered to arrange to have someone call the woman back.

### 6.3.7 Analysis

#### Participation rates

Women were counted as listening to a domain if they had reached the last item in that domain, or they had begun the domain but had validly skipped the remaining items not relevant to them. The participation rates were then tabulated for each completed domain of the IVR call using Equation 5.5 in Section 5.3.7 of Chapter 5.

#### Demography

Demographic variables (age, education, marital status, country of birth, aboriginality, employment, income) were tabulated and compared to 1996 ABS Census profiles<sup>208</sup> using chi-square goodness-of-fit tests<sup>225</sup>. Since the ABS Census estimates represent population level data, the ABS profiles were taken to be theoretical distributions.

Ages were classified into the groups *18–29*, *30–39*, *40–49*, *50–59*, *60–69*,  $\geq 70$ . Where participants refused to give their age or only gave their age as *<18*, *18–69*, or  $\geq 70$  years old, the data were not included in calculations involving age. Previous cervical screening research has found that Pap test rate estimates are only slightly affected by unknown age estimates (0.6%)<sup>253</sup>. Because of small category sample sizes, some levels of education and income were collapsed. For Education, the levels *no education*, *primary school*, and *secondary school* were

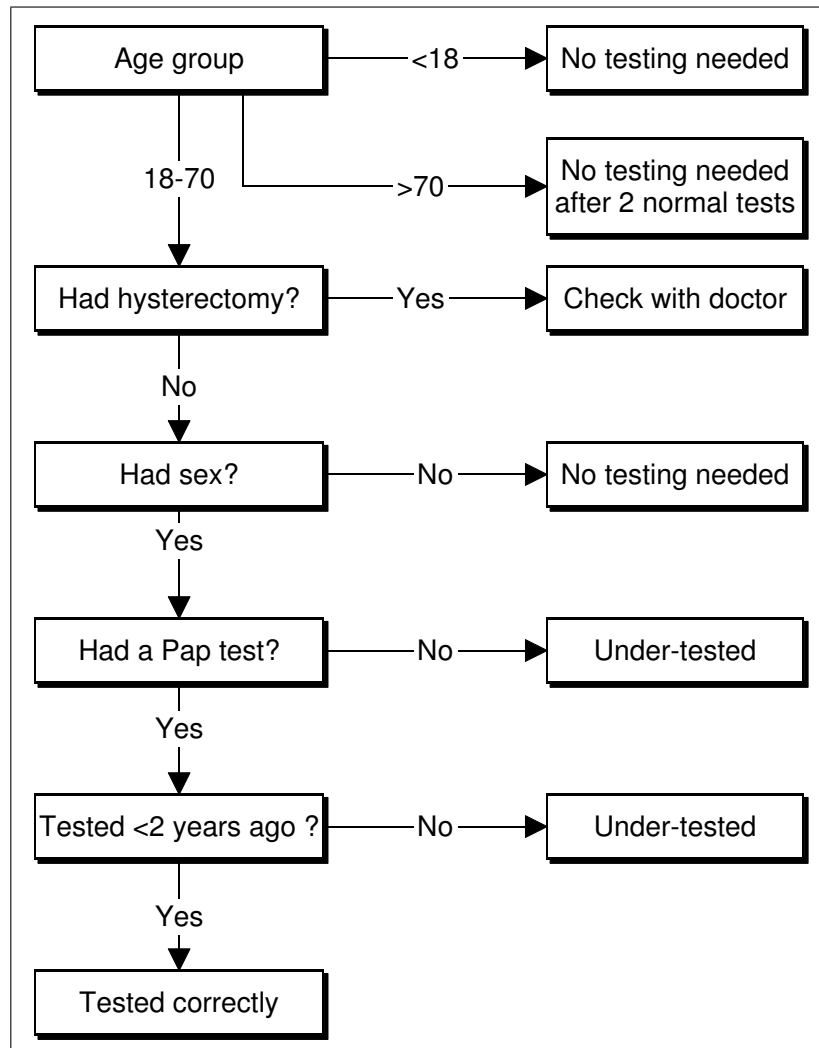


Figure 6.1: Determination of cervical screening status used within the Pap status domain.

Table 6.2: Pilot study messages played to women by cervical screening status.

<i>Status</i>	<i>Message</i>
Correctly screened	“Well done! Your Pap tests are up-to-date! All women aged eighteen to seventy years old who haven’t had a hysterectomy, and have <i>ever</i> had sexual intercourse, should have a Pap test every two years. Remember: a Pap test every two years can help prevent cervical cancer.”
Under-screened	“All women aged eighteen to seventy years old who haven’t had a hysterectomy, and have <i>ever</i> had sexual intercourse, should have a Pap test every two years. Please have a Pap test soon.”
Under eighteen	“All women should start having Pap tests at eighteen years of age, or two years after first having sex, whichever is later. Some women may need to start having Pap tests before eighteen years of age. It depends on your sexual activity. Ask your doctor for advice to see if you should start having regular Pap tests.”
Not had sex	“All women aged eighteen to seventy years old who haven’t had a hysterectomy, and have <i>ever</i> had sexual intercourse, should have a Pap test every two years. Since you’ve never had sexual intercourse, there’s no need for you to have a Pap test. Your doctor can give you more information about this.”
Had a hysterectomy	“Since you have had a hysterectomy, you should talk to your doctor to see if you need to have a Pap test or a similar test. Don’t assume you don’t need one: check with your doctor to be absolutely sure.”
Over seventy	“All women aged eighteen to seventy years old who haven’t had a hysterectomy, and have ever had sexual intercourse, should have a Pap test every two years. However, since you are over seventy you can stop having Pap tests if you have had two normal Pap tests within the last five years. Your doctor can tell you more.”

Table 6.3: Pilot study barrier messages available to under-screened women.

<i>Barrier</i>	<i>Message</i>
Discomfort/fear of pain	“Having a Pap test can feel uncomfortable but it should not hurt. Choose a doctor or nurse who you feel comfortable with. It’s important to have regular tests. Remember: a Pap test every two years can help prevent Cervical cancer.”
Embarrassment	“Being embarrassed is understandable, but most doctors and nurses are aware of this and will make sure you’re comfortable and you have enough privacy. It’s important that you find a doctor or nurse who you feel comfortable with. You can take a friend for support. It’s important to have regular Pap tests.”
Fear of finding cancer	“This is understandable, but a Pap test can give peace of mind. If Pap tests are done regularly, they can detect early changes that can easily be treated before cancer develops. Women who do have regular tests are less likely to develop cervical cancer. Remember: a Pap test every 2 years can help prevent cervical cancer.”
Lack of time	“Having enough time can be a very real problem. However it only takes a few minutes to do a Pap test. Most women visit their doctors at least once a year. Perhaps on one of these visits you could talk to your doctor about having a Pap test. A few minutes every 2 years can help save your life.”
Cost	“The cost of your Pap test can vary depending on where you have it done. Some doctors, clinics and health care centres provide free Pap tests. Others may charge but some may be refundable through Medicare.”



collapsed to *school or none*, and the levels of income were collapsed to less than \$20,000, and \$20,000 or greater.

### **CATI follow-up**

Chi-square tests<sup>226</sup> were used to compare the follow-up CATI data for participants and those who had hung up on the IVR call. Due to small sample sizes, it was necessary to further collapse levels of age, education, marital status, and income.

The kappa ( $\kappa$ ) statistic was used to obtain an estimate of agreement between the data obtained during the IVR call and the follow-up CATI. The  $\kappa$  statistic is a measure of agreement between different measurements on the same items<sup>226,230</sup>.

The self-reported cervical screening rate for the age groups *18–49* and *50–69* was compared to NSW PTR data<sup>252</sup> using chi-square goodness-of-fit tests<sup>225</sup>. Dependence of self-reported screening rates on demographic variables was examined using chi-square tests<sup>226</sup> or Cochran-Armitage trend tests<sup>231</sup>.

### **Barrier messages and information messages domains**

Those women who were under-screened according to the Pap status domain could listen to the barrier messages domain, or could skip ahead. The barrier messages domain was followed by the information messages domains. All women could listen to messages in the information messages domains, or could skip ahead. The proportions of those listening to domain messages were compared between the age groups *18–49* and *50–69* years old by chi-square tests<sup>226</sup>. Chi-square tests<sup>226</sup> were used to compare education, marital status, and income

between listening or not listening to domain messages.

### **Follow-up CATI self-report rates**

The self-reported screening rate obtained in the follow-up CATI was compared to NSW PTR data using chi-square goodness-of-fit tests<sup>225</sup>. The CATI self-report and IVR self-report screening statuses were compared using McNemar's test<sup>319</sup>. Logistic regression was used to model the CATI self-report screening against IVR self-report rate and number of days between the two contacts. The covariates marital status, employment, and income were also added singly.

### **Acceptability**

Chi-square tests were used to assess the association of the acceptability ratings with age, income, education, and employment.

### **Cost**

The cost was calculated by accumulating call cost over all connected calls and dividing by the number of participants.

## **6.4 Results**

The IVR system was active in September 2000. Calls were made within the hours 0830–2230 hours and numbers were called an average 2.7 times before contact.

### 6.4.1 Participation rates

#### IVR call

Of the 5000 numbers in the sample, 174 corresponded to disconnected lines, 30 were fax machines, and 85 were recorded as out-of-scope after receiving letter notifications. Of the remaining 4711 numbers, 216 were not rung after the information letters were returned, and 3 were omitted due to notification of the death of the recipient, leaving 4492 numbers.

As indicated in Table 6.4, a total of  $489 \div 4492 = 11\%$  households listened to the entire IVR call,  $661 \div 4492 = 15\%$  households and  $549 \div 4492 = 12\%$  women listened to at least the Pap status domain.

A total of  $106 \div 661 = 16\%$  of households chose to listen to at least one of the information messages, which were offered to all participants. Of the 106 households,  $49 \div 661 = 7.4\%$  listened to a single message,  $20 \div 661 = 3.0\%$  listened to two messages,  $9 \div 661 = 1.4\%$  listened to three messages, and  $28 \div 661 = 4.2\%$  listened to all five messages. A total of  $61 \div 661 = 9\%$  of households listened to the contact numbers, which were offered to all participants.

A total of  $16 \div 661 = 2\%$  of households requested that the Hunter Cervical Screening Coordinator ring them to provide additional information, and of these, seven recorded a spoken question in advance.

A total of  $489 \div 4492 = 11\%$  households and  $430 \div 4492 = 10\%$  of women listened to the entire call.

#### CATI followup

Those who had listened to the entire IVR call along with a random sample of 500 of those who had not completed listening were contacted in a follow-up

Table 6.4: Pilot study participation rates by the last domain completed.

<i>Last domain fully completed</i>	<i>Households listening<sup>a</sup></i>		<i>Women listening<sup>b</sup></i>	
	<i>n</i>	<i>%</i>	<i>n</i>	<i>%</i>
Pap status	661	15	549	12
Barrier messages	644	15	532	12
Demography	617	14	505	11
Information messages	505	11	444	10
All domains	489	11	430	10

<sup>a</sup> Number of households in which someone listened to the IVR call.<sup>b</sup> Number of women who listened to the IVR call.

CATI. Of the 988 households attempted, 587 (59%) completed the follow-up interview. The CATI follow-up response rate was significantly higher for those who had listened to the IVR than those who had not (87% versus 64%  $\chi^2=54.6$ ,  $df=1$ ,  $P<0.01$ ,  $n=758$ ).

### 6.4.2 Demography

As detailed in Table 6.5, the demographic composition of the IVR sample differed from the 1996 Census<sup>208</sup> on all variables except country of birth. Compared to the ABS data, the sample of those who listened to the IVR call under-represented younger and older women, over-represented women with post-school qualifications, who were married, of aboriginal descent, employed, and with lower incomes. Compared to the ABS data, the sample of those who did not listen to the IVR call under-represented younger and older women, and over-represented women who were married, and of aboriginal descent.

It was possible that in some cases individuals who listened to the IVR call may not have been the same people as those who completed the follow-up interview. To assess this the kappa ( $\kappa$ ) statistic was used as a measure of agreement

Table 6.5: Pilot study demographic composition of listeners to the IVR call from the IVR and CATI follow-up samples, and non-listeners in the CATI follow-up sample, compared to ABS census data.

<i>Variable</i>	<i>IVR</i>	<i>CATI Follow-up</i>		<i>ABS</i>
		<i>Listeners</i>	<i>Non-listeners</i>	
	%	%	%	%
<b>Age</b>				
18–29	12	8	13	21
30–39	24	24	24	21
40–49	26	27	26	20
50–59	21	24	18	13
60–69	16	13	17	11
≥70	2	4	3	15
P <sup>a</sup>	<0.01	<0.01	<0.01	
<b>Education</b>				
School <sup>b</sup>	64	74	81	85
Vocational	28	18	11	7
University	9	8	8	9
P	<0.01	<0.01	0.09	
<b>Marital status</b>				
Married	79	66	70	57
Never <sup>c</sup>	5	22	17	12
Div./Sep. <sup>d</sup>	9	6	7	14
Widowed	7	5	7	16
P	<0.01	<0.01	<0.01	
<b>Country of birth</b>				
Australia	92	91	93	90
P	0.16	0.46	0.17	
<b>Aboriginal/Torres Strait Islander</b>				
ATSI	4	2	3	0.3
P	<0.01	<0.01	<0.01	
<b>Employment</b>				
Employed	84	46	43	37
P	<0.01	<0.01	0.07	
<b>Income</b>				
<\$20,000	37	38	40	47
P	<0.01	<0.01	0.08	
<b>n</b> <sup>e</sup>	553	394	197	

<sup>a</sup> Chi-square test P-values comparing the sample profiles to the ABS Census data profile.

<sup>b</sup> School or none.

<sup>c</sup> Never married.

<sup>d</sup> Divorced or separated.

<sup>e</sup> Number of participants reporting demographic information.

Table 6.6: Pilot study agreement ( $\kappa$ ) between IVR and CATI demographic variables.

<i>Variable</i>	<i>Kappa</i>	<i>95% CI</i>
Age	0.98	{95, 100}
Education	0.66	{58, 73}
Marital status	0.58	{49, 66}
Employment	0.68	{60, 75}
Income	0.76	{69, 83}

between the IVR and CATI results. As a convention, values of the  $\kappa$  statistic in the range 0.41–0.60 may be regarded as indicating moderate agreement, 0.61–0.80 as substantial agreement, and 0.81–1.00 as almost perfect agreement<sup>226</sup>. As shown in Table 6.6, the agreement as measured by the kappa ( $\kappa$ ) statistic ranged from 0.58 to 0.98. These values ranged from moderate for marital status, and substantial to almost perfect for the remaining variables. This provided evidence that the number of cases where different individuals provided data in the IVR and CATI contacts was small enough that it was not likely to seriously affect the interpretation of the follow-up data.

### 6.4.3 IVR self-report screening status

Of the women who completed the Pap status domain, there were seven (1%) women aged over 70 years, 10 (2%) who had not had sexual intercourse, 114 (21%) who had had hysterectomies, 274 (50%) who were correctly screened, and 144 (26%) who were under-screened. For those who gave their age, 94 (68%) women aged 18–49, and 44 (59%) aged 50–69 were correctly-screened.

The proportion of correctly screened 18–49 year old women was significantly higher than NSW PTR data<sup>252</sup> (68% vs 56%;  $\chi^2=19.1$ ,  $df=1$ ,  $P<0.01$ ,  $n=296$ ). However, the proportion of correctly screened 50–69 year old women did not

differ from register data (59% vs 53%;  $\chi^2=1.7$ ,  $df=1$ ,  $P=0.19$ ,  $n=108$ ).

#### 6.4.4 Barrier messages

Those women who were under-screened according to the Pap status domain responses were presented with a set of barrier messages relating to discomfort/fear of pain, embarrassment, fear of finding cancer, lack of time, and cost. Of the women that completed the Pap status domain, 84 (59%) listened to barrier messages, as shown in Table 6.7. The most common barrier message selected was embarrassment, but there was a sizable proportion who did not select one.

The distribution of selected barrier messages did not differ between the 18–49, 50–69 and age-refused groups ( $\chi^2=9.2$ ,  $df=10$ , Monte Carlo  $P=0.52$ , 99% CI: {0.504, 0.530},  $n=143$ ). The distribution of selected barrier messages also did not differ among those who gave their age ( $\chi^2=3.5$ ,  $df=4$ , Monte Carlo  $P=0.49$ , 99% CI: {0.480, 0.506},  $n=131$ ). There was also no significant association between listening to any barrier message and education ( $\chi^2=1.4$ ,  $df=2$ , Monte Carlo  $P=0.30$ , 99% CI: {0.287, 0.310},  $n=127$ ), marital status ( $\chi^2=6.9$ ,  $df=3$ , Monte Carlo  $P=0.07$ , 99% CI: {0.060, 0.073},  $n=132$ ), employment (Fisher’s exact test,  $\chi^2=0.01$ ,  $P=1.00$ ,  $n=124$ ), and income (Cochran-Armitage trend test,  $T=+1.4$ , one-sided Monte Carlo  $P=0.09$ , 99% CI: {0.082, 0.097},  $n=66$ ).

#### 6.4.5 Information messages

Information messages were offered to all women. A total of 63 (13%) women listened to at least one information message. The proportion of eligible women who listened to an information message did not differ between those aged 18–49 and 50–69 (9% versus 15%; Fisher’s exact test,  $\chi^2=2.6$ ,  $P=0.11$ ,  $n=351$ ). For

Table 6.7: Pilot study barrier messages selected by age group of women.

<i>Barrier</i>	<i>Aged 18–49</i>	<i>Aged 50–69</i>	<i>Age refused</i>
	%	%	%
Cost	8	2	0
Discomfort/fear of pain	5	0	8
Embarrassment	28	21	33
Fear of finding cancer	5	2	0
Lack of time	18	21	25
None selected	36	53	33
<b>n<sup>a</sup></b>	88	43	12

<sup>a</sup> Number of women selecting or skipping barrier messages.

women aged 18–49, the proportions of under-screened and screened women who listened to information messages did not differ significantly (12% versus 8%; Fisher’s exact test,  $\chi^2=0.9$ ,  $P=0.33$ ,  $n=257$ ). However, for women aged 50–69, a significantly greater proportion of under-screened than screened women listened to the information messages (24% versus 8%; Fisher’s exact test,  $\chi^2=5.2$ ,  $P=0.02$ ,  $n=94$ ).

There was no significant association between listening to information messages and education (Fisher’s exact test,  $\chi^2=0.9$ ,  $P=0.39$ ,  $n=460$ ), marital status (Fisher’s exact test,  $\chi^2=0.01$ ,  $P=1.00$ ,  $n=472$ ), employment (Fisher’s exact test,  $\chi^2=0.02$ ,  $P=0.89$ ,  $n=458$ ), or income (Fisher’s exact test,  $\chi^2=0.1$ ,  $P=0.85$ ,  $n=251$ ).

A total of 57 (11%) women listened to the list of contact numbers.

There was no significant difference in the proportion of older and younger women listening to the contact numbers ( $\chi^2=2.9$ ,  $df=1$ ,  $P=0.09$ ,  $n=455$ ). However, a significantly larger proportion of under-screened than screened women listened to the contact numbers in the 18–49 age group (12% versus 2%; Fisher’s



exact test,  $\chi^2=9.8$ ,  $P<0.01$ ,  $n=257$ ) and also the 50–69 age group (24% versus 2%; Fisher’s exact test,  $\chi^2=11.3$ ,  $P<0.01$ ,  $n=94$ ).

However, no other association with demographic characteristics was found. No significant association was found between listening to the contact numbers and education ( $\chi^2=0.2$ ,  $df=1$ ,  $P=0.68$ ,  $n=460$ ), marital status ( $\chi^2=0.24$ ,  $df=1$ ,  $P=0.63$ ,  $n=472$ ), employment ( $\chi^2=0.11$ ,  $df=1$ ,  $P=0.74$ ,  $n=458$ ), and income ( $\chi^2=0.42$ ,  $df=1$ ,  $P=0.52$ ,  $n=251$ ).

A total of sixteen women recorded a request for the Hunter Cervical Screening Coordinator to ring them back and seven recorded a question to be answered by the Hunter Cervical Screening Coordinator. As detailed in Figure 6.2, the recorded questions were from women who either (a) had had hysterectomies, (b) were over 70 and who were unclear about their need for a Pap test, or (c) were inquiring about local access to Pap test facilities.

#### 6.4.6 IVR and follow-up CATI self-report screening status

During the CATI follow-up, women were asked about their screening status. This was then used to examine if the proportion of screened and under-screened women who had listened to the IVR message had differed. However, no significant difference was found (64% versus 69%;  $\chi^2=0.9$ ,  $df=1$ ,  $P=0.34$ ,  $n=465$ ).

Significant associations were found for the screening status obtained during the IVR and CATI interviews both for the 18–49 age group (McNemar’s test,  $S=26.0$ ,  $P<0.01$ ,  $n=201$ ) and also the 50–69 age group (McNemar’s test,  $S=6.2$ ,  $P=0.01$ ,  $n=73$ ).

For 18–49 year old women, the self-report screening rate was 85% using CATI and 72% using IVR, while for 50–69 year old women the rates were 73% and 60%,

1. "I'm wondering for a person aged seventy-two. Is it relevant to them?"
2. "I'm forty-three, female, had a partial hysterectomy, [but] still have ovaries. [My] mother recently passed away [due to] ovarian cancer. I don't have a cervix and I'm curious, as because doctors seem to be unsure of whether I should or shouldn't [have a Pap test]. They can't agree whether I should or should not have Pap smears. And the fact that my mother just passed away recently after just finding out she had ovarian cancer. So, if someone wants to call me back and talk to me about that would be nice."
3. "I'd like to know where I can go and see a female doctor to have a Pap smear test in my area. I live near Cessnock."
4. "Yes, I have had a hysterectomy and, yes, I was told in your instructions to go and see my own doctor. Is it necessary for me to have a Pap smear or not? Can you answer that question for me?"
5. "I had a hysterectomy eighteen years ago. I don't have a Pap test every two years. I've got to have one as often as I remember. Is that OK?"
6. "I would like to know if I do have to have a Pap smear. I have had a hysterectomy."
7. "Yes, I just like to know — I've been to a couple of doctors. I've a hysterectomy and I'd just like to know. Some say I do need a Pap smear. Some say I don't. So I'd just like to know whether I do or not 'cause I haven't had one for a few years now."

Figure 6.2: Pilot study texts of recorded questions.

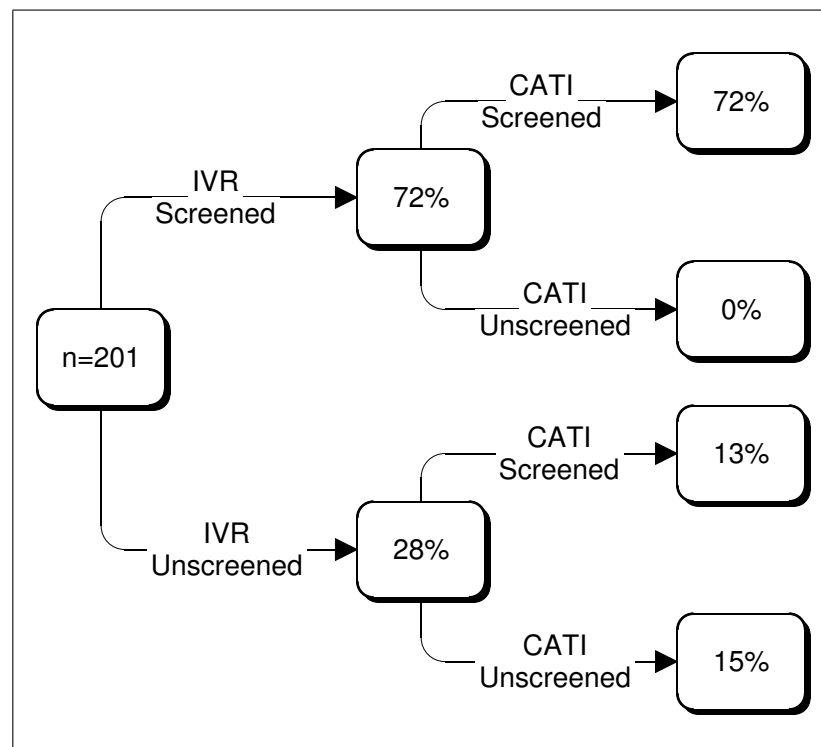


Figure 6.3: Pilot study self-report screening status for successive interviews of 18–49 aged women from IVR and CATI follow-up.

respectively. As shown in Figures 6.3 and 6.4, almost all of those who reported being correctly-screened in the IVR call later reported being correctly screened in the follow-up CATI call. However, 46% of under-screened women aged 18–49, and 38% of under-screened women aged 50–69, subsequently reported being correctly-screened in the CATI follow-up.

The extra screenings reported in the CATI followup interviews may have been due to normal screening behaviour in the two-month period between the IVR call and the CATI. However, during this period the proportion of screenings expected for Cessnock in this period (6%<sup>252</sup>) was less than that observed.

To further investigate this, logistic regression was used to model the CATI self-report status as a function of the IVR self-report status, the age group, and

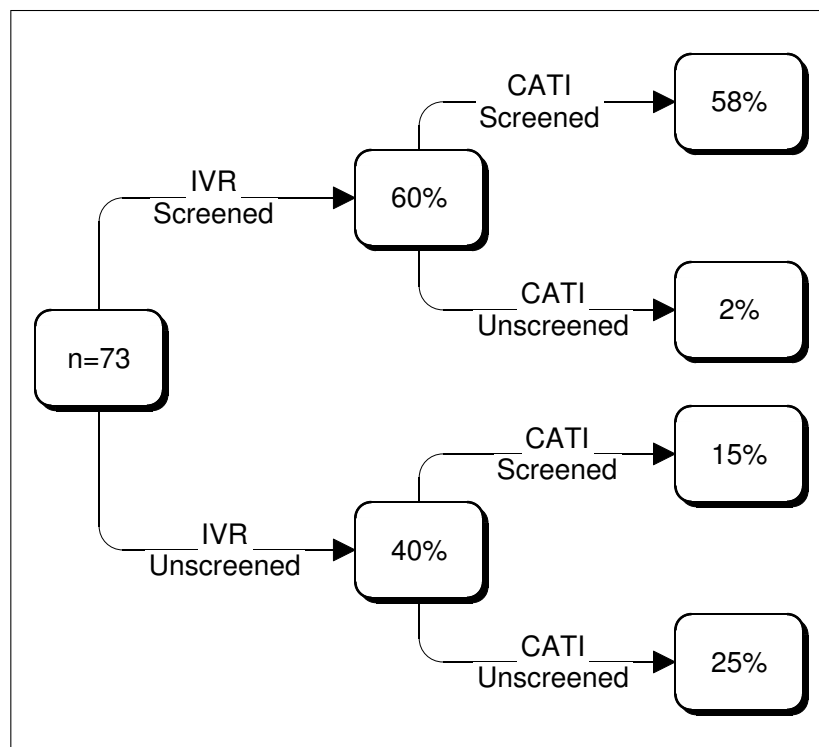


Figure 6.4: Pilot study self-report screening status for successive interviews for 50-69 aged women from IVR and CATI follow-up.

Table 6.8: Pilot study acceptability of the IVR call.

<i>Measure/Level</i>	<i>%</i>
<b>Ease</b>	
Very easy	53
Fairly easy	45
Not sure	1
Fairly hard	1
Very hard	0
<b>Enjoyable</b>	
Very enjoyable	2
Enjoyable	24
A little enjoyable	56
Not enjoyable at all	18
<b>Stressful</b>	
Not stressful	88
A little stressful	10
Stressful	2
Very stressful	0
<b>Likeable</b>	
Like it a lot	9
Like it a little	37
Neither like nor dislike it	48
Dislike it	5
Dislike it a lot	1
<b>n</b>	<b>378</b>

the time difference in days between the two reports. After adjusting for age group, the odds of self-report as correctly screened using CATI was 7.2 times (95% CI: {4.4, 11.6}) if a woman had also reported being correctly screened using IVR. However, the odds did not change significantly as the time difference increased ( $\chi^2=0.6$ ,  $df=1$ ,  $P=0.44$ ,  $n=272$ ), indicating that the screened proportion was not increasing with time.

#### 6.4.7 Acceptability

The acceptability of the IVR call is detailed in Table 6.8.

**Ease**

The majority of women (98%) thought that participating in the IVR call was fairly easy or very easy, but there was a significant association with age ( $\chi^2=52.9$ ,  $df=20$ ,  $P<0.01$ ,  $n=372$ ) with older women finding it less easy than younger ones (Kendall's  $\tau_b=-0.16$ ). A significant association was also found with education ( $\chi^2=24.0$ ,  $df=6$ , Monte Carlo  $P<0.01$ , 99% CI: {0.002, 0.005},  $n=374$ ) with more highly educated women finding it easier ( $\tau_b=0.12$ ). No association was found with income ( $\chi^2=18.2$ ,  $df=16$ , Monte Carlo  $P=0.27$ , 99% CI: {0.257, 0.280},  $n=230$ ) or employment ( $\chi^2=6.8$ ,  $df=3$ , Monte Carlo  $P=0.11$ , 99% CI: {0.108, 0.116},  $n=166$ ).

**Enjoyable**

A total of 82% found the IVR call enjoyable. No association was found with age ( $\chi^2=16.8$ ,  $df=15$ , Monte Carlo  $P=0.32$ , 99% CI: {0.306, 0.330},  $n=372$ ), education ( $\chi^2=6.8$ ,  $df=9$ , Monte Carlo  $P=0.63$ , 99% CI: {0.619, 0.644},  $n=374$ ), income ( $\chi^2=13.3$ ,  $df=12$ , Monte Carlo  $P=0.34$ , 99% CI: {0.329, 0.354},  $n=230$ ), or employment ( $\chi^2=2.2$ ,  $df=3$ , Monte Carlo  $P=0.50$ , 99% CI: {0.487, 0.513},  $n=166$ ).

**Stressful**

A total of 88% did not find the IVR call stressful. No association was found with age ( $\chi^2=14.3$ ,  $df=15$ , Monte Carlo  $P=0.44$ , 99% CI: {0.431, 0.457},  $n=375$ ), education ( $\chi^2=11.3$ ,  $df=9$ , Monte Carlo  $P=0.23$ , 99% CI: {0.215, 0.236},  $n=374$ ), or income ( $\chi^2=1.4$ ,  $df=4$ , Monte Carlo  $P=0.85$ , 99% CI: {0.836, 0.856},  $n=228$ ).

Table 6.9: Pilot study preference for alternative IVR call topics.

<i>Topic</i>	<i>n</i>	<i>%</i>
Losing weight	192	58
Nutrition	193	58
Physical activity	138	42
Illegal drugs	101	30
Falls & elderly	87	26
Smoking	77	23
Alcohol	51	15

### Likeable

Almost half (46%) liked the IVR call. No association was found with age ( $\chi^2=17.4$ ,  $df=20$ , Monte Carlo  $P=0.61$ , 99% CI: {0.593, 0.618},  $n=375$ ), education ( $\chi^2=9.7$ ,  $df=12$ , Monte Carlo  $P=0.56$ , 99% CI: {0.556, 0.581},  $n=374$ ), income ( $\chi^2=10.6$ ,  $df=16$ , Monte Carlo  $P=0.88$ , 99% CI: {0.883, 0.899},  $n=230$ ), or employment ( $\chi^2=4.9$ ,  $df=3$ , Monte Carlo  $P=0.20$ , 99% CI: {0.191, 0.212},  $n=166$ ).

### IVR health topic calls

A total of 90% agreed that IVR Pap call reminders were a good idea. No association was found with age ( $\chi^2=22.9$ ,  $df=20$ , Monte Carlo  $P=0.30$ , 99% CI: {0.291, 0.316},  $n=581$ ), income ( $\chi^2=19.9$ ,  $df=20$ , Monte Carlo  $P=0.47$ , 99% CI: {0.457, 0.483},  $n=472$ ), education ( $\chi^2=8.4$ ,  $df=12$ , Monte Carlo  $P=0.69$ , 99% CI: {0.683, 0.706},  $n=579$ ), or employment ( $\chi^2=35.5$ ,  $df=24$ , Monte Carlo  $P=0.06$ , 99% CI: {0.054, 0.066},  $n=472$ ).

As detailed in Table 6.9, when asked which of a set of health topic areas they would like to hear in an IVR call, the majority favoured ‘losing weight’ and ‘nutrition and diet’. Additional suggestions are shown in Table 6.10.

Table 6.10: Pilot study alternative health topics for IVR calls nominated by participants.

<i>Topic</i>	<i>n</i>	<i>Topic</i>	<i>n</i>
Breast cancer	27	Hepatitis	2
Diabetes	20	Illegal drug use	2
Children	18	Kidney disease	2
Cancer	16	Pregnancy	2
Mental health	13	Prostate cancer	2
Heart disease	9	Aging	1
Arthritis	7	Blood disease	1
Asthma	7	Convulsions	1
Blood pressure	7	Domestic violence	1
Adolescents	6	Eye disease	1
Menopause	6	Gambling	1
Nutrition and diet	6	Heart cancer	1
Osteoporosis	6	Lung Cancer	1
Stress	6	Migraine	1
Depression	4	Ovarian cancer	1
Immunization	4	Pain	1
Alzheimer's disease	3	Pap tests	1
Cervical cancer	3	Physical activity	1
Cholesterol	3	Quit smoking	1
Dementia	3	Reducing alcohol	1
Sexually transmitted diseases	3	Sex education	1
Bowel cancer	2	Skin cancer	1
Elderly	2	Speech pathology	1
General health	2	Women's health	1



### 6.4.8 Cost

The cost of hardware and software development was disregarded as being a capital expenditure not normally associated with a single project. The cost of printing, packaging, and posting the information letters was also disregarded as letters would not normally be used in an IVR intervention. The remaining costs were call charges and the salary of a health professional in cases where women wished to be rung back. The call charge (including Subscriber Trunk Dialling (STD) rates) per participant was  $\$2178 \div 661 = \$3.30$ .

To obtain the cost of an equivalent mail-pamphlet project, data from Newell<sup>320</sup> was used. Newell posted personal health record diaries to households. The diaries were accompanied by booklets that provided information about the prevalence of cancer and cardiovascular disease, recommendations of preventative, screening and detection behaviour, and further contact information. The diaries were meant for keeping record of an individual's screening history. The cost of printing, packaging, and posting was \$1.96 per household. Newell found that only 4.1% diaries were used, giving a cost per completed diary of  $\$1.96/4.1\% = \$47.80$ .

## 6.5 Discussion

### 6.5.1 Participation rates

#### IVR call

The proportion of women who listened to at least part of the IVR call was 15%. However, only 11% listened to the entire call. These outcome rates appear consistent with the response rates obtained by the IVR method in Chapter 5 and

suggest that cold-calling using IVR does not allow large proportions of a population to be accessed. This may be of less importance with an IVR intervention than a survey since the objective is not to obtain a representative sample, but to economically contact large numbers of individuals, and to preferentially contact individuals to whom the IVR call is most relevant. Even 15% of a population may correspond to a large number of individuals.

Many individuals dropped out progressively through the IVR call. These dropouts may have been due to misleading messages in the script, such as those beginning with

“Thank you . . .”

“That is the end of our questions . . .”

Some women may have misinterpreted these to indicate that the end of the IVR call had been reached and hung up without listening further.

Only a small number of women requested that someone ring them back and half of these recorded a question beforehand. This may have been because either the IVR call had already provided sufficient information, that the topic was sensitive, or that they did not want to talk to another person. The call-back facility would appear useful, although the proportion using it was small.

### **CATI follow-up**

The follow-up CATI was not just concerned with the IVR call but also considered other aspects of the multi-channel intervention in Cessnock, so the introduction spoken by the interviewers in the follow-up CATI did not mention the IVR call. However, the response rate for the follow-up was substantially higher for those who had listened to the IVR call than those who had hung up on

the IVR call. It was concluded that either listening to the IVR call increased participation in the follow-up, or that those who listened to the IVR call were people who were likely to participate in any survey.

### 6.5.2 Demography

The demographic profile of those who listened to the IVR call differed from ABS Census data. This difference was also seen in the CATI follow-up among those who had listened to the IVR call. Those who did not listen differed from the Census on fewer variables than the listeners, perhaps because the listeners may have represented a self-selected random sample. However, the listeners included fewer young or old women, more educated women, more married women, more aboriginal women, and more women with higher incomes.

### 6.5.3 IVR self-report screening status

The screening status of women in the 18–49 age group exceeded the population-level rate. The rate for women aged 50–69 agreed with population level rates, which may be due to the study having an insufficient power to detect any difference. However, the sample size had been sufficient to detect an over-report of 9% or higher.

### 6.5.4 Barrier messages

Previously identified barriers to screening (see Section 6.3.4) were addressed by messages: cost, discomfort/fear of pain, embarrassment, fear of finding cancer, and lack of time. Women who were under-screened were asked to listen to an appropriate barrier message. The most common barrier message selected was

embarrassment, which was consistent with other studies<sup>296,299,304–306,308,310–314</sup>.

No association was found between choice of barrier message and age, or between listening to any barrier message and demographic variables (education, marital status). This may have resulted from a lack of statistical power. However, the effect sizes being tested were all small according to the criteria suggested by Cohen<sup>321</sup>. It was concluded that there was no evidence that selection of barrier message or listening to any message depended on the woman's background.

### 6.5.5 Information messages

Although information messages were offered to all women, only a small proportion chose to listen to them. More older than younger women listened although the difference was only close-to-significant. More under-screened than screened younger women listened, but the effect was non-significant. Since the detectable difference was small ( $\approx 8\%$ ) it was concluded that the differential was not useful. However, it was of considerable interest that among older women, those who were under-screened were significantly more likely to listen than those who were screened. This was of interest since the 50–69 age group was a recognized risk group who would ideally be preferentially targeted with cervical screening promotions.

There was no association between listening to an information message and other demographic variables (education, marital status, employment, and income). The differences were small according to the criteria suggested by Cohen<sup>321</sup>. It was concluded that there was no evidence that, apart from age, listening to information messages depended on the woman's background.

The number of women listening to contact numbers was small. It was of interest that under-screened women were more likely to listen than correctly-screened women. This differential was more accentuated for older women than younger women, possibly because the older women perceived a greater personal risk from cervical cancer.

As with the information messages, there was no association between listening to the contact numbers and other demographic variables. It was concluded that there was no evidence that listening to the contact numbers depended on the woman's background.

The recorded questions mostly addressed whether women who had had hysterectomies needed to continue testing. Since this requires a knowledge of the woman's medical history, any telephone contact cannot properly deal with this issue other than to advise the woman to seek further advice from a medical practitioner.

#### **6.5.6 IVR and follow-up CATI self-report screening status**

In both age groups, the overall self-report Pap test rate was 13% higher for CATI than IVR. However, only in younger women did the IVR Pap test rates exceed population estimates, while the CATI Pap test self-report rates for both groups were higher than population data.

Where women reported being correctly screened in the IVR call, they also tended to report being correctly screened in the CATI. However, more than a third of younger women and a fifth of older women who reported being under-screened in the IVR call subsequently reported being correctly screened in the CATI. This may have been the result of: (a) normal screening between the IVR

call and the CATI; (b) women reporting screening that had occurred earlier as being more recent ('telescoping')<sup>322</sup>; (c) increased screening as a direct result of the IVR call or other aspects of the intervention; (d) over-reporting in the CATI.

The first explanation may be disregarded since the number of screens to be expected in the period between the IVR and CATI calls would have been much smaller than that observed.

The second explanation was not tenable since there was no reason to expect telescoping to occur during the IVR contact but not during the CATI contact.

The third explanation was eliminated by the result of the logistic regression. The logistic regression indicated that the most important predictor of screening in the CATI was being screened in the IVR call. The time difference between the two events was not a significant factor indicating that more women were not being screened over time.

The remaining explanation was that women over-reported their screening status to the CATI interviewer. As shown in Table 6.11, other studies have also found over-reporting of self-reported cervical screening in comparison to medical records or Pap test registry data of comparable magnitudes to the over-reporting observed in this study (13%). It may be speculated that some of the over-reporting in these studies may also have resulted from socially-desirable reporting, although other causes such as telescoping<sup>322</sup> or lack of understanding the nature of a Pap test<sup>323</sup> may also have contributed.

A complicating factor in this work was that the method of assessing cervical screening self-report differed between the IVR call and follow-up CATI. Since the CATI approach was different to that used in the IVR call, the difference in

Table 6.11: Over-report of Pap tests by women reported in cervical screening studies.

<i>Study</i>	<i>Over-report(%)</i>	<i>Validation method</i>	<i>Screen period</i>
Walter et al.(1988) <sup>322</sup>	20	Medical records	5 years
Sawyer et al.(1989) <sup>324</sup>	21	Medical records	3 years
Bowman et al.(1991) <sup>325</sup>	23	Medical records	3 years
Michielutte et al.(1991) <sup>323</sup>	71	Provider reports	Concurrent
McKenna et al.(1992) <sup>326</sup>	46	Medical records	3 years
Gordon et al.(1993) <sup>327</sup>	22	Medical records	2 years
Bowman et al.(1995) <sup>278</sup>	4	Registry data	6 months
Kottke et al.(1995) <sup>328</sup>	25	Medical records	1 year
Montaño and Phillips(1995) <sup>329</sup>	9	Medical records	1 year
Suarez et al.(1995) <sup>330</sup>	39	Medical records	2 years
Bowman et al.(1997) <sup>331</sup>	39	Medical records	3 years
Campbell et al.(1997) <sup>288</sup>	20	Medical records	6 months
Hancock et al.(1998) <sup>247</sup>	10	Registry data	2 years
McGovern et al.(1998) <sup>332</sup>	34	Medical records	1 year
Mamoon et al.(2001) <sup>253</sup>	13	Registry data	2 years

estimate between the two contacts may be confounded with differences due to the question-wording.

### 6.5.7 Acceptability

The IVR call was acceptable with most women finding it easy, enjoyable, un-stressful, and almost half liking it. More educated women and younger women found it easier, but this could be the result of greater familiarity with technology in those groups. The remaining tests of association with demographic variables were non-significant despite having powers of at least 50%<sup>321</sup>. It was concluded that the IVR call was broadly acceptable.

The majority of women also agreed that IVR health-related calls were a good idea and favoured further topics being developed. There was also no indication of any association with agreeing that IVR health-related calls were a good idea and demographic status.

### 6.5.8 Cost

A comparable alternative to IVR calls was mailed pamphlets. However, the IVR call was not only less expensive than pamphlets per completed call, but also allowed accurate recording of the number of those who listened to the calls without expensive followup.

## 6.6 Conclusions

The IVR call was successfully used by a large number of women to obtain information on cervical cancer and screening. The proportion of those called who listened to the call was only 15%. However, the use of an IVR call meant that it was possible to enumerate all the contacts made, unlike in other methods such as radio broadcasts.

The demographic composition of the participants did not match the population. Instead, there was an excess of middle-aged, employed women with a post-school education.

The system obtained a lower degree of over-reporting of cervical screening status compared to other studies, which may have resulted from the confidentiality of the IVR method. Those women who were unscreened and the older age group used the system to a greater extent than screened women or younger women. This was pleasing since these are recognized risk groups towards whom a cervical screening health promotion intervention should ideally be directed. Lastly, use of the system did not depend on womens' backgrounds, and the IVR call proved acceptable and inexpensive.

The successful outcome of the pilot study suggested that the IVR method



might be able to increase screening rates since those most at risk were also the heaviest users of the system. However, the pilot was an unreplicated uncontrolled trial and no conclusion could be reached on the effect it may have had on screening rates. Accordingly, a controlled trial was planned that would evaluate its efficacy in increasing screening rates. This is described in Chapter 7.

## Chapter 7

# A brief intervention for cervical screening using Interactive Voice Response: A randomized controlled trial

### 7.1 Introduction

The results of the pilot study (Chapter 6) were encouraging. The purpose of the pilot was to investigate the acceptability of the Interactive Voice Response (IVR) message and the proportion of those called who would listen to it. The general lack of association with demographic characteristics implied a broad applicability of the IVR method across diverse population groups. The pilot study's results indicated that the proportion of those answering the call was about 15%, and women in higher risk groups were more likely to listen to the IVR message. However, this did not imply that at-risk women would be more likely to be subsequently screened. The proportion of under-screened women who listened to the IVR message and subsequently screened is defined herein as

the efficacy of the method. Since the pilot investigated acceptability it used an unreplicated design. Therefore the next step was to estimate the efficacy of the IVR method by conducting a trial using a replicated and controlled design.

## 7.2 Aims

The aims of this project were to:

1. Conduct a controlled and replicated trial of a brief advice IVR intervention to increase cervical screening by a nominal 1.0% over three months;
2. Assess the efficacy of a brief advice IVR intervention in promoting cervical screening;
3. Compare the change in screening rates obtained with a brief advice IVR intervention to other cervical screening interventions;
4. Determine the cost per additional screen using a brief advice IVR intervention; and
5. Compare the cost per additional screen with a brief advice IVR intervention to that of other cervical screening interventions.

## 7.3 Methods

### 7.3.1 Design

A brief advice IVR cervical screening intervention was delivered to 17,008 randomly selected households between April–July 2001 in fifteen Hunter postcodes and the results compared to another fifteen control postcodes.

Cessnock Statistical Local Area (SLA) (SLAs are defined in Section 4.5.3 of Chapter 4) was excluded from the study because it had already been the setting for a cervical screening intervention as described in Chapter 6.

### 7.3.2 Setting

The Hunter region is located in New South Wales (NSW) as shown in Figure 7.1. Excluding Cessnock Statistical Local Area (SLA) it has a population of 462,704<sup>208</sup>. The area is covered by the Statistical Local Areas (SLAs) shown in Figure 7.2.

The population is 88% Australian-born, with a small proportion born overseas: United Kingdom, 3.7%; New Zealand, 1.0%; Germany, 0.5%; Italy, 0.3%; Netherlands, 0.3%; Poland, 0.3%; Philippines, 0.2%; Macedonia, 0.2%; other countries or unknown, 5.0%<sup>208</sup>. The Hunter area has 233,756 female residents of which 148,757 are aged 18–69.

The Hunter area has an Accessibility/Remoteness Index of Australia (ARIA) (ARIA scores are defined in Section 4.6.3 of Chapter 4) score<sup>212</sup> of 1.348, indicating relatively unrestricted accessibility (see Table 4.12 in Chapter 4). Fifty-two percent of Area Health Services in NSW have an ARIA score indicating greater access to resources.

### 7.3.3 Sample

Prior to selecting a sample, the population and screening rate needed to be defined. Eligible women were defined as those aged 18–69 who had not had a hysterectomy. The target population was defined as all eligible women. The screening rate was defined as the number of women aged 18–69 in the population

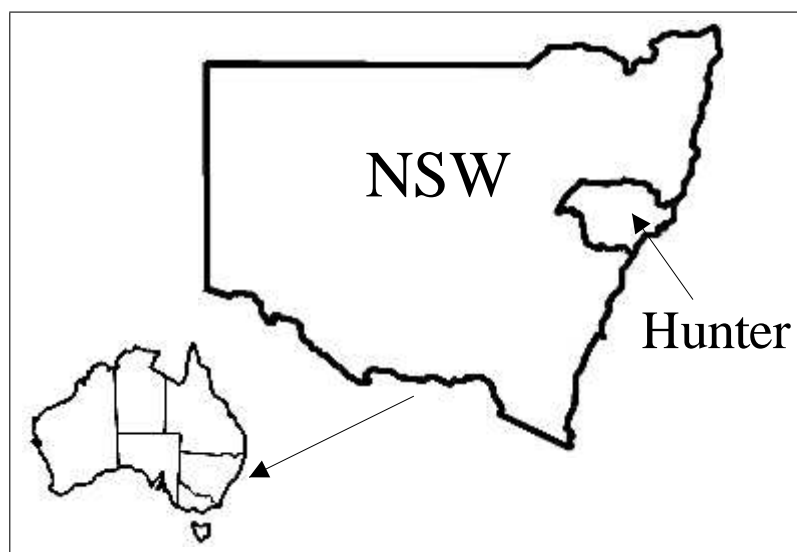


Figure 7.1: The Hunter region and NSW, Australia.

who had had a Pap test at least once in the reporting period divided by the size of the target population.

The sample needed to be of sufficient size to enable a large enough number of eligible women to be contacted. In selecting these women, small geographic sampling areas were preferred to larger ones. This was because smaller areas would have less variation in factors that may have been correlated with screening rates, such as the Index of Socio-Economic Disadvantage (SED)<sup>333</sup>. The smallest geographic areas for which cervical screening rates could be obtained were postcodes. Using smaller geographic areas than postcodes may have led to the possibility of identifying individual women from the screening rate statistics, which could theoretically have resulted in an infringement of confidentiality. This limitation meant that the intervention could be delivered to individual randomly selected postcodes, but not smaller geographic areas.

The intervention was to be delivered to households in one group of postcodes while another control group of postcodes would not receive any calls. It was

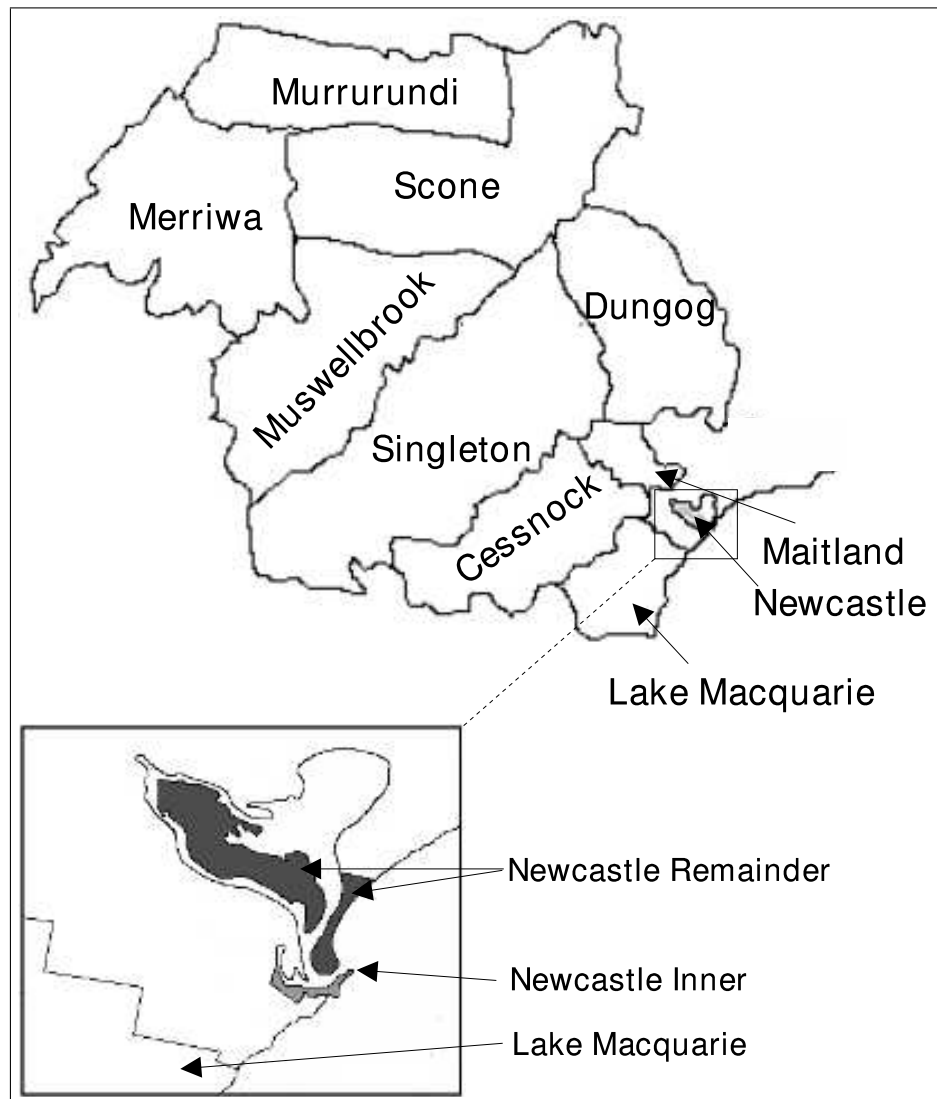


Figure 7.2: Statistical local areas of the Hunter region.

also noted that there was little risk of intervention contamination since randomly selected postcodes were unlikely to be contiguous or even geographically close.

### Selecting postcodes

SLAs were used as a means of selecting postcodes that spanned the Hunter area. Using the ABS National Localities Index (NLI)<sup>295</sup>, a list of 49 postcodes was constructed for localities found in the SLAs Dungog, Lake Macquarie, Maitland, Merriwa, Murrurundi, Muswellbrook, Newcastle - Inner, Newcastle - Remainder, Scone, Singleton, which, along with the Cessnock SLA, make up the administrative area of the Hunter Area Health Service (HAHS). Postcodes within the Cessnock SLA were not considered in this trial since it had already been the target of the cervical screening promotion by the Hunter Centre for Health Advancement (HCHA) as described in Chapter 6.

Postcodes that did not contain households (The University of Newcastle and Williamstown Air Force base), that only contained a small population (Newcastle West,  $n=18$ <sup>334</sup>), or which were the target of imminent cervical screening interventions (Lucy Bates, personal communication, 2000), were removed, leaving 30 postcodes.

Socioeconomic status was known to be correlated with cervical screening rates<sup>333</sup>. This association was used in assigning postcodes to control and intervention groups so that the two groups would be more likely to have similar screening rates than if the postcodes had been assigned purely at random. A preferable method of dividing the postcodes into the two groups would have been to use the postcode screening rate itself, but this data was unavailable. The Index of Socio-Economic Disadvantage (SED) for postcodes was obtained

from the Australian Bureau of Statistics (ABS) Socio-Economic Indexes for Areas (SEIFA) data set<sup>335</sup>. The SED is derived from attributes such as low income, low educational attainment, high unemployment and jobs in relatively unskilled occupations<sup>336</sup>. A high score on the SED occurs when a postcode has few families of low income and few people with little training and in unskilled occupations. A low score on the index occurs when a postcode has many low income families and people with little training and in unskilled occupations<sup>336</sup>.

Use of the SED index alone would have resulted in groups with very dissimilar population sizes, making the control and intervention groups less comparable. Therefore, the number of women aged 18–69 in each postcode (NWPC) was also used as an aid in forming the two groups. The NWPC was obtained from ABS Census profile data<sup>208</sup>.

Because two variables, NWPC and SED, were used to assign postcodes to the groups, the resulting groups were not likely to be exactly equal with respect to either NWPC or SED. However, it was intended that the groups would be approximately balanced with respect to both.

Principal Components Analysis (PCA) was used as an objective method to assign the postcodes to the groups using NWPC and SED. PCA is a means of replacing a data set by another equivalent data set that has reduced dimensionality<sup>337</sup>. This is done by calculating new variables, called *principal components*, that represent uncorrelated linear combinations of the eigenvalues of the original data set. The set of principal components summarize the same variance as the original data set. If the proportion of the total variance is explained by a small number of principal components then the dimensionality has been reduced to that number of dimensions with little loss of variance. PCA allowed the two



variables SED and NWPC to be replaced by a single variable.

Prior to the PCA, the two variables, SED and NWPC, were standardized to a mean of zero and a variance of one using Equation 7.1 to avoid the variable with the largest magnitude being unduly influential.

$$x'_i = \frac{(x_i - \bar{x})}{s_x} \quad (7.1)$$

where:  $x'_i$  = standardized value of  $x_i$ ,

$x_i$  = the original value of either SED or NWPC,

$\bar{x}$  = mean of  $x$  across all postcodes, and

$s_x$  = standard deviation of  $x$  across all postcodes.

From the PCA, the first principal component accounted for 59% of the variance in the data set, indicating that most of the variance was being explained by the first component. To assign postcodes to treatment and control groups, the postcodes were first sorted by the first principal component. Then, following a random start, alternate pairs were assigned to control and intervention groups.

### **Sample size**

Since population level data would be used rather than sample data, the sample size was selected so as to obtain a particular effect level rather than a statistically detectable amount. The effect to be measured was a change in the screening rate in intervention postcodes compared to control postcodes over the months following the intervention. To obtain a target effect of 1.0% it was necessary to contact a sufficient number of households so that enough women could be prompted to have a Pap test who would not otherwise have done so.

The selected postcodes (see Table 7.1) had a total population of 75,532

women aged 18–69<sup>208</sup>. The three monthly cervical screening rate for women aged 20–69 for the Hunter area was 8.9% of the eligible female population<sup>252</sup>. To obtain an increase of 1.0% in the three monthly rate in the intervention postcodes compared to the control postcodes, an additional  $75,532 \times 0.01 \div 2 = 378$  women would be needed to be screened in the intervention postcodes.

It was necessary to contact a sufficient number of households in the intervention group using an IVR message in order to prompt the additional 378 women to be screened. The number of households had to be large enough to allow for the:

1. Percentage of households with telephones;
2. Percentage of households that contained 18–69 year old women;
3. Percentage of women who would listen to the IVR message;
4. Percentage of women who would not have had a hysterectomy;
5. Percentage of women who would not have been screened within the last two years; and
6. Efficacy of the IVR intervention.

Three of these quantities were available from the pilot data (Chapter 6): (a) 90% of households had telephones; (b) 86% of households contained women aged 18–69; and (c) 15% listened to the Pap status domain of the IVR message. The percentage of women who would not have had a hysterectomy was estimated to be 85% (Karen Gillham, personal communication, 2001; 1997 NSW Health Survey).

The percentage of women aged 18–69 unscreened in the previous two years was estimated to be 39%<sup>252</sup>. The efficacy of the IVR intervention was unknown and was the subject of the trial. It was assumed to be 50%.

Therefore, in order for an additional 378 women to be screened, the number of households to contact was given by:

$$19,643 = \frac{378}{0.90 \times 0.86 \times 0.15 \times 0.85 \times 0.39 \times 0.50} \quad (7.2)$$

This figure was rounded up to 20,000 households.

If the percentage of women who had not had a hysterectomy was the same in the two age groups (18–49 and 50–69) then the resulting increase in screening rates would be the same for each: 1.0%, given that all the other estimates were as stated above. However, the hysterectomy rates in the age groups 20–49 and 50–69 were 7% and 33%, respectively (Karen Gillham, personal communication, 2001; 1997 NSW Health Survey). This resulted in expected screening rate increases for the age groups 20–49 and 50–69 of 1.1% and 0.8%, respectively.

### Sample selection

Households were randomly selected in proportion to the postcode population from the intervention postcodes using an electronic version of the telephone White Pages<sup>215</sup>. The sampling program is shown in Appendix Section E.1. Due to rounding approximations the sampling program selected 19,994 households rather than 20,000. The sample details by postcode are shown in Table 7.1.

Figure 7.3 shows the selected postcodes. The extent of the SLAs in the Hunter area has been overlaid as a grey shadow. All but one of the selected postcodes at least partly overlap the Hunter SLAs; the postcode 2787 appears

Table 7.1: Intervention and control group assignment, number of eligible women, Index of Socio-Economic Disadvantage, and number of selected households for selected postcodes.

<i>Postcode</i>	<i>Assignment</i>	<i>Women</i> <sup>a</sup>	<i>SED</i> <sup>b</sup>	<i>Sample Size</i>
2259	Intervention	14181	981	4954
2264	Control	3025	984	
2267	Intervention	807	975	373
2286	Control	1061	996	
2290	Intervention	10132	1002	4155
2292	Intervention	509	923	349
2293	Intervention	576	826	280
2294	Control	490	750	
2295	Control	1567	928	
2296	Control	419	801	
2297	Control	426	876	
2298	Intervention	2853	951	1336
2299	Intervention	3107	946	1561
2305	Intervention	3634	1033	1828
2307	Control	1100	983	
2311	Intervention	349	1015	167
2315	Intervention	2613	957	1957
2328	Intervention	794	1026	290
2329	Intervention	569	1005	326
2333	Intervention	3583	997	1630
2336	Intervention	647	1021	327
2338	Control	461	973	
2339	Control	188	1022	
2340	Control	12339	1002	
2343	Control	1563	1004	
2420	Control	1420	1009	
2421	Control	278	1057	
2423	Intervention	949	964	461
2787	Control	1611	1034	
2850	Control	4281	972	

<sup>a</sup> Number of women aged 18–69 in each postcode.

<sup>b</sup> Index of socio-economic disadvantage.

distant from the others. This resulted from a data error found at a later time in the NLI<sup>295</sup>. This postcode was omitted from the analyses.

### 7.3.4 Procedure

Approval was obtained for this study prior to commencement from the University of Newcastle Human Research Ethics Committee (HREC) and Hunter Area Research Ethics Committee (HAREC).

#### Participant instruction

Information letters printed on letterhead stationery and addressed to the name identified from the telephone directory were posted using DL-sized\* envelopes with the institution details (Hunter Centre for Health Advancement (HCHA), Hunter Health, and The University of Newcastle) plainly marked, one week before the first call. The letters (Appendix Section E.2) were in English since only 4.0% of the Hunter Statistical Division population speak another language at home<sup>208</sup>. The letters (Figure E.1 of Appendix Section E.2) explained the project aims, stated that each household would receive an IVR call about cervical screening, and that participation would be voluntary. An information sheet (Figure E.2 of Appendix Section E.2) was included giving instructions about using the system. No remuneration was offered.

#### Participant recruitment

The Generalized Electronic Interviewing System (GEIS) running on the Recorded Voice System (RVS) (The RVS is defined in Section 3.4.2 of Chapter 3)

---

\*Envelope size 220 x 110mm.

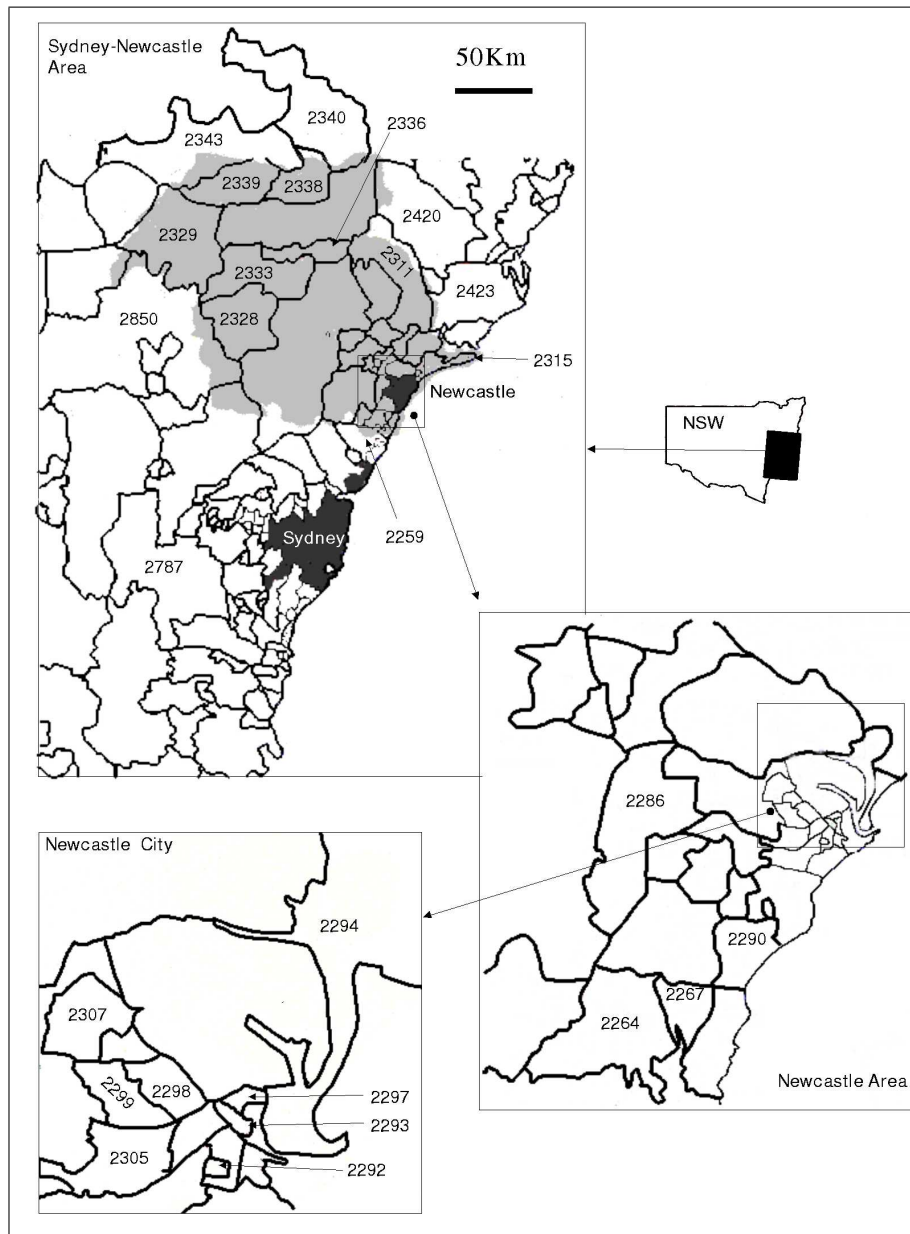


Figure 7.3: Postcodes selected overlaid with statistical local areas of the Hunter region.

initiated calls to households. When a call was answered, GEIS played an introductory message and then asked the person to press key ‘5’, although any key actually sufficed. If no key press was detected after three repetitions the system thanked the person and hung up. Otherwise GEIS described the survey purpose before asking if there were any women aged 18–69 in the household. Since there may have been more than one eligible woman in the household, it asked to speak to the woman with the most recent birthday<sup>209</sup>. If the current time was not convenient or the woman was unavailable GEIS explained it could call back the next day at the same time or some other time of the woman’s convenience. Self-identified businesses were recorded as out of scope. GEIS distinguished answering machines from genuine individuals by the length of the salutation:

“Hello?”

“Hello, we’re not in at present...”

### **Brief intervention**

In the pilot (Chapter 6) an IVR brief intervention was conducted in the Cessnock SLA as part of a community-based intensive awareness campaign of which an IVR message was a component (Section 6.3.4 in Chapter 6). In this trial the IVR message as a brief intervention was evaluated alone. The IVR intervention consisted of calls made by GEIS to households in intervention postcodes. Households in control postcodes did not receive telephone calls or letters.

As with the pilot, the intervention was developed in discussion with the cervical screening coordinator. There was no follow-up Computer-Assisted Telephone Interview (CATI) because the acceptability of the method had already

been examined as described in Chapter 6.

GEIS conducted the following sequence of operations:

1. Explained the nature of the call;
2. Identified if women aged 18–69 were present;
3. Selected one eligible woman;
4. Determined her screening status;
5. Delivered a message that either congratulated her on being correctly screened, a message of encouragement if she was under-screened, or another message appropriate to her status;
6. Offered to deliver additional messages designed to counter common barriers to screening (*Cost, Discomfort/Fear of pain, Embarrassment, Fear of finding cancer, and Lack of time*)
7. Offered to provide additional information on cervical screening and cancer (*Lack of knowledge of cancer, Lack of knowledge of test, and Limited access to medical services*);
8. Offered to read out contact numbers and an Internet address where she could obtain more information;
9. Offered to arrange for someone to ring her back if she still had questions; and then
10. Offered to record a question the woman may wish answered by the person making the call-back (the Hunter Cervical Screening Coordinator (HCSC)).



**Health Insurance Commission (HIC) data**

Population data was used rather than self-report to avoid a social desirability bias, inaccurate response, or sample bias<sup>247,288</sup>. Cervical screening statistics were obtained from the HIC. The HIC makes available statistics from the Medicare Benefits Scheme on the items in the Medicare Benefits Schedule, including cervical screening services, by demographic variables, such as age, and geographic variables, such as postcodes. The HIC data has previously been found to account for 90% of cervical screens in NSW<sup>338</sup> and has been found to be accurate when compared to other data sources<sup>247</sup>. The HIC does not record tests taken by non-Medicare providers (community centres, health nurses, and public hospitals)<sup>253</sup>, but this was not expected to be important in this study since there was no reason to suppose that the efficacy of the IVR method would be associated with screening location.

Data request protocols were prepared for submission to the HIC as shown in Appendix Section E.3. The HIC statistics consisted of monthly summaries of the numbers of cervical and vaginal smears from the postcodes in the intervention and control groups for women whose previous screening had been within the last year, between one and two years ago, more than two years ago, or who had no previous screening record. The report period spanned six months prior to and following intervention: October 2000 – January 2002.

### 7.3.5 Apparatus

#### Equipment

The equipment used was the same as in Chapter 5. Voice recordings were spoken in English by a single female staff member.

#### Software

The IVR system was implemented using GEIS (see Chapter 3). Answers to questions were provided by selecting one of a set of options, entering a number or date, or entering an open-ended response. If the woman chose to record a question for later follow-up, this was recorded as a file on the RVS's hard disk and also sent as an e-mail attachment to the HCSC. For numeric answers, absolute and reasonable limits prevented range errors. Invalid responses triggered an appropriate message. After three repetitions of an item, a non-response caused GEIS to hang up and record a refusal. Women could refuse to answer a particular item and if they wished they could return to earlier questions and modify their answers.

### 7.3.6 Measures

#### Script

The script for the system is shown in Supplementary Materials Section 7.1 and was almost the same as the pilot script (Supplementary Materials Section 6.1). Minor changes were made to address problems identified in the pilot: the sensitivity of age and attrition through the message. In the pilot, the participant had been asked to enter her exact age. This was changed so that she needed to only

nominate an age group ( $<18$ ,  $18-49$ ,  $50-69$ ,  $\geq 70$ ). Those items that appeared to increase attrition in the pilot script were reworded to remove wording that might suggest the message had ended prematurely.

The script handled all aspects of the call, including initiating the interview, identifying an eligible woman, delivering the message, providing information, and offering to arrange for a call-back. If the woman indicated she wanted a call-back she was asked if she wanted to record a question to be answered. This was done using an open-ended question which allowed a message of up to thirty seconds to be recorded. After the interview terminated, GEIS generated an e-mail that was sent to the Hunter Cervical Screening Coordinator (HCSC). If a question had been recorded, this was included with the e-mail as an attachment. The HCSC was a local staff member responsible for cervical screening promotion in the Hunter region.

### **Domains**

The script (Supplementary Section 7.1) contained domains concerned with Pap status, barrier messages, demographic items, information items, and contact numbers.

In the Pap status domain, the woman's cervical screening status was determined using the same method as in Chapter 6 (see Figure 6.1 in Chapter 6). Those women who completed the Pap status domain were classified as either:

- Less than 18 years old;
- Over 70 years old;
- Not had sexual intercourse;

- Had a hysterectomy;
- Under-screened;
- Correctly screened.

If the woman was under-screened, a series of messages were offered that addressed known barriers to screening (Cost, Discomfort/Fear of pain, Embarrassment, Fear of finding cancer, and Lack of time). The demography domain included questions on education, marital status, country of birth, aboriginality, employment status, and income. The information messages domain included messages covering a range of topics concerning additional barriers to screening (Lack of knowledge of cancer, Lack of knowledge of test, and Limited access to medical services). In the contact numbers domain were messages about numbers to ring for more information, as well as an opportunity to have someone call back.

### 7.3.7 Analysis

#### Participation rates

The last item that was played was used to determine where each woman dropped out or if they completed the whole message. A participant was counted as listening to a domain if they had reached the last item in that domain, or had begun the domain but had validly skipped the remaining items because they were not relevant. The participation rates were then tabulated for each completed domain of the IVR message by using Equation 5.5 of Chapter 5.

### Demography

Demographic variables (age, education, marital status, country of birth, aboriginality, employment, income) were tabulated. Since the ABS Census estimates represent population level data, the demographic variables' profiles were compared to 1996 ABS Census profiles<sup>208</sup> using chi-square goodness-of-fit tests<sup>225</sup> after classifying the variable levels.

Ages of participants were classified into the groups *18–49* and *50–69*. Where participants refused to give their age the data were not included in calculations involving age. For Education, the levels *no education*, *primary school*, and *secondary school* were collapsed to *school or none* to ensure sample sizes were at least 20 per level. Also because of small category sample sizes, the levels of income were collapsed to less than *\$30,000*, and *\$30,000* or greater, which produced roughly equal frequencies.

### Efficacy

Although this study attempted to encourage screening in under-screened women aged 18–69, data sources for screening rates only reported on women aged 20–69. To ensure consistency with other data sources, cervical screening rates for the age groups 20–49 and 50–69 were calculated. As previously mentioned (Section 7.3.3), the effect was defined as the change in the screening rate in intervention postcodes compared to control postcodes following the intervention. The effect was separately calculated for women who had been screened within the previous two years, more than two years previously, and not previously screened. If the intervention was efficacious, the screening rate was expected to increase in the intervention postcodes, while no such increase was expected in

the control postcodes. To adjust for existing screening rate differences between the control and intervention postcodes, the differences in pre-intervention period counts of the two groups was subtracted from the post-intervention change. The cumulative change for each month following intervention is referred to herein as the *effect*. The effect for month  $j$  following intervention was calculated as shown in Equation 7.3.

$$E_j = \frac{1}{15} \sum_{p=1}^{15} \left\{ \frac{1}{N_{Ip}} \left( \sum_{i=+1}^{+j} S_{Ipi} - \sum_{i=-1}^{-j} S_{Ipi} \right) - \frac{1}{N_{Cp}} \left( \sum_{i=+1}^{+j} S_{Cpi} - \sum_{i=-1}^{-j} S_{Cpi} \right) \right\} \quad (7.3)$$

where:  $E_j$  = effect for month  $j$ ,

$j = 1 \dots 6$ ,

$i = -1 \dots -6$  months prior to intervention,

$p$  = postcode,

$N_{Cp}$  = number of eligible women for the  $p$ th Control postcode,

$N_{Ip}$  = number of eligible women for the  $p$ th Intervention postcode,

$S_{Cpj}$  = number of screens in month  $j$  for the  $p$ th Control postcode, and

$S_{Ipj}$  = number of screens in month  $j$  for the  $p$ th Intervention postcode.

The effect was plotted by month. The effect was compared to other cervical screening intervention studies identified in a literature search conducted for articles published between 1988–2001 using the electronic databases Psycinfo<sup>14</sup>, Medline<sup>15</sup>, Sociofile<sup>16</sup>, and Health Reference Center<sup>17</sup>, with the search terms “vagina\$ smear\$, pap test\$, (papanicolaou adj2 (smear or test\$)), (cervical adj2 (smear or screen\$)), pap smear, intervention\$”.

### Cost per additional screen

The cost per additional screen was calculated by dividing the accumulated call cost by the number of additional screens expected in the sampled households. An additional calculation was made for a hypothetical sample that only included unscreened women aged 50–69 who had not had a hysterectomy. The cost per additional screen was compared to those studies identified in the above literature search that reported screening cost.

## 7.4 Results

The IVR system was active between April–July 2001 and made a total of 28,087 calls.

### 7.4.1 Participation rates

Of the 19,994 numbers in the sample 2297 were not rung. These included three that were excluded due to prior notification that they were not households, 2055 after information letters were returned, 119 due to prior refusals by letter or telephone, 35 due to notification of the death of the recipients, and 85 due to technical difficulties. This left 17,697 numbers. Of these, 533 were disconnected lines and 156 were fax machines, leaving 17,008 numbers.

As indicated in Table 7.2, a total of  $1795 \div 17,008 = 11\%$  households listened to the entire IVR message,  $2253 \div 17,008 = 13\%$  listened to at least the Pap status domain. The participation rates varied significantly between post-codes for the Pap status domain ( $\chi^2=78.0$ ,  $df=14$ ,  $P<0.01$ ,  $n=17,008$ ), barrier messages domain ( $\chi^2=83.6$ ,  $df=14$ ,  $P<0.01$ ,  $n=17,008$ ), demography domain

Table 7.2: Participation rates by last domain completed and post code, and in comparison to the Cessnock pilot.

<i>Postcode</i>	<i>Domain (%)</i>					<i>All (%)</i>
	<i>Pap status<sup>a</sup></i>	<i>Barrier<sup>b</sup></i>	<i>Demography<sup>c</sup></i>	<i>Information<sup>d</sup></i>	<i>Contact<sup>e</sup></i>	
2259	14	13	13	12	11	10
2267	10	10	10	10	9	9
2290	16	16	15	14	13	13
2292	9	8	8	8	7	6
2293	7	7	7	7	6	6
2298	13	12	12	11	10	10
2299	9	9	9	8	8	7
2305	10	10	9	9	8	8
2311	11	11	11	10	10	10
2315	11	11	11	10	9	8
2328	18	18	17	17	16	14
2329	13	13	13	13	12	11
2333	12	12	11	11	10	9
2336	15	15	14	14	12	11
2423	14	14	13	13	12	11
n	2253	2226	2133	2017	1880	1795
Pilot <sup>f</sup> (%)	15	15	14	11	-	11

<sup>a</sup> Pap status domain.<sup>b</sup> Barrier messages domain.<sup>c</sup> Demography domain.<sup>d</sup> Information messages domain.<sup>e</sup> Contact numbers domain.<sup>f</sup> Cessnock Pilot (Chapter 6).

( $\chi^2=77.7$ ,  $df=14$ ,  $P<0.01$ ,  $n=17,008$ ), information messages domain ( $\chi^2=77.3$ ,  $df=14$ ,  $P<0.01$ ,  $n=17,008$ ), contact numbers domain ( $\chi^2=76.5$ ,  $df=14$ ,  $P<0.01$ ,  $n=17,008$ ), and for the whole message ( $\chi^2=75.0$ ,  $df=14$ ,  $P<0.01$ ,  $n=17,008$ ).

Compared to the domains that were used in the Cessnock pilot, the rates were lower for the Pap status domain (Fisher's exact test,  $\chi^2=12.4$ ,  $P<0.01$ ,  $n=22,129$ ), barrier messages domain (Fisher's exact test,  $\chi^2=9.8$ ,  $P<0.01$ ,  $n=22,129$ ), and demography domain (Fisher's exact test,  $\chi^2=12.4$ ,  $P<0.01$ ,  $n=22,129$ ), but did not differ for the information messages domain (Fisher's exact test,  $\chi^2=0.9$ ,  $P=0.79$ ,  $n=22,129$ ) and overall (Fisher's exact test,  $\chi^2=2.1$ ,  $P=0.15$ ,  $n=22,129$ ).



Table 7.3: Age, education, marital status, country of birth, employment status, and household income of listeners to the IVR message compared to the ABS census.

<i>Variable</i>	<i>N</i>	<i>%</i>	<i>ABS</i>	<i>P Value</i> <sup>a</sup>
<b>Age</b>				<0.01
18–49	968	55	69	
50–69	800	45	31	
<b>Education</b>				<0.01
School or none	1000	58	75 <sup>b</sup>	
Vocational	441	26	9	
University	292	17	16	
<b>Marital status</b>				<0.01
Divorced / Separated	155	9	9 <sup>b</sup>	
Married / de facto	1387	80	59	
Never married	92	5	28	
Widowed	94	5	5	
<b>Country of birth</b>				
Australia	1539	89	87	0.18
<b>Employment</b>				<0.01
Employed	650	38	50 <sup>b</sup>	
Unemployed	85	5	5	
Not in labour force	970	57	45	
<b>Household income</b>				0.18
\$0–\$29,999	463	45	47 <sup>c</sup>	

<sup>a</sup> Chi-square goodness-of-fit test comparing sample proportions to ABS profiles.

<sup>b</sup> Females aged 15–69.

<sup>c</sup> All females.

### 7.4.2 Demography

As shown in Table 7.3, the demographic composition of participants to the IVR message differed from the 1996 Census<sup>208</sup> on most variables. Compared to the ABS data, the sample over-represented older women (50–69 y.o.), those with post-school qualifications, and married women.

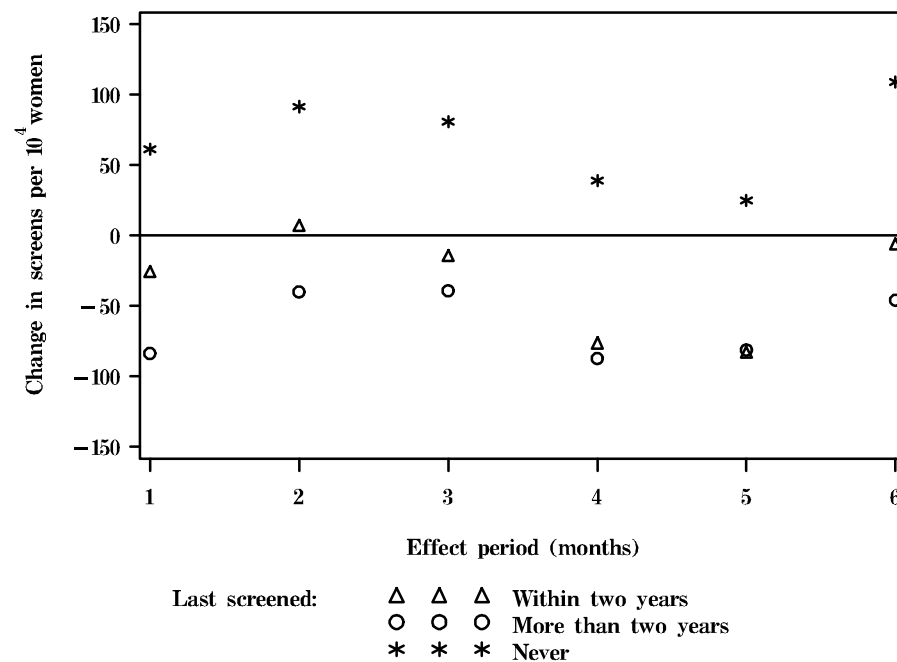


Figure 7.4: Effect (change in the screening rate following the intervention in intervention postcodes compared to control postcodes) by month and screening status for women aged 20–49 years.

### 7.4.3 Efficacy

At three months, an overall increase of 43 per 10<sup>4</sup> women (0.43%) screens occurred in the intervention postcodes compared to the control postcodes. However, the change in screening rates depended on the age and previous screening status of women. The change in screening rates is shown in Figure 7.4 for women aged 20–49, and in Figure 7.5 for women aged 50–69.

As shown in Figure 7.4, the screening rates for women aged 20–49 who had not had a previous screen increased in intervention postcodes compared to control postcodes for at least six months following intervention. However, the rates for women who had been previously screened decreased.

As shown in Figure 7.5, by the third month the rates for all women aged

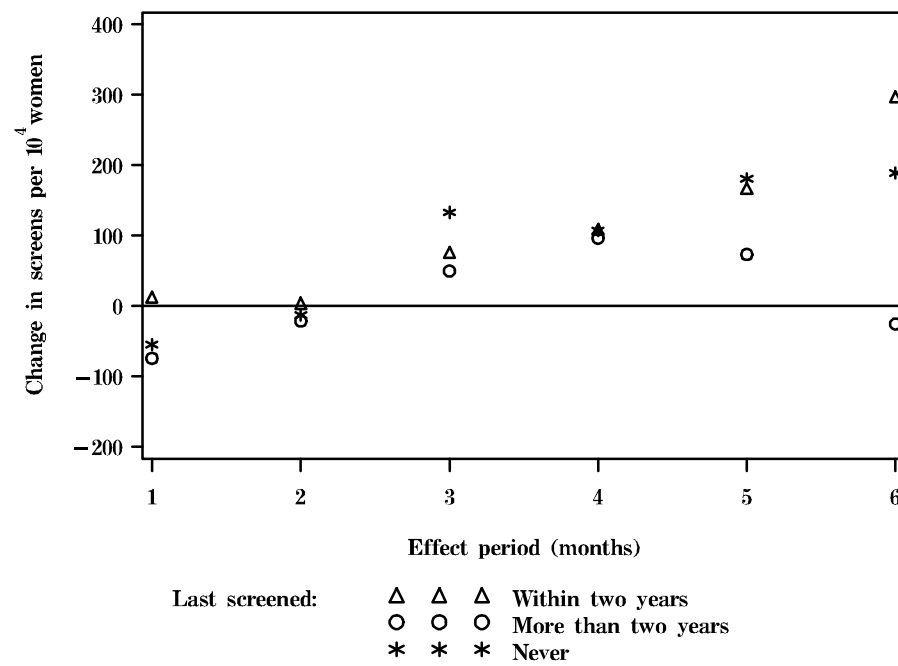


Figure 7.5: Effect (change in the screening rate following the intervention in intervention postcodes compared to control postcodes) by month and screening status for women aged 50–69 years.

50–69 increased. The maximum increases occurred at four months for women screened more than two years previously, and six months for those without a previous screen.

As shown in Table 7.4, the maximum effect for all women aged 20–49 of 16 additional screenings per  $10^4$  women (0.16%) occurred at two months following intervention, while the maximum for women aged 50–69 of 135 per  $10^4$  women (1.35%) occurred at six months.

The efficacy was defined above (Section 7.1) to be the proportion of under-screened women who listened to the IVR message and who were subsequently screened. The assumed efficacy was 50% and the target screening rate increase at three months for younger women was 1.1% (Section 7.3.3). Using the observed increase in the unscreened younger women at three months of 0.81% (Table 7.4), the efficacy was determined to be  $37\% = 0.81\% \times 50\% \div 1.1\%$  (Section 7.3.3). Since the rate for the younger women screened more than two years ago did not increase, the efficacy for this group was zero. For the older women, the target increase was 0.8% and the observed increase was 1.34% for the unscreened women and 0.49% for women screened more than two years ago. The efficacies for these two groups were  $84\% = 1.34\% \times 50\% \div 0.8\%$  and  $31\% = 0.49\% \times 50\% \div 0.8\%$ , respectively.

#### 7.4.4 Cost per additional screen

As noted in Chapter 6, the cost of hardware and software development could be disregarded as these were capital expenditures and would not normally be associated with a single project. The cost of printing, packaging, and posting the information letters were relevant to this project, but might not necessarily

Table 7.4: Effect per  $10^4$  women by age group, previous screening, and effect period<sup>a</sup>.

<i>Month</i>	<i>20–49 Yr.</i>			
	<i>L.E. 2 Yr.<sup>b</sup></i>	<i>G.T. 2 Yrs.<sup>c</sup></i>	<i>None Previous<sup>d</sup></i>	<i>All</i>
1	-26	-84	61	-18
2	7	-40	92	16
3	-14	-39	81	5
4	-76	-87	39	-46
5	-83	-81	25	-51
6	-6	-46	109	14
	<i>50–69 Yr.</i>			
	<i>L.E. 2 Yr.</i>	<i>G.T. 2 Yrs.</i>	<i>None Previous</i>	<i>All</i>
1	13	-74	-54	-36
2	4ns	-21	-13ns	-18
3	76	50	133	75
4	109	96	107	103
5	167	73	181	128
6	297	-26	189	135

<sup>a</sup> ns=Not Significant; t-test for  $H_0$ : Effect=0.<sup>b</sup> Women screened within the last two years.<sup>c</sup> Women screened more than two years ago.<sup>d</sup> Women not previously screened.

occur in an IVR intervention. The remaining cost was for call charges. The details of the cost calculation are shown in Table 7.5. The cost calculation in the table uses the additional number of women screened in the sample.

The cost per additional screen was \$388.23. If the sample had been pre-selected to only consist of unscreened women aged 50–69 who had not had a hysterectomy then the cost per additional screen would have been \$34.29.

## 7.5 Discussion

### 7.5.1 Participation rates

The proportion of women who listened to at least part of the IVR message was 13%, slightly less than in the Cessnock pilot, although the two rates did not differ

Table 7.5: Calculation of cost for the IVR brief intervention.

<i>Item</i>	<i>n</i>	<i>\$</i>
Number of telephone calls <sup>a</sup>	28,087	
Telephone unit call cost		0.33
Telephone cost		9186.24
<i>Case I: random household selection<sup>b</sup></i>		
Number eligible	5633	
Number of additional screenings <sup>c</sup>	56	
Cost per screening (tel. cost)		388.23
Cost per screening (total cost)		980.23
<i>Case II: preselecting women aged 50–69 requiring a Pap test<sup>d</sup></i>		
Number eligible	19,994	
Number of additional screenings <sup>e</sup>	624	
Cost per screening (tel. cost)		34.29
Cost per screening (total cost)		86.57

<sup>a</sup> Number of calls for a sample size of 19,994 households.<sup>b</sup> A random sample of households contain women who are aged 20–69 (85%), without a hysterectomy (85%), and unscreened (39%).<sup>c</sup> Assume 0.42% additional screenings are expected with IVR.<sup>d</sup> All households in the sample contain women who are aged 50–69, without a hysterectomy, and unscreened.<sup>e</sup> Assume 1.34% additional screenings are expected with IVR.

significantly. The rates differed from the Cessnock pilot for some domains, but since the rates also differed between postcodes this difference can be attributed to regional variation. The overall rate of 13% confirmed the conclusion reached in Chapter 6 that cold-calling using IVR does not allow large proportions of a population to be accessed. Many individuals continued to drop out progressively through the message, although 77% were retained overall, compared to 73% in the Cessnock pilot.

### 7.5.2 Demography

As in Chapter 6, the demographic composition of listeners to the IVR message differed from the 1996 ABS Census<sup>208</sup> on most variables. However, the previous chapter also concluded that those more at risk were more likely to listen to the IVR message. This conclusion is supported by the greater representation and effect noted in this study in the 50–69 age group.

### 7.5.3 Efficacy

The literature search identified a number of studies that described cervical screening interventions. These studies are listed in Table 7.6. Many of these studies<sup>254,278–280,283,289,339–350</sup> compared pre-intervention and post-intervention rates at distinct times.

A strength of this study was that the intervention dates for the postcodes were generally staggered, which minimized the effect of any possible external phenomena such as unrelated or unknown screening campaigns. Many of the studies<sup>280,343–350</sup> identified in the literature search lacked validly replicated designs. By contrast, this study used multiple postcodes that were randomly

Table 7.6: Methods used, effects, and effect determination methods observed in cervical screening intervention studies.

<i>Study</i>	<i>Method</i> <sup>a</sup>	<i>Effect</i> <sup>bc</sup>	<i>O.D.M.</i> <sup>d</sup>	<i>Note</i> <sup>e</sup>
Hyndman et al. <sup>351</sup>	1, 5, 6	4.4ns, 8.9, 13.6	3	
Jenkins et al. <sup>346</sup>	3	0.0ns	4	
Paskett et al. <sup>343</sup>	3	21.0	4	2.
Suarez et al. <sup>350</sup>	3	5.4–6.2ns	4	
Suarez et al. <sup>347</sup>	3	-0.7ns	4	
Hancock et al. <sup>283</sup>	3	1.2	2	
Allen et al. <sup>341</sup>	4	2.8ns	4	
Gotay et al. <sup>348</sup>	4	2.0	4	
Margolis et al. <sup>340</sup>	4	12.9	3	
Sung et al. <sup>342</sup>	4	-2.0ns	4	
Whitman et al. <sup>344</sup>	4	36.0	3	
Byles and Sanson-Fisher <sup>279</sup>	5	not reported	2	
Del Mar et al. <sup>339</sup>	5	4.0ns	1	
Buehler and Parsons <sup>352</sup>	5	0.9ns	1	4.
Somkin et al. <sup>353</sup>	5, 2	10.3, 13.7	3	
Byles et al. <sup>280</sup>	5, 11	3.1, 2.9	2	1.
McDowell et al. <sup>354</sup>	6, 1, 12	-0.7ns, 8.9ns, 2.4ns	3	
Shelley et al. <sup>349</sup>	7	30	2	
Mitchell et al. <sup>289</sup>	7	6.7	1	1.
Byles et al. <sup>254</sup>	7, 5	(28, -14ns, 24), (28ns, 11ns, 54)	4	3, 5.
Mandelblatt et al. <sup>345</sup>	8	32.7	3	
Ward and Proude <sup>355</sup>	8	-28.0ns	4	
Bowman et al. <sup>278</sup>	9, 6, 10	1.6ns, -0.5ns, 12.7	2	
Lantz et al. <sup>356</sup>	10, 12	17.9	2	
Yancey et al. <sup>357</sup>	13	4.3–7.5	3	
Burack et al. <sup>358</sup>	14, 1	1.0ns, 1.0ns	3	

<sup>a</sup> 1=chart reminders, 2=chart reminders & information letters, 3=community intervention, 4=health workers, 5=information letters, 6=invitation letters to attend a specific clinic, 7=media, 8=nurse-delivered intervention, 9=pamphlets, 10= physician letters, 11=prompt cards, 12=telephone calls, 13=video, 14=reminder letters for regular screening.

<sup>b</sup> Effect corresponding to each listed method expressed as percentage change in screening rates.

<sup>c</sup> ns=No significant change.

<sup>d</sup> Outcome Determination Method: 1=Cytology registry, 2=Medical insurance claims, 3=Medical records, 4=Surveys.

<sup>e</sup> 1=effect expressed as change over 3 months, 2=effect expressed as change over 3 years, 3=effect expressed as relative changes, 4=effect change restricted to unscreened women, 5=effect reported for 3 different locality types.



assigned to the control and intervention groups and which were not necessarily geographically close. Lastly, unlike many other studies, this study used health service data rather than self-report<sup>254,341–343,346–348,350,355</sup>.

These aspects of the design were likely to have increased the reliability of the efficacy estimate. However, the efficacy was obtained by comparing the expected and maximum observed screening rates. This was an indirect calculation that involved several quantities each of which had some unknown degree of error (Section 7.3.3). Therefore, the estimated efficacy was an approximation to the actual efficacy.

Many of the effects reported by the studies in Table 7.6 were non-significant and the remainder were variable in magnitude with the largest being 36%. The variability arose from the variety of methodologies used and also because some were changes in screening rates amongst samples while others were of populations. A number of different intervention methods were used, but there was no clear indication that particular methods were more efficacious than others. There was also no clear difference between those studies that determined the effect using self-report and those using an objective measure, such as a cytology registry, medical insurance claims, or medical records.

The results of this study suggested that some of the variability in the effect observed in the listed studies might have occurred due to differential response by age and screening status. To investigate this, the studies were further examined for age and screening status effects.

As shown in Table 7.7, few studies provided indications of age and screening status effects. Apart from Bowman et al.<sup>278</sup>, larger effects were reported for the older age groups, which is consistent with this study. Three studies reported

Table 7.7: Age effects, screening status effects, and cost per additional screen observed in cervical screening intervention studies.

<i>Study</i>	<i>Age Effect<sup>a</sup></i>	<i>Screening Status Effect<sup>b</sup></i>	<i>Cost per screening</i>
351			14.75, 45.35, 48.21 <sup>c</sup>
343	$\left\{ \begin{array}{ll} < 65: & OR = 1.0 \\ \geq 65: & OR = 1.9 \end{array} \right.$		
347	$\left\{ \begin{array}{ll} 40-49: & -15 \\ 50-64: & 1 \\ 65-69: & 11 \end{array} \right.$		
283		Screening increases were observed for under-screened women, but none for correctly-screened women.	
340	$\left\{ \begin{array}{ll} 40-59: & 9 \\ \geq 60: & 18 \end{array} \right.$	Screening increases were observed for under-screened women, but none for correctly-screened women.	750.00 <sup>d</sup>
352	$\left\{ \begin{array}{ll} < 40: & 0 \\ \geq 40: & 10 \end{array} \right.$		
353	$\left\{ \begin{array}{ll} 20-49: & 11, 13 \\ 50-64: & 9.0, 5.1^e \end{array} \right.$		
354			12.34, 14.23, 11.26 <sup>f</sup>
349		Those not screened within previous two years increased by 50%.	
278	$\left\{ \begin{array}{ll} 18-34: & 3.62 \\ 35-54: & 1.10 \\ 55-70: & 1.00 \end{array} \right.$		

<sup>a</sup> Age groups with effects expressed as percentages, except where indicated otherwise.<sup>b</sup> Effects for specific screening status groups.<sup>c</sup> Cost (\$AU) of chart reminders, invitations, and appointment letters.<sup>d</sup> Cost (\$US) of staff salaries.<sup>e</sup> Effects for letters and chart reminders, respectively.<sup>f</sup> Cost (\$CAN) of reminder calls, chart reminders, and telephone calls.

Table 7.8: Estimated percentage effect rates for varying efficacy and listening percentages.

<i>Efficacy</i>	<i>Percentage Listening</i>			
	15% <sup>a</sup>	50%	75%	100%
22% <sup>b</sup>	1	3	4	6
50%	2	6	10	13
75%	3	10	14	19
100%	4	13	19	26

<sup>a,b</sup> Observed values.

an interaction between effect and previous screening status in which a larger effect was observed for under-screened women. This study found that the effect on under-screened women depended on their ages, with older women showing a positive change and younger women screened more than two years previously showing a negative change. These results suggest that cervical screening interventions should be customized for different age and screening status groups.

Apart from the variability of the effect reported by the listed studies, it was also noted that increases in screening magnitude were larger than in this study. However, the effect size obtained in this study was close to the pre-planned target value. The maximum possible effect may be obtained from the calculations of Section 7.3.3. Of the listed variables only efficacy and the percentage of participants are potentially modifiable. As shown in Table 7.8 both variables limit the effect magnitudes that may be obtained and if the value of either could be increased then the effect may be improved to a maximum of 26%. This limit arises because the IVR method can only contact households with telephones, and of those households, only some have eligible under-screened women who have not had a hysterectomy.

However, recontacting non-responding households at regular intervals may increase the proportion of participants. This speculation is supported by the rise in acceptability with increasing exposure to IVR technology found in Chapter 5. The efficacy of letter interventions has been found to diminish with repeated application<sup>279</sup>. Although this may also occur with an IVR intervention over time, its novelty may still attenuate the fall in efficacy.

Another method of increasing the effect may be to combine IVR calls with existing intervention methods that have been shown to be effective in encouraging women to have Pap tests. Three possible methods are discussed below:

1. Mass media campaigns<sup>254,289,349,359,360</sup>;
2. Community interventions<sup>283,318,343,344,361,362</sup>; and
3. Information and invitation letters<sup>278,280,353,356,363,364</sup>.

**Media campaigns** Media campaigns could be used to increase awareness of an IVR call campaign beforehand. This might increase the percentage of those listening to the calls. Byles et al.<sup>254</sup> found a campaign involving television media, letters and physicians increased cervical screening rates in three rural areas. Alexander and McCullough<sup>359</sup> found that mass media methods were more successful in increasing cervical screening among a low income population of 85,029 women than posters and flyers. Mitchell et al.<sup>289</sup> found that the use of ethnic mass media increased cervical screening by 6.7% in postcodes with high proportions of non-English speakers than postcodes with low proportions of non-English speakers. Byles et al.<sup>360</sup> found a television campaign resulted in a screening rate increase of 5% in the intervention regions compared to the control regions. Shel-

ley et al.<sup>349</sup> found a 30% increase in screening rates in NSW following a mass media campaign that was significantly greater than increases in other states where the campaign was not run.

These studies indicate that it is possible to increase screening rates using the media. If a media intervention were to be combined with an IVR approach then the proportion of participants may potentially be increased. This may occur if the media intervention explained the nature of the calls so that they may be less confusing when received. As noted above, the proportion of participants was a limiting factor in this study. While a media approach may increase this proportion, the potential effect as shown in Table 7.8 is still limited by the efficacy of the method. It can be seen from Table 7.8, that if a media intervention resulted in 50% of eligible women listening then the effect would still only be 3%.

**Community interventions** A community intervention approach used by Paskett et al.<sup>343</sup> produced an increase in screening of 21% in an intervention city compared to a control city. In another community intervention, Dignan et al.<sup>361</sup> found a significant increasing screening trend in an intervention county compared to a control county, and in the same study, Michielutte et al.<sup>362</sup> found a slight increase in follow-up of abnormal smears for black women. Hancock et al.<sup>283</sup> found a 1.2% increase in screening rates following a community intervention in ten intervention towns compared to control towns. In a pair of towns, a community intervention reported by Whitman et al.<sup>344</sup> resulted in a screening rate increase of 36%.

The IVR study described in Chapter 6 was embedded within a larger

community intervention<sup>318</sup>. This approach could be extended to provide a more comprehensive package in which the IVR calls are highlighted by the other intervention methods. For example an inbound IVR facility may be advertised by meetings, flyers, and letters. Similarly, the use of outbound calls may be explained before they are received. These approaches may increase the proportion of participants. The use of an Internet web page may also provide a novel approach for information provision that may be integrated with other aspects of a community intervention, including an IVR system to allow women to schedule calls and obtain other information.

**Letter interventions** There exists the possibility of combining an IVR call with a letter approach. Bowman et al.<sup>278</sup> found screening increased by 12.7% among women who were sent a letter by their physician compared to a control group. A meta-analysis reported by Tseng et al.<sup>363</sup> found a pooled odds ratio that indicated that women who received a letter were 3.4 times more likely to rescreen than those who did not receive a letter. Mitchell et al.<sup>364</sup> found that women who received an invitation letter were 3.0 times more likely to screen than those not receiving a letter. Somkin et al.<sup>353</sup> found that women in a health maintenance organization receiving letters were more likely to screen than those not receiving letters. Byles et al.<sup>280</sup> found screening rates increased by 3% in two postcode areas following a mail campaign. A combination of letters and telephone calls was used by Lantz et al.<sup>356</sup> to increase screening rates by 18% among clinic patients compared to control women.

In this study, the IVR method's effect was of modest size, and even if the efficacy and the percentage of participants were maximized, the effect

would be no more than approximately 26% (Table 7.8). However, if the method were to be combined with a letter approach that targeted risk groups, such as identified older unscreened women, then these limitations would not apply and the effect size could be larger. This may be done by using existing databases currently used to mail letter reminders for cervical screening appointments, such as are conducted by the Pap Test Register (PTR) and which already store the womens' ages and screening statuses. An IVR call may also reach those women who move but retain their telephone number.

#### 7.5.4 Cost per additional screen

Of the studies listed in Table 7.6, four studies reported cost, but only the three studies shown in Table 7.7 reported the cost per additional screening. The three studies all involved increasing screening among clinic patients and so are difficult to compare to this study. The cost for the study by McDowell et al.<sup>354</sup> was adjusted for inflation through to 2001 for Canadian dollars<sup>†</sup> and then converted to Australian dollars<sup>‡</sup>. The cost for the study by Margolis et al.<sup>340</sup> was adjusted for inflation through to 2001 for United States dollars<sup>§</sup> and then converted to Australian dollars<sup>¶</sup>. The cost for the study by Hyndman et al.<sup>351</sup> was already in Australian dollars, but was adjusted for inflation through to 2001<sup>||</sup> dollars. The adjusted costs are shown along with this study's cost in Table 7.9.

The cost reported by McDowell et al.<sup>354</sup> and Hyndman et al.<sup>351</sup> were less

---

<sup>†</sup>Using rates from <http://www.bank-banque-canada.com>.

<sup>‡</sup>Using rates from <http://www.rba.gov.au>.

<sup>§</sup>Using rates from <http://minneapolisfed.org>.

<sup>¶</sup>Using rates from <http://www.rba.gov.au>.

<sup>||</sup>Using rates from <http://www.abs.gov.au>.

Table 7.9: Cost per additional screening for cervical screening methods identified in literature search studies compared to this study.

<i>Intervention Method</i>	<i>Cost per screening for each study<sup>a</sup></i>				
	Study:	351	354	340	This Study
Chart reminders		16.01	17.79		
Information letters		49.22			
Invitation letters		52.33	15.43		
Telephone calls			14.08		
Lay advisors				1290.00	
IVR (Case I) <sup>b</sup>					388.23
IVR (Case II) <sup>c</sup>					34.29

<sup>a</sup> Converted to Australian dollars and adjusted to 2001 Values.<sup>b,c</sup> Case I and II are defined in Table 7.5.

and the cost reported by Margolis et al.<sup>340</sup> exceeded the cost of this study (Case I in Table 7.9). The cost reported by McDowell et al.<sup>354</sup> included pre-paid reply envelopes and follow-up letters, and Hyndman et al.<sup>351</sup> included Pap test collection, McDowell et al.<sup>354</sup> and Margolis et al.<sup>340</sup> reported staff salary cost. The comparison studies all involved promoting cervical screening among women attending a clinic. This allowed the intervention to be directed only to specific women due for a Pap test, unlike in this study where calls were directed to households. Case II in Table 7.9 shows the cost that would have been obtained in this study for screening an identified risk group: older unscreened women without a hysterectomy. This cost favourably compares to the other studies, which suggests that the IVR method should be used to target identified individuals. For example, it may be speculated that IVR interventions may economically replace or complement existing letter reminders for cervical screening appointments, such as are conducted by the Pap Test Register (PTR).



## 7.6 Conclusions

One of the aims of this study was to increase cervical screening by a nominal 1.0%. This was partly achieved with an increase of 0.43% overall, but an increase for older women of 1.34%. This was a desirable outcome since older women are considered to be an at-risk group<sup>240</sup>. While 37% of unscreened younger women, 84% of older unscreened women, and 31% of older women screened more than two years previously appear to have been prompted to screen by the IVR message, the increases in screening rates were modest and were limited by the proportions of eligible under-screened women without hysterectomies who might be contacted by ringing households at random. The same factor also resulted in a considerable cost per additional screen.

The overall conclusion was that the IVR method obtained an increase in screening rates when directed towards risk groups, such as under-screened older women, but that a means was required of identifying these women in advance in order for this to be done economically. The use of a cervical screening register, such as that operated by the PTR, would be one means by which these women could be located. By linking an IVR system to the PTR cervical screening register it would be possible to directly and economically contact women, and provide an efficacious complementary approach to the existing letter reminder system.

# Bibliography

- [1] William L. Nicholls II. Computer-assisted telephone interviewing: A general introduction. In Robert M. Groves, Paul P. Biemer, Lars E. Lyberg, James T. Massey, William L. II Nicholls, and Joseph Waksberg, editors, *Telephone Survey Methodology*, pages 377–385. John Wiley & Sons, New York, 1988.
- [2] Sonia A. Alemagno, Deborah Cochran, Thomas E. Feucht, Richard C. Stephens, John M. Butts, and Stephanie A. Wolfe. Assessing substance abuse treatment needs among the homeless: A telephone-based interactive voice response system. *American Journal of Public Health*, 86(11):1626–1628, 1996.
- [3] Roger Tourangeau and Tom W. Smith. Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly*, 60:275–304, 1996.
- [4] Bradley O. Boekeloo, Laura Schiavo, David L. Rabin, Richard T. Conlon, Carol S. Jordan, and Diane J. Mundt. Self-reports of HIV risk factors by patients at a sexually transmitted disease clinic: Audio vs written questionnaires. *American Journal of Public Health*, 84(5):754–760, 1994.

- [5] John D. Piette, Stephen J. McPhee, Morris Weinberger, Connie A. Mah, and Frederic B. Kraemer. Use of automated telephone disease management calls in an ethnically diverse sample of low-income patients with diabetes. *Diabetes Care*, 22(8):1302–1309, 1999.
- [6] John D. Piette. Patient education via automated calls. A study of English and Spanish speakers with diabetes. *American Journal of Preventive Medicine*, 17(2):138–141, 1999.
- [7] Elizabeth Decker Tanke, Claudine M. Martinez, and Von O. Leirer. Use of automated reminders for tuberculin skin test return. *American Journal of Preventive Medicine*, 13(3):189–192, 1997.
- [8] Michael Havice. How response rates compare for human and digitized phone surveys. *Journalism Quarterly*, 66(1):137–142, 1989.
- [9] James C. Mundt, M.W. Perrine, John S. Searles, and Dan Walter. An application of interactive voice response (IVR) technology to longitudinal studies of daily behavior. *Behavior Research Methods, Instruments, and Computers*, 27(3):351–357, 1995.
- [10] James C. Mundt, John S. Searles, M.W. Perrine, and John E. Helzer. Cycles of alcohol dependence: Frequency-domain analyses of daily drinking logs for matched alcohol-dependent and nondependent subjects. *Journal of Studies on Alcohol*, 56:491–499, 1995.
- [11] Kim Bartel Sheehan and Mariea Grubbs Hoy. Using e-mail to survey internet users in the United States: Methodology and assessment. *Journal of Computer Mediated Communication [On-line]*, 4(3):Available: <http://www.ascusc.org/jcmc>, 1999.

- [12] Mick P. Couper. Web surveys. A review of issues and approaches. *Public Opinion Quarterly*, 64:464–494, 2000.
- [13] David R. Schaefer and Don A. Dillman. Development of a standard e-mail methodology. Results of an experiment. *Public Opinion Quarterly*, 62:378–397, 1998.
- [14] American Psychological Association. *PsycINFO [computer file]*. American Psychological Association, Washington, DC, 2001.
- [15] National Library of Medicine. *MEDLINE [computer file]*. National Library of Medicine (U.S.), Bethesda, Md., 2001.
- [16] Sociological Abstracts. *SocioFile [computer file]*. Sociological Abstracts, San Diego, 2001.
- [17] Information Access Co. *Health reference center [computer file]*. Information Access Co., Foster City, Calif., 2001.
- [18] American Statistical Association. *Current index to statistics extended database [computer file]*. American Statistical Association & the Institute of Mathematical Statistics, Bethesda, 2001.
- [19] Bell & Howell. *ABI Inform Global [computer file]*. Bell & Howell, Wooster, Ohio, 2000.
- [20] Academic Press. *IDEAL [computer file]*. Academic Press, San Diego, CA, 2001.
- [21] Institute for Scientific Information. *Social sciences citation index [computer file]*. Institute for Scientific Information, Philadelphia, Pa., 1989.

- [22] Michael J. Havice. Measuring nonresponse and refusals to an electronic telephone survey. *Journalism Quarterly*, 67(3):521–530, 1990.
- [23] Michael J. Havice. Touch-tone polling in a university setting. *College and University*, 65(3):227–234, 1990.
- [24] Michael J. Havice and Mark J. Banks. Live and automated telephone surveys: A comparison of human interviewers and an automated technique. *Journal of the Market Research Society*, 33(2):91–102, 1991.
- [25] Von O. Leirer, Daniel G. Morrow, Elizabeth Decker Tanke, and Grace M. Pariante. Elders' nonadherence: Its assessment and medication reminding by voice mail. *The Gerontologist*, 31(4):514–520, 1991.
- [26] Polly A. Phipps and Alan R. Tupek. Assessing measurement errors in a touchtone recognition survey. *Survey Methodology*, 17(1):15–26, 1991.
- [27] Umesh H. Patel and Charles F. Babbs. A computer-based, automated, telephonic system to monitor patient progress in the home setting. *Journal of Medical Systems*, 16(2-3):101–112, 1992.
- [28] Lee Baer, Michael W. Brown-Beasley, James Sorce, and Ana I. Henriques. Computer-assisted telephone administration of a structured interview for obsessive-compulsive disorder. *American Journal of Psychiatry*, 150:1737–1738, 1993.
- [29] Farrokh Alemi, Richard Stephens, Theodore Parran, Shirley Llorens, Pallav Bhatt, Ali Ghadiri, and Eric Eisenstein. Automated monitoring of outcomes: Application to treatment of drug abuse. *Medical Decision Making*, 14(2):180–187, 1994.

- [30] Robert W. Linkins, Eugene F. Dini, Grace Watson, and Peter A. Patriarca. A randomized trial of the effectiveness of computer-generated telephone messages in increasing immunization visits among preschool children. *Archives of Pediatrics and Adolescent Medicine*, 148:908–914, 1994.
- [31] Elizabeth Decker Tanke and Von O. Leirer. Automated telephone reminders in tuberculosis care. *Medical Care*, 32(4):380–389, 1994.
- [32] Lee Baer, Douglas G. Jacobs, Peter Cukor, John O’Laughlen, Joseph T. Coyle, and Kathryn M. Magruder. Automated telephone screening survey for depression. *Journal of the American Medical Association*, 273(24):1943–1944, 1995.
- [33] M.W. Perrine, James C. Mundt, John S. Searles, and Laurie S. Lester. Validation of daily self-reported alcohol consumption using interactive voice response (IVR) technology. *Journal of Studies on Alcohol*, 56(5):487–490, 1995.
- [34] Sid J. Schneider, Marc D. Schwartz, and Jeffrey Fast. Computerized, telephone-based health promotion: I. smoking cessation program. *Computers in Human Behavior*, 11(1):135–148, 1995.
- [35] John S. Searles, M.W. Perrine, James C. Mundt, and John E. Helzer. Self-report of drinking using touch-tone telephone: Extending the limits of reliable daily contact. *Journal of Studies on Alcohol*, 56(4):375–382, 1995.
- [36] A. M. Albisser, R. I. Harris, S. Sakkal, I.D. Parson, and S.C. En Chao.

- Diabetes intervention in the information age. *Medical Informatics*, 21(4): 297–316, 1996.
- [37] Farrokh Alemi, Richard C. Stephens, Rajshekhar G. Javalghi, Hayne Dyches, John Butts, and Ali Ghadiri. A randomized trial of a telecommunications network for pregnant women who use cocaine. *Medical Care*, 34(10):OS10–OS20, 1996.
- [38] J.P. Harding, L.R. Hamm, R.S.B. Ehsanullah, A.T. Heath, S.C. Sorrells, J. Haw, G.E. Dukes, S.G. Wolfe, A.W. Mangel, and A.R. Northcutt. Use of a novel electronic data collection system in multicenter studies of irritable bowel syndrome. *Alimentary Pharmacology and Therapeutics*, 11(6):1073–1076, 1997.
- [39] Kenneth A. Kobak, Leslie vH. Taylor, Susan L. Dottl, John H. Greist, James W. Jefferson, Diane Burroughs, David J. Katzelnick, and Monica Mandell. Computerized screening for psychiatric disorders in an outpatient community mental health clinic. *Psychiatric Services*, 48(8):1048–1057, 1997.
- [40] Kenneth A. Kobak, Leslie vH. Taylor, Susan L. Dottl, John H. Greist, James W. Jefferson, Diane Burroughs, Julie M. Mantle, David J. Katzelnick, Randal Norton, Henry J. Henk, and Ronald C. Serlin. A computer-administered telephone interview to identify mental disorders. *Journal of the American Medical Association*, 278(11):905–910, 1997.
- [41] James C. Mundt, Philip F. Kelleher, M.W. Perrine, and John S. Searles. Psychological performance assessment via interactive voice response sys-

- tems. *Behavior Research Methods, Instruments, and Computers*, 29(4):506–518, 1997.
- [42] John D. Piette and Connie A. Mah. The feasibility of automated voice messaging as an adjunct to diabetes outpatient care. *Diabetes Care*, 20(1):15–21, 1997.
- [43] John H. Greist, Issac M. Marks, Lee Baer, J. Richard Parkin, Peter A. Manzo, Julia M. Mantle, Keith W. Wenzel, Christopher J. Spierings, Kenneth A. Kobak, Susan L. Dottl, Terese M. Bailey, and Leslie Forman. Self-treatment for obsessive compulsive disorder using a manual and a computerized telephone interview: A U.S.-U.K. study. *M.D. Computing*, 15(3):149–157, 1998.
- [44] Paula Kadison, Elise M. Pelletier, Edgar L. Mounib, Paul Oppedisano, and Harry T. Poteat. Improved screening for breast cancer associated with a telephone-based risk assessment. *Preventive Medicine*, 27(3):493–501, 1998.
- [45] I.M. Marks, L. Baer, J.H. Greist, J.M. Park, M. Bachofen, A. Nakagawa, K.W. Wenzel, J.R. Parkin, P.A. Manzo, S.L. Dottl, and J.M. Mantle. Home self-assessment of obsessive-compulsive disorder. *British Journal of Psychiatry*, 172:406–412, 1998.
- [46] Luigi F. Meneghini, A. Michael Albisser, Ronald B. Goldberg, and Daniel H. Mintz. An electronic case manager for diabetes control. *Diabetes Care*, 21(4):591–596, 1998.
- [47] James C. Mundt, Kenneth A. Kobak, Leslie V.H. Taylor, Julia M. Mantle,



- James W. Jefferson, David J. Katzelnick, and John H. Greist. Administration of the Hamilton depression rating using interactive voice response technology. *MD Computing*, 15(1):31–39, 1998.
- [48] Deborah J. Osgood-Hynes, John H. Greist, Isaac M. Marks, Lee Baer, Susan W. Heneman, Keith W. Wenzel, Peter A. Manzo, J. Richard Parkin, Christopher J. Spierings, Susan L. Dottl, and Helene M. Vitse. Self-administered psychotherapy for depression using a telephone-accessed computer system plus booklets: An open U.S.-U.K. study. *Journal of Clinical Psychiatry*, 59(7):358–365, 1998.
- [49] Martin Bachofen, Akiko Nakagawa, Isaac M. Marks, Je-Min Park, John H. Greist, Lee Baer, Keith W. Wenzel, J. Richard Parkin, and Susan L. Dottl. Home self-assessment and self-treatment of obsessive-compulsive disorder using a manual and a computer-conducted telephone interview: Replication of a U.K.-U.S. study. *Journal of Clinical Psychiatry*, 60(8):545–549, 1999.
- [50] Kenneth A. Kobak, John H. Greist, James W. Jefferson, James C. Mundt, and David J. Katzelnick. Computerized assessment of depression and anxiety over the telephone using interactive voice response. *MD Computing*, 16(3):64–68, 1999.
- [51] Sunming Cheung, Steven M. Entine, and Jerome H. Klotz. Microcomputer voice-response telephone entry for balanced clinical trial randomization. *Journal of Medical Systems*, 1(2):165–169, 1977.
- [52] David Steel, Joe Vella, and Peter Harrington. Quality issues in telephone

- surveys: Coverage, non-response and quota sampling. *Australian Journal of Statistics*, 38(1):15–34, 1996.
- [53] Lois Oksenberg, Lerita Coleman, and Charles F. Cannell. Interviewers' voices and refusal rates in telephone surveys. *Public Opinion Quarterly*, 50:97–111, 1986.
- [54] Lois Oksenberg and Charles Cannell. Effects of interviewer vocal characteristics on nonresponse. In Robert M. Groves, Paul P. Biemer, Lars E. Lyberg, James T. Massey, William L. II Nicholls, and Joseph Waksberg, editors, *Telephone Survey Methodology*, pages 257–269. John Wiley & Sons, New York, 1988.
- [55] Arpad Barath and Charles F. Cannell. Effect of interviewer's voice intonation. *Public Opinion Quarterly*, 40:370–373, 1976.
- [56] Claire G  linas-Chebat, Jean-Charles Chebat, and Alexander Vaninsky. Voice and advertising: Effects of intonation and intensity of voice on source credibility, attitudes towards the advertised service and the intent to buy. *Perceptual and Motor Skills*, 83(1):243–262, 1996.
- [57] Stephen M. Smith and David R. Shaffer. Celerity and cajolery: Rapid speech may promote or inhibit persuasion through its impact on message elaboration. *Personality and Social Psychology Bulletin*, 17(6):663–669, 1991.
- [58] Norman Miller, Geoffrey Maruyama, Rex Julian Beaber, and Keith Valone. Speed of speech and persuasion. *Journal of Personality and Social Psychology*, 34(4):615–624, 1976.

- [59] Richard A. Page and Joseph L. Balloun. The effect of voice volume on the perception of personality. *The Journal of Social Psychology*, 105:65–72, 1978.
- [60] John D.M. Laver. Voice quality and indexical information. *British Journal of Disorders of Communication*, 3:43–54, 1968.
- [61] R. Platek and G.B. Gray. On the definitions of response rates. *Survey Methodology*, 12(1):17–27, 1986.
- [62] William G. Cochran. *Sampling techniques*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York, third edition edition, 1977.
- [63] J.M. White and R. Humeniuk. Alcohol-related violence: Exploring the relationship. report no. 2. In *National Symposium on Alcohol Misuse and Violence*, Canberra, 1993.
- [64] D.R. English, C.D.J. Holman, E Milne, M.G. Winter, G.K. Hulse, J.P. Codde, C.I. Bower, B. Corti, N. de Klerk, M.W. Knuiman, J.J. Kurinczuk, G.F. Lewin, and G.A. Ryan. *The quantification of drug caused morbidity and mortality in Australia*. Commonwealth Department of Human Services and Health, Canberra, 1995.
- [65] Priya Raghubir and Geeta Menon. Asking sensitive questions: The effects of type of referent and frequency wording in counterbiasing methods. *Psychology and Marketing*, 13(7):633–652, 1996.
- [66] Norman M. Bradburn, Seymour Sudman, Ed Blair, and Carol Stocking.

- Question threat and response bias. *Public Opinion Quarterly*, 42:221–234, 1978.
- [67] John C. Duffy and Jennifer J. Waterton. Under-reporting of alcohol consumption in sample surveys: The effect of computer interviewing in field-work. *British Journal of Addiction*, 79:303–308, 1984.
- [68] John W. Welte and Marcia Russell. Influence of socially desirable responding in a study of stress and substance abuse. *Alcoholism: Clinical and Experimental Research*, 17(4):758–761, 1993.
- [69] Debra Javeline. Response effects in polite cultures: A test of acquiescence in Kazakhstan. *Public Opinion Quarterly*, 63(1):1–28, 1999.
- [70] Catherine E. Ross and John Mirowsky. Socially-desirable response and acquiescence in a cross-cultural survey of mental health. *Journal of Health and Social Behavior*, 25(2):189–197, 1984.
- [71] Michael F. Fleming, Michael Bruno, Kristen Barry, and Norman Fost. Informed consent, deception, and the use of disguised alcohol questionnaires. *Am. J. Drug. Alcohol Abuse*, 15(3):309–319, 1989.
- [72] Leslie G. Carr and Neal Krause. Social status, psychiatric symptomatology, and response bias. *Journal of Health and Social Behavior*, 19(1):86–91, 1978.
- [73] Kevin Clancy and Walter Gove. Sex differences in mental illness: An analysis of response bias in self-reports. *American Journal of Clinical Hypnosis*, 14(3):205–216, 1972.

- [74] William S. Aquilino. Telephone versus face-to-face interviews for household drug surveys. *International Journal of the Addictions*, 27(1):71–91, 1992.
- [75] William S. Aquilino and Leonardo A. Lo Sciuto. Effects of interview mode on self-reported drug use. *Public Opinion Quarterly*, 54:362–395, 1990.
- [76] William G. Axinn. The influence of interviewer sex on responses to sensitive questions in Nepal. *Social Science Research*, 20(3):303–318, 1991.
- [77] Charles F. Cannell and Floyd J. Fowler. Comparison of a self-enumerative procedure and a personal interview: A validity study. *Public Opinion Quarterly*, 27:250–264, 1963.
- [78] John Colombotos. Personal versus telephone interviews: Effect on responses. *Public Health Reports*, 84(9):773–782, 1969.
- [79] Martin David. The validity of income reported by a sample of families who received welfare assistance during 1959. *Journal of the American Statistical Association*, 57:680–685, 1962.
- [80] Barbara S. Dohrenwend, John Colombotos, and Bruce P. Dohrenwend. Social distance and interviewer effects. *Milbank Quarterly*, 47(1, Pt 2): 213–226, 1969.
- [81] Barbara Snell Dohrenwend, John Colombotos, and Bruce P. Dohrenwend. Social distance and interviewer effects. *Public Opinion Quarterly*, 32(3): 410–422, 1968.
- [82] Michael Fendrich and Connie M. Vaughn. Diminished lifetime substance

- use over time: An inquiry into differential underreporting. *Public Opinion Quarterly*, 58:96–123, 1994.
- [83] Dorothy S. Fidler and Richard E. Kleinknecht. Randomized response versus direct questioning: Two data collection methods for sensitive questions. *Psychological Bulletin*, 84:1045–1049, 1977.
- [84] Walter R. Gove and Michael R. Geerken. Response bias in survey of mental health: An empirical investigation. *American Journal of Sociology*, 82(6):1289–1317, 1977.
- [85] John H. Greist, Marjorie H. Klein, Harold P. Erdman, Jane K. Bires, Sandra M. Bass, Paula E. Machtinger, and Dean G. Kresge. Comparison of computer and interviewer-administered versions of the diagnostic interview schedule. *Hospital and Community Psychiatry*, 38(12):1304–1315, 1987.
- [86] Ramon Henson, Charles F. Cannell, and Aleda V. Roth. Effects of interview mode on reporting of moods, symptoms, and need for social approval. *Journal of Social Psychology*, 105(1):123–129, 1978.
- [87] Joseph R. Hochstim. A critical comparison of three strategies of collecting data for household health surveys. *Journal of the American Statistical Association*, 62:976–989, 1967.
- [88] Weldon T. Johnson and John D. Delamater. Response effects in sex surveys. *Public Opinion Quarterly*, 40:165–181, 1976.
- [89] Patricia Kissinger, Janet Rice, Thomas Farley, Shelley Trim, Kayla Jewitt, Victor Margavio, and David H. Martin. Application of computer-assisted

- interviews to sexual behavior research. *American Journal of Epidemiology*, 149(10):950–954, 1999.
- [90] Barbara S. Mensch and Denise B. Kandel. Underreporting of substance use in a national longitudinal youth cohort: Individual and interviewer effects. *Public Opinion Quarterly*, 52:100–124, 1988.
- [91] Torbjørn Moum. Mode of administration and interviewer effects in self-reported symptoms of anxiety and depression. *Social Indicators Research*, 45(1-3):279–318, 1998.
- [92] Stanley Presser and Linda Stinson. Data collection mode and social desirability bias in self-reported religious attendance. *American Sociological Review*, 63(1):137–145, 1998.
- [93] Wendy L. Richman, Sara Kiesler, Suzanne Weisband, and Fritz Drasgow. A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology*, 84(5):754–775, 1999.
- [94] Norbert Schwarz, Fritz Strack, Hans-J. Hippler, and George Bishop. The impact of administration mode on response effects in survey measurement. *Applied Cognitive Psychology*, 5(3):193–212, 1991.
- [95] You Poh Seng. Errors in age reporting in statistically underdeveloped countries. *Population Studies*, 18:164–182, 1964.
- [96] Eleanor Singer and Luane Kohnke-Aguirre. Interviewer expectation effects: A replication and extension. *Public Opinion Quarterly*, 43(2):245–260, 1979.

- [97] Harvey A. Skinner and Barbara A. Allen. Does the computer make a difference? Computerized versus face-to-face self-report assessment of alcohol, drug, and tobacco use. *Journal of Consulting and Clinical Psychology*, 51(2):267–275, 1983.
- [98] Herbert Stember and Herbert Hyman. How interviewer effects operate through question form. *International Journal of Opinion and Attitude Research*, 3(28):493–512, 1949.
- [99] Jennifer J. Waterton and John C. Duffy. A comparison of computer interviewing techniques and traditional methods for the collection of self-report alcohol consumption in a field survey. *International Statistical Review*, 52:173–182, 1984.
- [100] Cynthia Webster. Hispanic and anglo interviewer and respondent ethnicity and gender: The impact on survey response quality. *Journal of Marketing Research*, 33(1):62–72, 1996.
- [101] Ronald L. Akers, James Massey, William Clarke, and Ronald M. Lauer. Are self-reports of adolescent deviance valid? Biochemical measures, randomized response, and the bogus pipeline in smoking behavior. *Social Forces*, 62:234–251, 1983.
- [102] Philip Ash and Edward Abramson. The effect of anonymity on attitude-questionnaire response. *Journal of Abnormal and Social Psychology*, 47:722–723, 1952.
- [103] Stephen M. Corey. Signed versus unsigned attitude questionnaires. *Journal of Educational Psychology*, 28:144–148, 1937.



- [104] John D. DeLamater and Patricia MacCorquodale. The effects of interview schedule variations on reported sexual behavior. *Sociological Methods and Research*, 4(2):215–236, 1975.
- [105] Richard I. Evans, William B. Hansen, and Maurice B. Mittlemark. Increasing the self-reports of smoking behavior in children. *Journal of Applied Psychology*, 62:521–523, 1977.
- [106] Joseph C. Gfroerer and Arthur L. Hughes. The feasibility of collecting drug abuse data by telephone. *Public Health Reports*, 106(4):384–393, 1991.
- [107] M.J. Leutgart and Ann Haberman Armstrong. Methodological issues in drug surveys: Anonymity, recency, and frequency. *International Journal of the Addictions*, 8(4):683–689, 1973.
- [108] Hongjie Liu, Jie Xie, Wenzhou Yu, Weisheng Song, Zhenya Gao, Zhixin Ma, and Roger Detels. A study of sexual behavior among rural residents of China. *Journal of Acquired Immune Deficiency Syndromes*, 19(1):80–88, 1998.
- [109] Toni Makkai and Ian McAllister. Measuring social indicators in opinion surveys: A method to improve accuracy on sensitive questions. *Social Indicators Research*, 27(2):169–186, 1992.
- [110] Janet H. Malvin and Joel M. Moskowitz. Anonymous versus identifiable self-reports of adolescents drug attitudes, intentions, and use. *Public Opinion Quarterly*, 47:557–566, 1983.

- [111] Willard C. Olson. The waiver of signature in personal reports. *Journal of Applied Psychology*, 20:442–450, 1936.
- [112] Thomas P. Petzel, James E. Johnson, and Jack McKillip. Response bias in drug surveys. *Journal of Consulting and Clinical Psychology*, 40(3): 437–439, 1973.
- [113] Eleanor Singer. Informed consent: Consequences for response rate and response quality in social surveys. *American Sociological Review*, 43:144–162, 1978.
- [114] Tom W. Smith. Discrepancies between men and women in reporting numbers of sexual partners: A summary from four countries. *Social Biology*, 39(3–4):203–211, 1992.
- [115] C.F. Turner, L. Ku, S.M. Rogers, L.D. Lindberg, J.H. Pleck, and F.L. Sonenstein. Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science*, 280:867–873, 1998.
- [116] Stephen A. Ayidiya and McKee J. McClendon. Response effects in mail surveys. *Public Opinion Quarterly*, 54:229–247, 1990.
- [117] Carol Fuller. Effect of anonymity on return rate and response bias in a mail survey. *Journal of Applied Psychology*, 59:292–296, 1974.
- [118] James R. Hebert, Yunsheng Ma, Lynn Clemow, Ira S. Ockene, Gordon Saperia, Edward J. Stanek III, Philip A. Merriam, and Judith K. Ockene. Gender differences in social desirability and social approval bias in dietary self-report. *American Journal of Epidemiology*, 146(12):1046–1055, 1997.

- [119] Gavriel Salvendy. Some variables which affect mail survey response rate. *Studia Psychologica*, 18(3):250–252, 1976.
- [120] Jack Siemiatycki. A comparison of mail, telephone, and home interview strategies for household health surveys. *American Journal of Public Health*, 69(3):238–245, 1979.
- [121] Richard C. Wildman. Effects of anonymity and social setting on survey responses. *Public Opinion Quarterly*, 41:74–79, 1977.
- [122] Pamela A. Gillies, Brian Wilcox, Claire Coates, Fanney Kristmundsdóttir, and Donald J. Reid. Use of objective measurement in the validation of self-reported smoking in children aged 10 and 11 years: Saliva thiocyanate. *Journal of Epidemiology and Community Health*, 36:205–208, 1982.
- [123] Lauri Keso and Mikko Salaspuro. Comparative value of self-report and blood tests in assessing outcome amongst alcoholics. *British Journal of Addiction*, 85:209–215, 1990.
- [124] Carl A. Latkin and David Vlahov. Socially desirable response tendency as a correlate of accuracy of self-reported HIV serostatus for HIV seropositive injection drug users. *Addiction*, 93(8):1191–1197, 1998.
- [125] Stephen Magura, Douglas F. Goldsmith, Cathy Casriel, Paul J. Goldstein, and Douglas S. Lipton. The validity of methadone clients' self-reported drug use. *International Journal of the Addictions*, 22:727–749, 1987.
- [126] J.E. Peachey and B.M. Kapur. Monitoring drinking behavior with the alcohol dipstick during treatment. *Alcoholism: Clinical and Experimental Research*, 10:663–666, 1986.

- [127] Aihide Yoshino and Motoichiro Kato. Influence of social desirability response set on self-report for assessing the outcome of treated alcoholics. *Alcoholism: Clinical and Experimental Research*, 19(6):1517–1519, 1995.
- [128] Michael Fendrich, Timothy Johnson, Joseph S. Wislar, and Catherine Nageotte. Accuracy of parent mental health service reporting: Results from a reverse record-check study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38(2):147–155, 1999.
- [129] Robert H. Hardt and Sandra Peterson-Hardt. On determining the quality of the delinquency self-report method. *Journal of Research in Crime and Delinquency*, 14:247–257, 1977.
- [130] Stephen A. Maisto, Linda C. Sobell, and Mark B. Sobell. Comparison of alcoholic's self-reports of drinking behavior with reports of collateral informants. *Journal of Consulting and Clinical Psychology*, 47:106–112, 1979.
- [131] Georg E. Matt, Monica Garcia, Wesley W. Primicias, Lailani Frericks, and Fabio De Faria. Exploring biases in self-reported exercising behavior: Heuristics based on recency, frequency, and preference. *Perceptual and Motor Skills*, 88(1):126–128, 1999.
- [132] Paul L. Erdos and James Regier. Visible vs. disguised keying on questionnaires. *Journal of Advertising Research*, 17(1):13–18, 1977.
- [133] Karl E. Bauman and Clyde W. Dent. Influence on an objective measure on self-reports of behaviour. *Journal of Applied Psychology*, 67:623–628, 1982.

- [134] Pamela Campanelli, T.E. Dielman, and Jean T. Shope. Validity of adolescent's self-reports of alcohol use and misuse using a bogus pipeline procedure. *Adolescence*, 22:7–22, 1987.
- [135] Edward E. Jones and Harold Sigall. The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, 76:349–364, 1971.
- [136] John B. Lowe, Richard A. Windsor, Bryn Adams, Joyce Morris, and Yoland Reese. Use of a bogus pipeline method to increase accuracy of self-reported alcohol consumption among pregnant women. *Journal of Studies on Alcohol*, 47:173–175, 1986.
- [137] David M. Murray, Catherine M. O'Connell, Linda A. Schmid, and Cheryl L. Perry. The validity of smoking self-reports by adolescents: A reexamination of the bogus pipeline procedure. *Addictive Behaviors*, 12:7–15, 1987.
- [138] Roger Tourangeau, Tom W. Smith, and Kenneth A. Rasinski. Motivation to report sensitive behaviors on surveys: Evidence from a bogus pipeline experiment. *Journal of Applied Social Psychology*, 27(3):209–222, 1997.
- [139] Guy Bégin and Michel Boivin. Comparison of data gathered on sensitive questions via direct questionnaire, randomized response technique, and a projective method. *Psychological Reports*, 47(3, Pt 1):743–750, 1980.
- [140] Robert J. Fisher. Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research*, 20:303–315, 1993.
- [141] Myung-Soo Jo, James E. Nelson, and Pamela Kiecker. A model for control-

- ling social desirability bias by direct and indirect questioning. *Marketing Letters*, 8(4):429–437, 1997.
- [142] James H. Frey. An experiment with a confidentiality reminder in a telephone survey. *Public Opinion Quarterly*, 50:267–269, 1986.
- [143] Frederic G. Reamer. Protecting research subjects and unintended consequences: The effect of guarantees of confidentiality. *Public Opinion Quarterly*, 43:497–506, 1979.
- [144] Kathryn M. Woods and J. Regis McNamara. Confidentiality: Its effects on interviewee behavior. *Professional Psychology*, 11(5):714–720, 1980.
- [145] Gilbert Becker and Donald A. Bakal. Subject anonymity and motivational distortion in self-report data. *Journal of Clinical Psychology*, 26:207–209, 1970.
- [146] Ned A. Rosen. Anonymity and attitude measurement. *Public Opinion Quarterly*, 24:675–679, 1960.
- [147] James Carifo and Ronald Biron. Collecting sensitive data anonymously: The CDRGP technique. *Journal of Alcohol and Drug Education*, 23:47–66, 1978.
- [148] Abdel-Latif A. Abul-Ela, Bernard G. Greenberg, and Daniel G. Horvitz. A multiproportions randomized response model. *Journal of the American Statistical Association*, pages 990–1008, 1967.
- [149] Robert F. Boruch. Relations among statistical methods for assuring confidentiality of social research. *Social Science Research*, 1:403–414, 1972.

- [150] Cathy Campbell and Brian L. Joiner. How to get the answer without being sure you've asked the question. *The American Statistician*, 27(5): 229–231, 1973.
- [151] John C. Duffy and Jennifer J. Waterton. Randomized response models for estimating the distribution function of a quantitative character. *International Statistical Review*, 52(2):165–171, 1984.
- [152] Ralph E. Folsom, Bernard G. Greenberg, Daniel G. Horvitz, and James R. Abernathy. The two alternate questions randomized response model for human surveys. *Journal of the American Statistical Association*, 68(343): 525–530, 1973.
- [153] Michael S. Goodstadt, Gaynoll Cook, and Valerie Gruson. The validity of reported drug use: The randomized response technique. *International Journal of the Addictions*, 13:359–367, 1978.
- [154] Bernard G. Greenberg, James R. Abernathy, and Daniel G. Horvitz. A new survey technique and its application in the field of public health. *Milbank Memorial Fund Quarterly*, 48:39–55, 1970.
- [155] Kenneth J. Levy. Reducing the occurrence of omitted or untruthful responses when testing hypotheses concerning populations. *Psychological Bulletin*, 83(5):759–761, 1976.
- [156] James E. Reinmuth and Michael D. Geurts. The collection of sensitive information using a two-stage, randomized response model. *Journal of Marketing Research*, 12:402–407, 1975.
- [157] Stanley L. Warner. Randomized response: A survey technique for elimi-

- nating evasive answer bias. *Journal of the American Statistical Association*, 60:63–69, 1965.
- [158] Paul S. Maxim. *Quantitative Research Methods in the Social Sciences*. Oxford University Press, Oxford, 1999.
- [159] Harold P. Erdman, Marjorie H. Klein, John H. Priest, Stacey S. Skare, Jack J. Husted, Lee N. Robins, John E. Helzer, Evelyn Goldring, Merle Hamburger, and J. Phillip Miller. A comparison of two computer-administered version of the NIMH diagnostic interview schedule. *J Psychiat. Res.*, 26(1):85–95, 1992.
- [160] Lenore Kohlmeier, Michelle Mendez, Jennifer McDuffie, and Mary Miller. Computer-assisted self-interviewing: A multimedia approach to dietary assessment. *American Journal of Clinical Nutrition*, 65(suppl):1275S–1281S, 1997.
- [161] D. Koson, C. Kitchen, M. Kochen, and D. Stodolosky. Psychological testing by computer: Effect on response bias. *Educational and Psychological Measurement*, 30:803–810, 1970.
- [162] Andrew J. Supple, William S. Aquilino, and Debra L. Wright. Collecting sensitive self-report data with laptop computers: Impact on the response tendencies of adolescents in a home interview. *Journal of Research on Adolescence*, 9(4):467–488, 1999.
- [163] Kevin Clancy and Walter Gove. Sex differences in mental illness: An analysis of response bias in self-reports. *American Journal of Sociology*, 80(1):205–216, 1974.



- [164] Frank J. Ivis, Susan J. Bondy, and Edward M. Adlaf. The effect of question structure on self-reports of heavy drinking: Close-ended versus open-ended questions. *Journal of Studies on Alcohol*, 58:622–624, 1997.
- [165] S. Stephen Kegeles, Clinton F. Fink, and John P. Kirscht. Interviewing a national sample by long-distance telephone. *Public Opinion Quarterly*, 33:412–419, 1969.
- [166] Nancy Bates. Data quality issues in a multi-mode census: Results from the mail and telephone mode test (MTMT). In *Proceedings of the Section on Survey Research Methods of the American Statistical Association.*, pages 736–741, 1993.
- [167] Richard J. Fox, Melvin R. Crask, and Kim Jooghoon. Mail survey response rate: A meta-analysis of selected techniques for inducing response. *Public Opinion Quarterly*, 52:467–491, 1988.
- [168] L. M. Sullivan, K. A. Dukes, L. Harris, R. S. Dittus, S. Greenfield, and S. H. Kaplan. A comparison of various methods of collecting self-reported health outcomes data among low-income and minority patients. *Medical Care*, 33(4 Suppl):AS183–94, 1995.
- [169] Ruth E. Little, Charles N. Uhl, Robert F. Labbe, Janis L. Abkowitz, and E. Louise R. Phillips. Agreement between laboratory tests and self-reports of alcohol, tobacco, caffeine, marijuana and other drug use in post-partum women. *Social Science and Medicine*, 22(1):91–98, 1986.
- [170] Lynne Hancock. *Drug use in the Australian community: Prevalence, sociodemographic characteristics of users, and context of use*. Phd, University of Newcastle, Australia, 1991.

- [171] Judith Droitcour, Rachel A. Caspar, Michael L. Hubbard, and Trena M. Ezzati. The item count method technique as a method of indirect questioning: A review of its development and a case study application. In Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz, and Seymour Sudman, editors, *Measurement Errors in Surveys*, pages 185–210. John Wiley & Sons, New York, 1991.
- [172] Robert G. Orwin and Robert F. Boruch. RRT meets RDD: Statistical strategies for assuring response privacy in telephone surveys. *Public Opinion Quarterly*, 46:560–571, 1982.
- [173] Australian Bureau of Statistics. *Household Use of Information Technology, Australia (Catalogue 8146.0)*. Australian Bureau of Statistics, Canberra, 1998.
- [174] Monash University. General support, 23 May 2002 2002. Retrieved 23 May 2002 from: [http://www.celts.monash.edu.au/html/general\\\_support.html](http://www.celts.monash.edu.au/html/general\_support.html).
- [175] Monash University. The information, 24 May 1997 2002. Retrieved 2002 from: <http://www.lib.monash.edu.au/libnews/issue24.htm>.
- [176] University of Melbourne. Exam results on the web head for another record, 2002. Retrieved 23 May 2002 from: <http://www.unimelb.edu.au/ExtRels/Media/UN/archive/1997/220/examresult%sonthewebheadfor.html>.
- [177] B. Bonevski, R.W. Sanson-Fisher, Elizabeth M. Campbell, and Malcolm C. Ireland. Do general practice patients find computer health risk

- surveys acceptable? A comparison with pen-and-paper method. *Health Promotion Journal of Australia*, 7(2):100–106, 1997.
- [178] R.W. Millard and J.R. Carver. Cross-sectional comparison of live and interactive voice recognition administration of the SF-12 health status survey. *American Journal of Managed Care*, 5(2):153–159, 1999.
- [179] Anna M. Bardone, Dean D. Krahm, Brian M. Goodman, and John S. Searles. Using interactive voice response technology and timeline follow-back methodology in studying binge eating and drinking behavior: Different answers to different forms of the same question? *Addictive Behaviors*, 25(1):1–11, 2000.
- [180] John S. Searles, John E. Helzer, and Daniel E. Walter. Comparison of drinking patterns measured by daily reports and timeline follow back. *Psychology of Addictive Behaviors*, 14(3):277–286, 2000.
- [181] Sonia Alemagno, Scott Frank, Maghboeba Mosavel, and John Butts. Screening adolescents for health risks using interactive voice response technology: An evaluation. *Computers in Human Services*, 15(4):27–37, 1998.
- [182] Robert H. Friedman, Lewis E. Kazis, Alan Jette, Mary Beth Smith, John Stollerman, Jeanne Torgerson, and Kathleen Carey. A telecommunications system for monitoring and counseling patients with hypertension. impact on medication adherence and blood pressure control. *American Journal of Hypertension*, 9(4 Pt 1):285–292, 1996.
- [183] Richard D. Havel and Michael P. Wright. Automated interviewing for hepatitis B risk assessment and vaccination referral. *American Journal of Preventive Medicine*, 13(5):392–395, 1997.

- [184] Von O. Leirer, Elizabeth Decker Tanke, and Daniel G. Morrow. Automated telephone reminders for improving ambulatory care services. *Journal of Ambulatory Care Management*, 15(4):54–62, 1992.
- [185] SAS Institute Inc. *SAS/STAT User's Guide, Version 6, Fourth Edition, Volumes 1 and 2*. SAS Institute Inc., Cary, 8 edition, 1999.
- [186] SAS Institute Inc. *SAS/AF Software: Usage and Reference, Version 6, First Edition*. SAS Institute Inc., Cary, NC, first edition edition, 1991.
- [187] Artisoft Inc. Visual Voice Pro, (Version 5.0) [Computer software], 1998.
- [188] Richard A. Kulka and Michael F. Weeks. Towards the development of optimal calling protocols for telephone surveys: A conditional probabilities approach. *Journal of Official Statistics*, 4(4):319–332, 1988.
- [189] Telstra Corporation Limited. Telstra public switched telephone service (PSTS) section of the standard form of agreement, 2001. Retrieved 2001 from: <http://www.telstra.com.au/sfoa/>.
- [190] Australian Surveying and Land Information Group. Distance between two points, 2000. Retrieved 2 May 2001 from: <http://www.auslig.gov.au/geodesy/datums/distance.htm>.
- [191] SAS Institute Inc. *SAS/SHARE Software: Usage and Reference, Version 6, First Edition*. SAS Institute Inc., Cary, NC, first edition edition, 1991.
- [192] Harry Newton. *Newton's Telecom Dictionary*. Miller Freeman, Inc., New York, 15th. edition, 1999.
- [193] Marco Piccinelli, Elisabetta Tessari, Marco Bortolomasi, Orazio Piasere, Massimo Semenzin, Nicola Garzotto, and Michele Tansella. Efficacy of the

- alcohol use disorders identification test as a screening tool for hazardous alcohol intake and related disorders in primary care: A validity study. *British Medical Journal*, 314(7078):420–424, 1997.
- [194] Thomas F. Babor, Juan Ramon de la Fuente, John Saunders, and Marcus Grant. AUDIT. The alcohol use disorders identification test: Guidelines for use in primary medical care. Technical Report WHO/PSA/92.4, World Health Organization, 1992.
- [195] John B. Saunders, Olaf G. Aasland, Thomas F. Babor, Juan R. de la Fuente, and Marcus Grant. Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption — II. *Addiction*, 88:791–804, 1993.
- [196] Robert M. Schumacher, Mary L. Hardzinski, and Amy L. Schwartz. Increasing the usability of interactive voice response systems: Research and guidelines for phone-based interfaces. *Human Factors*, 37(2):251–264, 1995.
- [197] Graham Kalton. Survey sampling. In Samuel Kotz and Norman L. Johnson, editors, *Encyclopedia of Statistical Sciences*, volume 9, pages 111–119. John Wiley & Sons, New York, 1988.
- [198] Richard J. Rosen, Richard L. Clayton, and Lynda L. Pivetz. Converting mail reporters to touchtone data entry. In *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, pages 763–768. The American Statistical Association, Toronto, Canada, 1994.

- [199] William L. Nicholls II and Martin V. Appel. New CASIC technologies at the U.S. Bureau of the Census. In *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, pages 757–762, Toronto, Canada, 1994.
- [200] David O’Connell, Richard J. Rosen, and Richard L. Clayton. Long-term results of touchtone data entry in the current employment statistics survey program. In *Proceedings of the Section on Survey Research Methods of the American Statistical Association.*, pages 764–767. American Statistical Association, 1996.
- [201] JRB Software Inc. JRButils [Computer software], 2001.
- [202] Richard Halstead-Nussloch. The design of phone-based interfaces for consumers. In *CHI’89 Proceedings*, pages 347–352. Association for Computing Machinery, 1989.
- [203] Chris S. Craig. Goldwave (Version 4.02) [computer software], 1998.
- [204] Ray Meddis. *Statistics Using Ranks. A Unified Approach*. Basil Blackwell, Oxford, 1984.
- [205] Desktop Marketing Systems Pty Ltd. Marketing Pro (August 1999 edition) [computer data file], 1999.
- [206] Australian Bureau of Statistics. *Statistical Geography, Volume 1, Australian Standard Geographical Classification (ASGC) (Catalogue no. 1216.0)*. Australian Bureau of Statistics, Canberra, 2001.
- [207] Australian Surveying & Land Information Group. Interactive place name

- search, 2001. Retrieved 31 May 2001 from: <http://www.auslig.gov.au/mapping/names/names.htm>.
- [208] Australian Bureau of Statistics. 1996 Census of Population and Housing. Basic community profiles. In *AusStats [electronic resource]*. Available: <http://www.abs.gov.au/ausstats>. Australian Bureau of Statistics, Canberra, 1996.
- [209] Paul J. Lavrakas, Sandra L. Bauman, and Daniel M. Merkle. The last-birthday selection method & within-unit coverage problems. In *Proceedings of the Section on Survey Research Methods of the American Statistical Association.*, pages 1107–1112. American Statistical Association, 1993.
- [210] Lawrence M. Leemis and Kishor S. Trivedi. A comparison of approximate interval estimators for the Bernoulli parameter. *The American Statistician*, 50(1):63–68, 1996.
- [211] Frederick Wiseman and Philip McDonald. Non contact and refusal rates in consumer telephone surveys. *Journal of Marketing Research*, 17:478–484, 1979.
- [212] Department of Health and Aged Care. *Measuring Remoteness: Accessibility/Remoteness Index of Australia (ARIA) (Occasional paper no. 6)*. Department of Health & Aged Care, Canberra, 1999.
- [213] Pramod Adhikari and Amber Summerhill. 1998 National Drug Strategy: Household survey: Detailed findings. Drug Statistics Series No. 6. AIHW catalogue no. PHE 27, Australian Institute of Health & Welfare, 2000.

- [214] Joseph L. Fleiss. *Statistical Methods for Rates and Proportions*. Wiley, New York, 1973.
- [215] Desktop Marketing Systems Pty Ltd. Marketing Pro (July 2000 edition) [computer data file], 2000.
- [216] Larry D. Rosen, Deborah C. Sears, and Michelle M. Weil. Computerphobia. *Behavior Research Methods, Instruments, and Computers*, 19(2): 167–179, 1987.
- [217] Robert M. Groves. *Survey Errors and Survey Costs*. John Wiley & Sons, New York, 1989.
- [218] Don J. Bennett and David Steel. An evaluation of a large-scale CATI household survey using random digit dialling. *The Australian and New Zealand Journal of Statistics*, 42(3):255–270, 2000.
- [219] Australian Institute of Health and Welfare. *National Drug Strategy: Household Survey: Technical Report 1998*. Australian Institute of Health & Welfare, Department of Health & Aged Care, Canberra, 1999.
- [220] Frank J. Ivis, Edward M. Adlaf, and Jürgen Rehm. Incorporating the AUDIT into a general population telephone survey: A methodological experiment. *Drug and Alcohol Dependence*, 60(1):97–104, 2000.
- [221] Larry D. Rosen and Michelle M. Weil. Adult and teenage use of consumer, business, and entertainment technology: Potholes on the information superhighway? *Journal of Consumer Affairs*, 29(1):55–84, 1995.
- [222] Ed Blair, Seymour Sudman, Norman M. Bradburn, and Carol Stocking. How to ask questions about drinking and sex: Response effects in mea-



- suring consumer behavior. *Journal of Marketing Research*, 14:316–321, 1977.
- [223] James E. Gentle. Monte carlo methods. In Samuel Kotz and Normal L. Johnson, editors, *Encyclopedia of Statistical Sciences*, volume 5, pages 612–617. John Wiley & Sons, New York, 1985.
- [224] Leslie Kish. *Survey Sampling*. John Wiley, New York, 1995.
- [225] George W. Snedecor and William G. Cochran. *Statistical Methods*. The Iowa State University Press, Ames, Iowa, seventh edition, 1980.
- [226] B.S. Everitt. *The Analysis of Contingency Tables*. Chapman & Hall, London, 1992.
- [227] American Association for Public Opinion Research. Standard definitions: Final dispositions of case codes and outcome rates for surveys. Technical report, American Association for Public Opinion Research, 2000.
- [228] David F. Signorini. Sample size for Poisson regression. *Biometrika*, 78(2): 446–450, 1991.
- [229] F.B. Alt. Bonferroni inequalities and intervals. In Samuel Kotz and Normal L. Johnson, editors, *Encyclopedia of Statistical Sciences*, volume 1, pages 294–301. John Wiley & Sons, New York, 1982.
- [230] Maura E. Stokes, Charles S. Davis, and Gary G. Koch. *Categorical Data Analysis Using the SAS System*. SAS Institute Inc., Cary, NC, second edition, 2000.
- [231] Barry H. Margolin. Test for trend in proportions. In Samuel Kotz and

- Normal L. Johnson, editors, *Encyclopedia of Statistical Sciences*, volume 9, pages 334–336. John Wiley & Sons, New York, 1988.
- [232] Dagmar Krebs. Non-response to sensitive questions: Nationalism in Germany. In *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, pages 1199–1203. American Statistical Association, Toronto, Canada, 1994.
- [233] Jared B. Jobe, William F. Pratt, Roger Tourangeau, Alison K. Baldwin, and Kenneth A. Rasinski. Effects of interview mode on sensitive questions in a fertility survey. In Lars E. Lyberg, Paul P. Beimer, Martin Collins, Edith de Leeuw, Cathryn S. Dippo, Norbert Schwarz, and Dennis Trewin, editors, *Survey Measurement and Process Quality*, pages 311–329. John Wiley & Sons, New York, 1997.
- [234] William S. Aquilino. Interview mode effects in surveys of drug and alcohol use: A field experiment. *Public Opinion Quarterly*, 58:210–240, 1994.
- [235] Roger Tourangeau, Jobe B. Jobe, William F. Pratt, and Kenneth Rasinski. Design and results of the women’s health study. In *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, pages 49–58. The American Statistical Association, Toronto, Canada, 1994.
- [236] James Katz, Philip Aspden, and Warren A. Reich. Public attitudes toward voice-based electronic messaging technologies in the United States: A national survey of opinions about voice response units and telephone answering machines. *Behaviour and Information Technology*, 16(3):125–144, 1997.

- [237] Edward McQuarrie and Kichiro Iwamoto. Public opinion toward computers as a function of exposure. *Social Science Computer Review*, 8(2): 221–233, 1990.
- [238] Larry D. Rosen and Michelle M. Weil. What we have learned from a decade of research (1983–1993) on “The psychological impact of technology”. *Computers and Society*, 24:3–10, 1994.
- [239] M.S. Coates and E.A. Tracey. Cancer in New South Wales. Incidence and mortality, 1997., 2000. Retrieved 22 Jan 2001 from: <http://www.nswcc.org.au/pages/ccic/stats/index.htm>.
- [240] Anne Kricker, Jane Bell, Marylon Coates, Richard Taylor, and Bruce Armstrong. *Cancer of the cervix in NSW in 1972 to 1992*. NSW Cancer Council, Sydney, 1996.
- [241] Jeanne S. Mandelblatt, William F. Lawrence, Sharita Mizell Womack, Denise Jacobson, Bin Yi, Yi-ting Hwang, Karen Gold, James Barter, and Keerti Shah. Benefits and costs of using HPV testing to screen for cervical cancer. *Journal of the American Medical Association*, 287(18):2372–2381, 2002.
- [242] F.X. Bosch, M.M. Manos, N. Munoz, M. Sherman, A.M. Jansen, J. Peto, M.H. Schiffman, V. Moreno, R. Kurman, and K.V. Shah. Prevalence of human papillomavirus in cervical cancer: A worldwide perspective. International biological study on cervical cancer (IBSCC) study group. *Journal of the National Cancer Institute*, 87:796–802, 1995.
- [243] G.N. Papanicolaou and H.F. Traut. The diagnostic value of vaginal smears

- in carcinoma of the uterus. *American Journal of Obstetrics and Gynecology*, 42:193–206, 1941.
- [244] G.N. Papanicolaou and H.F. Traut. *Diagnosis of Uterine Cancer by Vaginal Smear*. The Commonwealth Fund, New York, 1943.
- [245] Commonwealth Department of Health and Family Services. *Screening for the prevention of cervical cancer*. Australian Government Publishing Service, Canberra, 1998.
- [246] Jeanne Mandelblatt, Harold Freeman, Deidre Winczewski, Kate Cagney, Sterling Williams, Reynold Trowers, Jian Tang, Karen Gold, Ting Hsiang Lin, and Jon Kerner. The costs and effects of cervical and breast cancer screening in a public hospital emergency room. *American Journal of Public Health*, 87(7):1182–1189, 1997.
- [247] Lynne Hancock, Rob Sanson-Fisher, and Lynne Kentish. Cervical cancer screening in rural NSW: Health Insurance Commission data compared to self-report. *Australian and New Zealand Journal of Public Health*, 22(3):307–312, 1998.
- [248] N.P. Gordon, T.G. Rundall, and L. Parker. Type of health care coverage and the likelihood of being screened for cancer. *Medical Care*, 36(5):636–645, 1998.
- [249] L. Suarez, R.A. Roche, D. Nichols, and D.M. Simpson. Knowledge, behaviour, and fears concerning breast and cervical cancer among low-income Mexican-American women. *American Journal of Preventive Medicine*, 13(2):137–142, 1997.

- [250] L.J. Burak and M. Meyer. Using the health belief model to examine and predict college women's cervical cancer screening beliefs and behavior. *Health Care for Women International*, 18(3):251–262, 1997.
- [251] H. Mitchell, G. Medley, and V. Higgins. An audit of women who died during 1994 from cancer of the cervix in victoria, Australia. *Australian and New Zealand Journal of Obstetrics and Gynaecology*, 36(1):73–76, 1996.
- [252] NSW Cervical Screening Program. *Screening rate data for NSW for the period ending 30 September 2000*. NSW Cervical Screening Program, Sydney, 2000.
- [253] Hassan Mamoon, Richard Taylor, Stephen Morrell, Gerard Wain, and Helen Moore. Cervical screening: Population-based comparisons between self-reported survey and registry-derived Pap test rates. *Australian and New Zealand Journal of Public Health*, 25(6):505–510, 2001.
- [254] Julie E. Byles, Rob Sanson-Fisher, Selina Redman, James A. Dickinson, and Stephen Halpin. Effectiveness of three community based strategies to promote screening for cervical cancer. *Journal of Medical Screening*, 1: 150–158, 1994.
- [255] T. Lancaster, W. Dobbie, K. Vos, P. Yudkin, M. Murphy, and G. Fowler. Randomized trial of nurse-assisted strategies for smoking cessation in primary care. *British Journal of General Practice*, 49(440):191–194, 1999.
- [256] R. Cordoba, M. T. Delgado, V. Pico, R. Altisent, D. Fores, A. Monreal, O. Frisas, and A. Lopez del Val. Effectiveness of brief intervention on

- non-dependent alcohol drinkers (EBIAL): A Spanish multi-centre study. *Family Practice*, 15(6):562–568, 1998.
- [257] Y. Tomson, A. Romelsjo, and H. Aberg. Excessive drinking—brief intervention by a primary health care nurse. A randomized controlled trial. *Scandinavian Journal of Primary Health Care*, 16(3):188–192, 1998.
- [258] S. M. Bacha and C. F. Rispoli. Myofunctional therapy: Brief intervention. *International Journal of Orofacial Myology*, 25:37–47, 1999.
- [259] G. Chang, L. Wilkins-Haug, S. Berman, and M. A. Goetz. Brief intervention for alcohol use in pregnancy: A randomized trial. *Addiction*, 94(10):1499–1508, 1999.
- [260] K. Evans, P. Tyrer, J. Catalan, U. Schmidt, K. Davidson, J. Dent, P. Tata, S. Thornton, J. Barber, and S. Thompson. Manual-Assisted Cognitive-Behaviour Therapy (MACT): A randomized controlled trial of a brief intervention with bibliotherapy in the treatment of recurrent deliberate self-harm. *Psychological Medicine*, 29(1):19–25, 1999.
- [261] L. M. Gentilello, F. P. Rivara, D. M. Donovan, G. J. Jurkovich, E. Daranciang, C. W. Dunn, A. Villaveces, M. Copass, and R. R. Ries. Alcohol interventions in a trauma center as a means of reducing the risk of injury recurrence. *Annals of Surgery*, 230(4):473–483, 1999.
- [262] L. Irvine, I. K. Crombie, R. A. Clark, P. W. Slane, C. Feyerabend, K. E. Goodman, and J. I. Cater. Advising parents of asthmatic children on passive smoking: Randomised controlled trial. *British Medical Journal*, 318(7196):1456–1459, 1999.

- [263] P. M. Monti, S. M. Colby, N. P. Barnett, A. Spirito, D. J. Rohsenow, M. Myers, R. Woolard, and W. Lewander. Brief intervention for harm reduction with alcohol-positive older adolescents in a hospital emergency department. *Journal of Consulting and Clinical Psychology*, 67(6):989–994, 1999.
- [264] R. Mullins, P. Livingston, and R. Borland. A strategy for involving general practitioners in smoking control. *Australian and New Zealand Journal of Public Health*, 23(3):249–251, 1999.
- [265] M. Thome and B. Alder. A telephone intervention to reduce fatigue and symptom distress in mothers with difficult infants in the community. *Journal of Advanced Nursing*, 29(1):128–137, 1999.
- [266] M. Aalto, R. Saksanen, P. Laine, R. Forsstrom, M. Raikaa, M. Kiviluoto, K. Seppa, and P. Sillanaukee. Brief intervention for female heavy drinkers in routine general practice: A 3-year randomized, controlled study. *Alcoholism: Clinical and Experimental Research*, 24(11):1680–1686, 2000.
- [267] B. Borsari and K. B. Carey. Effects of a brief motivational intervention with college student drinkers. *Journal of Consulting and Clinical Psychology*, 68(4):728–733, 2000.
- [268] D. K. Freeborn, M. R. Polen, J. F. Hollis, and R. A. Senft. Screening and brief intervention for hazardous drinking in an HMO: Effects on medical care utilization. *Journal of Behavioral Health Services and Research*, 27(4):446–453, 2000.
- [269] R. E. Glasgow, E. P. Whitlock, E. G. Eakin, and E. Lichtenstein. A brief smoking cessation intervention for women in low-income planned

- parenthood clinics. *American Journal of Public Health*, 90(5):786–789, 2000.
- [270] E. Lang, M. Engeland, and T. Brooke. Report of an integrated brief intervention with self-defined problem cannabis users. *Journal of Substance Abuse Treatment*, 19(2):111–116, 2000.
- [271] L. B. Manwell, M. F. Fleming, M. P. Mundt, E. A. Stauffacher, and K. L. Barry. Treatment of problem alcohol use in women of childbearing age: Results of a brief intervention trial. *Alcoholism: Clinical and Experimental Research*, 24(10):1517–1524, 2000.
- [272] J. E. Moore, M. Von Korff, D. Cherkin, K. Saunders, and K. Lorig. A randomized trial of a cognitive-behavioral program for enhancing back pain self care in a primary care setting. *Pain*, 88(2):145–153, 2000.
- [273] G. Paton-Simpson, I. R. McCormick, A. Powell, P. Adams, and D. Bunbury. Problem drinking profiles of patients presenting to general practitioners: Analysis of alcohol use disorders identification test (AUDIT) scores for the Auckland area. *New Zealand Medical Journal*, 113(1105):74–77, 2000.
- [274] R. Richmond, L. Kehoe, N. Heather, and A. Wodak. Evaluation of a workplace brief intervention for excessive alcohol consumption: The workscreen project. *Preventive Medicine*, 30(1):51–63, 2000.
- [275] M. M. Black, E. H. Siegel, Y. Abel, and M. E. Bentley. Home and videotape intervention delays early complementary feeding among adolescent mothers. *Pediatrics*, 107(5):E67, 2001.



- [276] R. D. Blondell, S. W. Looney, A. P. Northington, M. E. Lasch, S. B. Rhodes, and R. L. McDaniels. Can recovering alcoholics help hospitalized patients with alcohol problems? *Journal of Family Practice*, 50(5):447–447, 2001.
- [277] A. Peleg, L. Neumann, M. Friger, R. Peleg, and A. D. Sperber. Outcomes of a brief alcohol abuse prevention program for Israeli high school students. *Journal of Adolescent Health*, 28(4):263–269, 2001.
- [278] Jenny Bowman, Rob Sanson-Fisher, Catherine Boyle, Stephanie Pope, and Sally Redman. A randomised controlled trial of strategies to prompt attendance for a Pap smear. *Journal of Medical Screening*, 2(4):211–218, 1995.
- [279] Julie E. Byles and Rob W. Sanson-Fisher. Mass mailing campaigns to promote screening for cervical cancer: Do they work, and do they continue to work? *Australian and New Zealand Journal of Public Health*, 20(3):254–260, 1996.
- [280] Julie E. Byles, Selina Redman, Robert W. Sanson-Fisher, and Catherine A. Boyle. Effectiveness of two direct-mail strategies to encourage women to have cervical (Pap) smears. *Health Promotion International*, 10(1):5–16, 1995.
- [281] Jeanette E. Ward, Kate Boyle, Selina Redman, and Robert W. Sanson-Fisher. Increasing women’s compliance with opportunistic cervical cancer screening: A randomized trial. *American Journal of Preventive Medicine*, 7(5):285–291, 1991.

- [282] Jeanette Ward and Elizabeth M. Proude. Evaluation of doctors' reminders in emergency departments to encourage cervical screening. *Australian and New Zealand Journal of Public Health*, 23(1):95–98, 1999.
- [283] Lynne Hancock, Rob Sanson-Fisher, Janice Perkins, Ross Corkrey, Robert Burton, and Sandy Reid. Effect of a community action program on cervical cancer screening rates in rural Australian towns: The CART project. *Preventive Medicine*, 32:109–117, 2001.
- [284] Jeanette Ward, Neil Donnelly, and Phoebe Holt. Impact in general practice of the policies of the organised approach to preventing cancer of the cervix. *Australian and New Zealand Journal of Public Health*, 22(3):336–341, 1998.
- [285] J. Ward and R. Sanson-Fisher. Cervical cancer screening in general practice: Standards of care, barriers and strategies for change. *Cancer Forum*, 14(3):149–155, 1990.
- [286] Margaret Lesjak, Myna Hua, and Jeanette Ward. Cervical screening among immigrant Vietnamese women seen in general practice: Current rates, predictors and potential recruitment strategies. *Australian and New Zealand Journal of Public Health*, 23(2):168–173, 1999.
- [287] Jeanette Ward and Rob Sanson-Fisher. Accuracy of women's recall of opportunistic recruitment for cervical cancer screening in general practice. *Australian and New Zealand Journal of Public Health*, 21(3):335–336, 1997.
- [288] Elizabeth Campbell, Danna Peterkin, Richard Abbott, and John Rogers. Encouraging underscreened women to have cervical cancer screening: The

- effectiveness of a computer strategy. *Preventive Medicine*, 26:801–807, 1997.
- [289] Heather Mitchell, Sheila Hirst, Julie Anne Mitchell, Margaret Staples, and Nicolette Torcello. Effect of ethnic media on cervical cancer screening rates. *Australian and New Zealand Journal of Public Health*, 21(3):265–267, 1997.
- [290] Barbara Adamson and Ronda Taylor. *Cervical cancer screening: Knowledge, attitudes and practices of Vietnamese women working in factories*. Women’s Health Studies Discussion Papers and Research Reports. The University of Sydney, Sydney, 1990.
- [291] Allan P. Frank, Michael G. Wandell, Milton D. Headings, Marcus A. Conant, George E. Woody, and Catherine Michel. Anonymous HIV testing using home collection and telemedicine counselling: A multicenter evaluation. *Archives of Internal Medicine*, 157(3):309–315, 1997.
- [292] Tracey A. Lieu, Angela M. Capra, James Makol, Steven B. Black, and Henry R. Shinefield. Effectiveness and cost-effectiveness of letters, automated telephone messages, or both for underimmunized children in a health maintenance organization. *Pediatrics [On-line]*, 101(4):  
<http://www.pediatrics.org/cgi/content/full/101/4/E3>, 1998.
- [293] John D. Piette, Morris Weinberger, Stephen J. McPhee, Connie A. Mah, Frederic B. Kraemer, and Lawrence M. Crapo. Do automated calls with nurse follow-up improve self-care and glycemic control among vulnerable patients with diabetes? *American Journal of Medicine*, 108(1):20–27, 2000.

- [294] John D. Piette, Morris Weinberger, and Stephen J. McPhee. The effect of automated calls with telephone nurse follow-up on patient-centered outcomes of diabetes care: A randomized controlled trial. *Medical Care*, 38(2):218–230, 2000.
- [295] Australian Bureau of Statistics. *National Localities Index, Australia (Catalogue 1252.0)*. Australian Bureau of Statistics, Canberra, 2000.
- [296] R. Branoff, K. Santi, J. K. Campbell, R. Roetzheim, and M. Oler. A family practice residency cervical screening project: Perceived screening barriers. *Family Medicine*, 29(2):119–123, 1997.
- [297] L. R. Chavez, F. A. Hubbell, S. I. Mishra, and R. B. Valdez. The influence of fatalism on self-reported use of Papanicolaou smears. *American Journal of Preventive Medicine*, 13(6):418–424, 1997.
- [298] D. M. Elnicki, D. K. Morris, and W. T. Shockcor. Patient-perceived barriers to preventive health care among indigent, rural Appalachian patients. *Archives of Internal Medicine*, 155(4):421–424, 1995.
- [299] A. M. Navarro, K. L. Senn, R. M. Kaplan, L. McNicholas, M. C. Campo, and B. Roppe. Por la vida intervention model for cancer prevention in Latinas. *Journal of the National Cancer Institute. Monographs*, 18:137–145, 1995.
- [300] Elizabeth J. Rankow and Irene Tessaro. Cervical cancer risk and Papanicolaou screening in a sample of lesbian and bisexual women. *Journal of Family Practice*, 47(2):139–143, 1998.
- [301] T. L. Skaer, L. M. Robison, D. A. Sclar, and G. H. Harding. Knowledge,

- attitudes, and patterns of cancer screening: A self-report among foreign born hispanic women utilizing rural migrant health clinics. *Journal of Rural Health*, 12(3):169–177, 1996.
- [302] T. G. Hislop, H. F. Clarke, M. Deschamps, R. Joseph, P. R. Band, J. Smith, N. Le, and R. Atleo. Cervical cytology screening. How can we improve rates among first nations women in urban British Columbia? *Canadian Family Physician*, 42:1701–1708, 1996.
- [303] Stephen J. McPhee, Joyce A. Bird, Thomas Davis, Ngoc-The Ha, Christopher N. Jenkins, and Bich Le. Barriers to breast and cervical cancer screening among Vietnamese-American women. *American Journal of Preventive Medicine*, 13(3):205–213, 1997.
- [304] Ruth K. Peters, Moraye B. Bear, and Duncan Thomas. Barriers to screening for cancer of the cervix. *Preventive Medicine*, 18(1):133–146, 1989.
- [305] A. Seow, M. L. Wong, W. C. Smith, and H. P. Lee. Beliefs and attitudes as determinants of cervical cancer screening: A community-based study in Singapore. *Preventive Medicine*, 24(2):134–141, 1995.
- [306] K. M. Jennings. Getting a Pap smear: Focus group responses of African American and latina women. *Oncology Nursing Forum*, 24(5):827–835, 1997.
- [307] Sandra L. Welner. Screening issues in gynecologic malignancies for women with disabilities: Critical considerations. *Journal of Womens Health*, 7(3):281–285, 1998.
- [308] Margaret I. Fitch, Marlene Greenberg, Maureen Cava, Donna Spaner, and

- Karmel Taylor. Exploring the barriers to cervical screening in an urban Canadian setting. *Cancer Nursing*, 21(6):441–449, 1998.
- [309] Paula M. Lantz, Margaret E. Weigers, and James S. House. Education and income differentials in breast and cervical cancer screening. policy implications for rural women. *Medical Care*, 35(3):219–236, 1997.
- [310] C. K. Yu and J. Rymer. Women’s attitudes to and awareness of smear testing and cervical cancer. *British Journal of Family Planning*, 23(4):127–133, 1998.
- [311] Jill Cockburn, Victoria M. White, Sheila Hirst, and David Hill. Barriers to cervical screening in older women. *Australian Family Physician*, 21(7):973–978, 1992.
- [312] M. Deschamps, P. R. Band, T. G. Hislop, H. F. Clarke, J. M. Smith, and V. To Yee Ng. Barriers to cervical cytology screening in native women in british columbia. *Cancer Detection and Prevention*, 16(5-6):337–339, 1992.
- [313] L. P. Larsen and F. Olesen. Women’s knowledge of and attitude towards organized cervical smear screening. *Acta Obstetricia et Gynecologica Scandinavica*, 77(10):988–996, 1998.
- [314] G. R. Najem, F. Batuman, and A. M. Smith. Papanicolaou test status among inner-city adolescent girls. *American Journal of Preventive Medicine*, 12(6):482–486, 1996.
- [315] C. Lerman, C. Caputo, and D. Brody. Factors associated with inadequate

- cervical cancer screening among lower income primary care patients. *Journal of the American Board of Family Practice*, 3(3):151–156, 1990.
- [316] E. D. Paskett, J. Rushing, R. D’Agostino, C. Tatum, and R. Velez. Cancer screening behaviors of low-income women: The impact of race. *Womens Health*, 3(3-4):203–226, 1997.
- [317] Australian Bureau of Statistics. *Disability, New South Wales (Catalogue 4443.1)*. Australian Bureau of Statistics, Canberra, 2001.
- [318] Aye Thidar Htun, Lynne Hancock, Lucy Bates, Ross Corkrey, and Sue Green. The acceptability of a multi-channel community-based cervical screening intervention: A pilot for regional dissemination. *Unpublished manuscript*, 2001.
- [319] Quinn McNemar. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2):153–157, 1947.
- [320] Sally Anne Newell. *Assessing the effectiveness of personal health record booklets at increasing preventative and screening behaviours for cancer and cardiovascular disease*. Phd, University of Newcastle, 1995.
- [321] Jacob Cohen. *Statistical Power Analysis for the Behavioral Sciences*. Academic Press, New York, 1969.
- [322] S.D. Walter, E.A. Clarke, J. Hatcher, and L.W. Stitt. A comparison of physician and patient reports of Pap smear histories. *Journal of Clinical Epidemiology*, 41(4):401–410, 1988.
- [323] Robert Michielutte, Mark B. Dignan, H. Bradley Wells, Judy Bahnson, Marie Smith, Robert Wooten, and Lynn N. Hale. Errors in reporting

- cervical screening among public health clinic patients. *Journal of Clinical Epidemiology*, 44(4/5):403–408, 1991.
- [324] John A. Sawyer, Jo Anne Earp, Robert H. Fletcher, Fedora F. Daye, and Tonja M. Wynn. Accuracy of women’s self-report of their last Pap smear. *American Journal of Public Health*, 79(8):1036–1037, 1989.
- [325] Jennifer A. Bowman, Selina Redman, James A. Dickinson, Robert Gibberd, and Rob W. Sanson-Fisher. The accuracy of Pap smear utilization self-report: A methodological consideration in cervical screening research. *Health Services Research*, 26(1):98–107, 1991.
- [326] Matthew T. McKenna, Katherine Mallin, Marjorie Speers, and Richard Warnecke. Agreement between patient self-reports and medical records for Pap smear histories. *American Journal of Preventive Medicine*, 8(5):287–291, 1992.
- [327] Nancy P. Gordon, Robert A. Hiatt, and Dominique I. Lampert. Concordance of self-reported data and medical record audit for six cancer screening procedures. *Journal of the National Cancer Institute*, 85(7):566–570, 1993.
- [328] Thomas E. Kottke, Mary Alice Trapp, Maria M. Fores, Ann W. Kelly, Sin-ho Jung, Paul J. Novotny, and Laurel A. Panser. Cancer screening behaviors and attitudes of women in southeastern Minnesota. *Journal of the American Medical Association*, 273(14):1099–1105, 1995.
- [329] Daniel E. Montaña and William R. Phillips. Cancer screening by primary care physicians: A comparison of rates obtained by physician self-report,



- patient survey, and chart audit. *American Journal of Public Health*, 85 (6):795–800, 1995.
- [330] Lucina Suarez, Daniel A. Goldman, and Nancy S. Weiss. Validity of Pap smear and mammogram in a low-income hispanic population. *American Journal of Preventive Medicine*, 11(2):94–98, 1995.
- [331] Jennifer Ann Bowman, Rob Sanson-Fisher, and Sally Redman. The accuracy of self-reported Pap smear utilisation. *Social Science and Medicine*, 44(7):969–976, 1997.
- [332] Paul G. McGovern, Nicole Lurie, Karen L. Margolis, and Jonathon S. Slater. Accuracy of self-report of mammography and Pap smear in a low-income urban population. *American Journal of Preventive Medicine*, 14 (3):201–208, 1998.
- [333] Julia M. Shelley, Les M. Irwig, Judy M. Simpson, and Petra Macaskill. Who has Pap smears in New South Wales? Patterns of screening across sociodemographic groups. *Australian Journal of Public Health*, 18(4):406–411, 1994.
- [334] Australian Bureau of Statistics. *CLIB96 [Computer data file]*. Australian Bureau of Statistics, Canberra, 1997.
- [335] Australian Bureau of Statistics. *Socio-Economic Indexes for Areas*. Australian Bureau of Statistics, Canberra, 1991.
- [336] Australian Bureau of Statistics. *Information paper. 1996 Census of Population and Housing. Socio-Economic Indexes for Areas. (Catalogue 2039.0)*. Australian Bureau of Statistics, Canberra, 1998.

- [337] Richard A. Johnson and Dean W. Wichern. *Applied Multivariate Statistical Analysis*. Prentice-Hall Inc., Englewood Cliffs, New Jersey, second edition, 1988.
- [338] Julia Shelley, Les Irwig, Judy M. Simpson, and Petra Macaskill. Pap smear rates in New South Wales — 1984 to 1988. *Medical Journal of Australia*, 153(10):631–631, 1990.
- [339] Chris Del Mar, Paul Glasziou, Peter Adkins, Thuy Hua, and Mary Brown. Do personalised letters in Vietnamese increase cervical cancer screening among Vietnamese women? *Australian and New Zealand Journal of Public Health*, 22(7):824–825, 1998.
- [340] Karen L. Margolis, Nicole Lurie, Paul G. McGovern, Mary Tyrrell, and Jonathon S. Slater. Increasing breast and cervical cancer screening in low-income women. *Journal of General Internal Medicine*, 13(8):515–521, 1998.
- [341] Jennifer Dacey Allen, Anne M. Stoddard, Jeb Mays, and Glorian Sorensen. Promoting breast and cervical cancer screening at the workplace: Results from the woman to woman study. *American Journal of Public Health*, 91(4):584–590, 2001.
- [342] John F. C. Sung, Daniel S. Blumenthal, Ralph J. Coates, Jackie E. Williams, Ernest Alema-Mensah, and Jonathon M. Liff. Effect of a cancer screening intervention conducted by lay health workers among inner-city women. *American Journal of Preventive Medicine*, 13(1):51–57, 1997.
- [343] Electra D. Paskett, Cathy M. Tatum, Ralph D’Agostino Jr., Julia Rushing, Ramon Velez, Robert Michielutte, and Mark Dignan. Community-

- based interventions to improve breast and cervical cancer screening: Results of the Forsyth County Cancer Screening (FoCaS) Project. *Cancer Epidemiology, Biomarkers and Prevention*, 8(5):453–459, 1999.
- [344] Steven Whitman, Loretta Lacey, David Ansell, Jade Dell, Edwin Chen, and Clyde W. Phillips. An intervention to increase breast and cervical cancer screening in low-income African-American women. *Family and Community Health*, 17(1):56–63, 1994.
- [345] Jeanne Mandelblatt, Maryann Traxler, Patricia Lakin, Louys Thomas, Prem Chauhan, Stephen Matseoane, and Peter Kanetsky. A nurse practitioner intervention to increase breast and cervical cancer screening for poor, elderly black women. *Journal of General Internal Medicine*, 8(4):173–178, 1993.
- [346] Christopher N. H. Jenkins, Stephen J. McPhee, Joyce Adair Bird, Giao Qui Pham, Bang H. Nguyen, Thoa Nguyen, Ky Quoc Lai, Ching Wong, and Thomas B. Davis. Effect of a media-led education campaign on breast and cervical cancer screening among Vietnamese-American women. *Preventive Medicine*, 28(4):395–406, 1999.
- [347] Lucina Suarez, Rich Ann Roche, Lea Vonne Pulley, Nancy S. Weiss, Daniel Goldman, and Diane M. Simpson. Why a peer intervention program for Mexican-American women failed to modify the secular trend in cancer screening. *American Journal of Preventive Medicine*, 13(6):411–417, 1997.
- [348] Carolyn Cook Gotay, Richard O. Banner, Doris Segal Matsunaga, Nancy Hedlund, Rachelle Enos, Brian F. Issell, and Ho’oipo DeCambra. Impact

- of a culturally appropriate intervention on breast and cervical screening among native Hawaiian women. *Preventive Medicine*, 31(5):529–537, 2000.
- [349] J.M. Shelley, L.M. Irwig, J.M. Simpson, and P. Macaskill. A mass-media-led campaign to promote Pap smear screening. *Health Education Research*, 6(3):267–277, 1991.
- [350] Lucina Suarez, Donna C. Nichols, and Cody A. Brady. Use of peer role models to increase Pap smear and mammogram screening in Mexican-American and black women. *American Journal of Preventive Medicine*, 9(5):290–296, 1993.
- [351] Jilda C. G. Hyndman, Judith A.Y. Straton, Douglas A. Pritchard, and Helen Le Sueur. Cost-effectiveness of interventions to promote cervical screening in general practice. *Australian and New Zealand Journal of Public Health*, 20(3):272–277, 1996.
- [352] Sharon K. Buehler and Wanda L. Parsons. Effectiveness of a call/recall system in improving compliance with cervical screening: A randomized controlled trial. *Canadian Medical Association Journal*, 157(5):521–526, 1997.
- [353] Carol Somkin, Robert A. Hiatt, Leo B. Hurley, Elisabeth Gruskin, Lynn Ackerson, and Pamela Larson. The effect of patient and provider reminders on mammography and Papanicolaou smear screening in a large health maintenance organization. *Archives of Internal Medicine*, 157:1658–1664, 1997.
- [354] Ian McDowell, Claire Newell, and Walter Rosser. Computerized reminders

- to encourage cervical screening in family practice. *Journal of Family Practice*, 28(4):420–424, 1989.
- [355] Jeanette E. Ward and Elizabeth M. Proude. Evaluation of doctors' reminders in emergency departments to encourage cervical screening. *Australian and New Zealand Journal of Public Health*, 23(1):95–98, 1999.
- [356] Paula M. Lantz, Debra Stencil, MaryAnn T. Lippert, Sarah Beversdorf, Linda Jaros, and Patrick L. Remington. Breast and cervical screening in a low-income managed care sample: The efficacy of physician letters and phone calls. *American Journal of Public Health*, 85(6):834–836, 1995.
- [357] Antronette K. Yancey, Sora Park Tanjasiri, Mary Klein, and Janice Tunder. Increased cancer screening behavior in women of color by culturally sensitive video exposure. *Preventive Medicine*, 24(2):142–148, 1995.
- [358] Robert C. Burack, Phyllis A. Gimotty, Julie George, Scott McBride, Anita Moncrease, Michael S. Simon, Peter Dews, and Jennifer Coombs. How reminders given to patients and physicians affected Pap smear use in a health maintenance organization: Results of a randomized controlled trial. *Cancer*, 82:2391–2400, 1998.
- [359] Katherine Alexander and James McCullough. Application of marketing principles to improve participation in public health programs. *Journal of Community Health*, 6(3):216–222, 1981.
- [360] J.E. Byles, R.W. Sanson-Fisher, and S. Redman. Promoting screening for cervical cancer: Realizing the potential for recruitment by general practitioners. *Health Promotion International*, 11(4):299–308, 1996.

- [361] M. Dignan, R. Michielutte, H.B. Wells, and J. Bahnson. The Forsyth County cervical cancer prevention project – i. cervical cancer screening for black women. *Health Education Research: Theory and Practice*, 9(4): 411–420, 1994.
- [362] R. Michielutte, M. Dignan, J. Bahnson, and H.B. Wells. The Forsyth County cervical cancer prevention project – ii. compliance with screening follow-up of abnormal cervical smears. *Health Education Research: Theory and Practice*, 9(4):421–432, 1994.
- [363] Daniel S. Tseng, Elizabeth Cox, Mary Beth Plane, and Khin Mae Hla. Efficacy of patient letter reminders on cervical cancer screening: A meta-analysis. *Journal of General Internal Medicine*, 16(8):563–568, 2001.
- [364] Heather Mitchell, Sheila Hirst, Jill Cockburn, Dorothy J. Reading, Margaret Staples, and Gabriele Medley. Cervical cancer screening: A comparison of recruitment strategies among older women. *Medical Journal of Australia*, 155(2):79–82, 1991.

# Appendices

# Appendix A

## Appendix for Chapter 3

### A.1 GEIS compiler steps

The principal actions taken by the compiler are described below in the order in which they were carried out by GEIS. Not all steps necessarily applied to every script. If errors were found then the compiler exited.

The compiler steps were:

1. The current data sets were saved, if they existed;
2. The script text file was checked for non-text characters;
3. A check was made to ensure that required variables were present in the CONFID data set;
4. The FORMATS.SAS files were processed to create the GEIS and project formats\*;
5. The CALLS, CONTROL, and PROTECT data sets were created;
6. Any DO loops in the script were expanded;
7. The ANSWERS, PROTECT and COMBRESP data sets were created;
8. A series of checks were made using the SCRIPT data set:
  - Required variables had to be present;
  - Item names had to be valid and not duplicated;
  - Item options had to have acceptable values for each item type;
  - Any external data set referenced by the script had to exist. It also had to contain an ID variable and other referenced variables;
  - The location nominated to hold IVR voice recordings existed;
  - SAS formats and informats<sup>†</sup> had to have valid names;
9. Any automatic item formats were created;

---

\*Formats are used in SAS software to define the output appearance of data values.

<sup>†</sup>Informats are used in SAS software to define how data values should be read in.



10. All variables in the ANSWERS, PROTECT and COMBRESP data sets were created;
11. The AUTOCALL data set was created;
12. The SCHSTR variable in the CONTROL data set was created;
13. All cases with valid ID values were merged from the CONFID data set to the CONTROL, ANSWERS and PROTECT data sets;
14. Required variables were copied from CONFID to CONTROL;
15. Distances were inserted into the CONTROL data set;
16. Variables in the INTRVS data set were initialized;
17. For all LINK items, data from variables in external data sets were imported into the ANSWERS data set;
18. For each MAIL item, e-mail address information was imported to the ANSWERS data set;
19. SPSs were analyzed for syntax errors;
20. All script item information were stored in the DEFINIT catalogue;
21. For IVR scripts:
  - The PHRASES data set was created if it did not exist;
  - The PHRASES data set was updated with STANDARD data set phrases;
  - Item key words were added to the PHRASES data set;
22. For the script, the compiler checked:
  - The first item was of type TITL and only one TITL item was defined;
  - The item names were valid, not missing, and unique;
  - All secondary variable names were valid, not missing, and unique;
  - All STAT items did not precede items that modified survey data;
  - It was possible to reach each item from an earlier item;
  - All SPSs only referred to previous items and not later ones, or non-existent ones;
  - There were no self-references in the SPSs;
  - For LVLC items, all referenced items had to exist, and were of types CHCE or MULT, and each CHCE or MULT had the nominated levels listed in the LVLC item;
  - For all CHCE and MULT items the option codes were unique, not missing, and were numeric;
  - For MULT items, the maximum number of choices did not exceed the number of options;

- Data sets and variables referred to by LINK items existed and imported data was within specified absolute limits;
  - Data sets and variables referred to by external update links existed and the syntax of the ID-matching specification was valid;
  - CBDT items did not follow CBTM items;
  - The default ring-back time was set for CBTM items;
  - For the TABL type, the overall absolute upper limit was not less than the overall absolute lower limit;
  - For the TABL type, the maximum number of rows that could be selected did not exceed the actual number of rows;
  - For absolute limits set on individual rows of the TABL type, the upper limit was not less than the lower limit;
  - For the NUM type, the absolute upper limit was not less than the absolute lower limit;
  - For the NUM type, the upper value of the reasonableness limit was not less than the lower value;
  - For the NUM type, the reasonableness limits were contained within the absolute limits;
  - The last item in the script was not of type: TIME, LINK, NULL, CALC, or STAT;
  - STAT items were defined for at least the 'CQ' status code;
  - The GEIS formats were defined;
23. For IVR, the compiler checked:
- The maximum repetition count was defined for relevant items;
  - The settings for CALL and HUP items were valid;
  - The voice recording location could be found;
  - The key code values for CHCE and MULT items were unique and not missing;
  - The IVR keywords for each CHCE items were unique, not missing, and defined;
  - If an INFO item appeared last, the WAIT or PAUSE options were not set;
  - If an INFO item had the RESP option set, the response and non-response codes were also set;
  - If an OPEN item was set, the recording location existed;
  - The CALL, HUP, CBTM, and CBDT items were used;
24. The local data sets were created: LOCALANS, LOCALCTL, LOCALCNF, and LOCALPRT; and
25. Existing data were merged, if requested, by matching on ID.

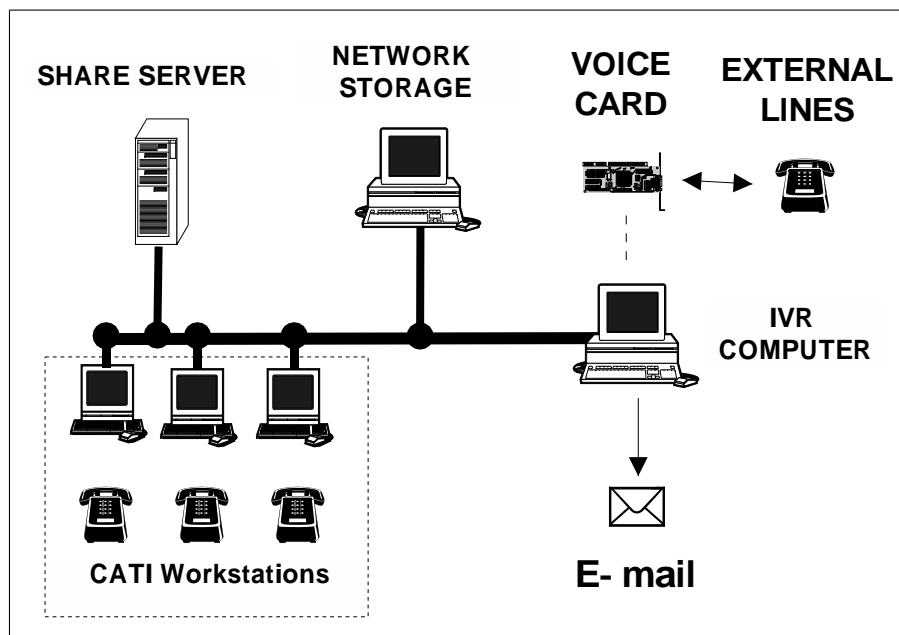


Figure A.1: Hardware set up.

## A.2 IVR equipment

To connect to the Public Telephone System a voice card was installed within the computer as illustrated in Figure A.1. A Dialogic® D/41H voice card was used allowing 4 lines to be connected. Visual Voice software allowed access to the voice card drivers.

## A.3 Interview STATUS variable codes

A number of codes were used by the STATUS variable to monitor the state of interviews. A full list is given in Table A.1 although not all codes had to be used in any particular project. The STATUS variable codes were defined by the FORMATS.SAS file shown in Supplementary Materials Section 3.2.

## A.4 The GEIS call-scheduling algorithm

The GEIS call-scheduling algorithm (Figure 3.13 of Chapter 3) was embedded in a loop that contained a series of checks. Each check consisted of applying a SAS WHERE statement to the CONTROL data set. To be selected, a case had to have the:

- SELECTED variable set to one;
- ELIGIBLE variable set to zero; and
- START variable set to a time and date before the current time and date.

Table A.1: A complete list of valid codes for the STATUS variable.

<i>Status</i>	<i>Meaning</i>
'AB'	Answering machine — business
'AH'	Answering machine — household
'AI'	Answering machine — household — no message left
'AM'	Answering machine — general
'AN'	Answering machine — household — message left
'AO'	Answering machine — unknown if household
'BG'	Business or Government
'BI'	Business or Government — Institution
'BQ'	Business or Government — Group quarters
'CB'	Callback arranged
'CE'	Callback — respondent eligible
'CQ'	Completed questionnaire
'CU'	Callback — eligibility unknown
'D3'	Refused by third party or household
'DF'	Dropped — friendly
'DO'	Dropped part-way
'DR'	Dropped before starting
'DT'	Disconnected tone
'DZ'	Dropped — other reason
'ER'	Error condition
'ET'	Engaged tone
'FM'	Fax machine
'IV'	Interview in progress
'LR'	Letter returned — address unknown
'NA'	No attempt made to contact
'OS'	Out of scope — ineligible
'OT'	Other reason for no contact
'PD'	Phone difficulties
'PQ'	Partly-completed
'RB'	Respondent has call blocking
'RD'	Respondent is dead
'RE'	No eligible respondent
'RI'	Respondent is incompetent
'RL'	Respondent not English-speaker
'RM'	Respondent has moved
'RS'	Respondent is too sick to help
'RT'	Ring tone
'RV'	Respondent on vacation
'TO'	Time out
'UA'	Unable to complete interview (special)
'UL'	Respondent unlocatable
'UN'	No ring — strange noise
'US'	No ring — silence
'WF'	Call forwarding
'WH'	Number changed — not a household
'WN'	Number changed
'WP'	Cell/mobile phone

Each check in the loop then applied an additional test in addition to these conditions. The first check in the loop tested for cases where the status was set to 'CB'. If there were no cases that matched this WHERE statement, the second check was applied and tested for cases that had status 'AM' and which were last rung some time ago. If there were still no cases that matched the WHERE statement, the third check tested for non-contacted cases that were last rung some time ago. Finally, GEIS checked for any unattempted cases, which were those with the STATUS variable set to 'NA'.

## A.5 GEIS data sets and catalogues

There were a number of data sets and one catalogue used by GEIS. The DEFINIT catalogue (Section A.5.1) was used to store all project settings as well as the compiled version of the script. The DISTANCE data set (Section A.5.2) was used to save distances from Newcastle to telephone exchanges in Australia. The SCRIPT data set (Section A.5.3) was used to store the script after it was imported.

There were five main data sets used to control interviews or collect interview data: CONFID (Section A.5.4), CONTROL (Section A.5.5), ANSWERS (Section A.5.6), PROTECT (Section A.5.7), and CALLS (Section A.5.8). The COMBRESP data set (Section A.5.9) was used to collect data while an interview was in progress.

There were also copies of four of the main data sets used to store data locally on workstations during interviews: LOCALCNF, LOCALCTL, LOCALANS, and LOCALPRT (Section A.5.10). These were used in conjunction with the COMBRESP dataset to allow automatic backups during an interview to guard against power failures.

Most system events were stored automatically in the interviewing log: ILOG (Section A.5.11).

For CATI scripts, interviewer log-in times were saved to the INTRVS data set (Section A.5.12). For IVR scripts, voice recording information was stored in the STANDARD and PHRASES data sets (Section A.5.13).

For the Hybrid method, the AUTOCALL data set (Section A.5.14) was used to coordinate calls between the interviewers and the IVR system. Although the Hybrid method is not discussed until Section 4.6 of Chapter 4, the AUTOCALL data set is described in this appendix for completeness.

The catalogues and data sets are described in greater detail below.

### A.5.1 The DEFINIT Catalogue

The DEFINIT catalogue stored project configuration as well as the compiled script. Its structure is shown in Table A.2.

### A.5.2 The DISTANCE data set

The DISTANCE data set contained the distances from Wallsend exchange to all the telephone exchanges in Australia. It was created by the candidate using Equation 3.1 of Chapter 4 and a list of the longitude and latitude of telephone

Table A.2: The structure of the DEFINIT catalogue<sup>a</sup>.

<i>Item</i>	<i>Description</i>
BACKUP	Details of last backup
CALLSHED	Stores call scheduling information
CONFSET	Compiler options settings
DIAG	Automatic flow chart storage location
EMAIL	E-mail settings
IVR	Control list used to control running IVR systems
SCRIPT	The compiled script
SCRPINIT	Initialization settings for the Script window
TREEINFO	Summary information on the last compilation

<sup>a</sup> This is an example only.

Table A.3: The contents of the DISTANCE data set.

<i>Telephone Prefix</i>	<i>Telephone Exchange</i>	<i>Latitude</i>	<i>Longitude</i>	<i>Distance</i>
02422	Wollongong	-34.425000	150.841667	187.128
02423	Kiama	-34.638889	150.781944	211.012
024250	Wollongong	-34.425000	150.841667	187.128
024251	Wollongong	-34.425000	150.841667	187.128
024252	Wollongong	-34.425000	150.841667	187.128
024253	Wollongong	-34.425000	150.841667	187.128
024254	Wollongong	-34.425000	150.841667	187.128
0242550	Helensburgh	-34.200000	150.983333	158.971
0242551	Wollongong	-34.425000	150.841667	187.128
<sup>a</sup>				

<sup>a</sup> The rest of the file is omitted.

exchanges in Australia obtained from Telstra<sup>189</sup>. Only a fragment of the contents of the DISTANCE data set is shown in Table A.3 since the actual file is very lengthy.

### A.5.3 The **SCRIPT** data set

The **SCRIPT** data set was created when importing a script. Its structure is shown in Table A.4. The data set contained the information used by the compiler. Each observation in the data set corresponded to an item in the script. The **SURVTYPE** variable in the **SCRIPT** data set determined the interview mode. Valid codes for the **SURVTYPE** variable are shown in Table A.5.

Table A.4: Structure of the **SCRIPT** data set.

<i>Variable<sup>a</sup></i>	<i>Type<sup>b</sup></i>	<i>Label</i>
QTYPE	Char	Item type
SURVTYPE	Char	Survey type
VOXLINES	Char	IVR keyword
QNAME	Char	Item name
VNAME	Char	Variable name
SP	Num	Number of SPSs
QNLEVELS	Num	Number of CHCE or MULT levels
MAXMULT	Num	Maximum number of MULT level selections
NLINES	Num	Number of lines to display
NVARS	Num	Number of extra variables used by some item types
NCALC	Num	Number of calculation statements used by CALC type
PAUSE	Num	Delay in milliseconds for an IVR item to pause
NLOOK	Num	Number of previous and following items for evaluating SPSs
REASMIN	Num	Reasonable minimum for numeric data entry (NUM type)
REASMAX	Num	Reasonable maximum for numeric data entry (NUM type)
ABSMIN	Num	Absolute minimum for numeric data entry (NUM type)
ABSMAX	Num	Absolute maximum for numeric data entry (NUM type)
DEFCBDT	Num	Default number of days when arranging a call-back
STAT	Num	Level counted by LVLC item
QINFORM	Char	Item informat
QFORMAT	Char	Item format
JUMP	Char	Flags if CHCE item jumps ahead
MODULE	Char	Module name
SUBMODUL	Char	Submodule name
LABEL	Char	Item label
ARESP	Char	Add respondent flag
AUTO	Char	Automatic dialling flag

Table A.4 continued on next page.

Table A.4 continued from previous page.

<i>Variable</i> <sup>a</sup>	<i>Type</i> <sup>b</sup>	<i>Label</i>
NONAME	Char	Respondent name shown or hidden on screen
ENABLE	Char	Disable/Enable setting status by STAT item
RECLOC	Char	Location to store voice recordings
PATH	Char	Path to sub-script; path to AUTOCALL library
VARSTAT	Char	Statistic calculated by TABL item
STATUS	Char	Status code
AUXST	Char	Auxiliary status code
NOWAIT	Char	Controls behaviour of items in IVR surveys
CC	Num	MULT/CHCE input/output format
SELPRT1–SELPRT5	Char	Self protection statements
LINE1–LINE50	Char	Item question lines
QCODE1–QCODE50	Num	CHCE/MULT option codes
QCHCE1–QCHCE50	Char	CHCE/MULT option texts
VCHCE1–VCHCE50	Char	CHCE IVR level keywords
MLTLB1–MLTLB50	Char	MULT level label
CALC1–CALC50	Char	CALC statement
LVAR1–LVAR50	Char	Names of extra variables in CALC items
NLVAR1–NLVAR50	Char	Linked variable new name in LINK items
LWARD1–LWARD50	Char	Dynamic/Static flag in LINK items
VART1–VART50	Char	Char/Num type in LINK items
DSNAM1–DSNAM50	Char	Name of external data set in LINK items
QKEY1–QKEY50	Num	IVR phone keypad numbers to be used
TABL1–TABL100	Char	Label for TABL row
RMIN1–RMIN100	Num	Minimum numeric value for TABL row
RMAX1–RMAX100	Num	Maximum numeric value for TABL row
RVAL1–RVAL100	Num	Numeric value for TABL row button
BTN1–BTN100	Char	Flag a TABL row to be a button or entry field
EUPDATE	Char	External update flag
KK	Num	IVR key code column
NCHAR	Num	Number of characters in OPEN type
NROWS	Num	Number of rows in TABL type

<sup>a</sup> Variables are listed in the actual physical order in which they appear in the data set.<sup>b</sup> Num=numeric variable. Char=character variable.



Table A.5: Codes used by the SURVTYPE variable in the CONFID data set.

<i>Code</i>	<i>Meaning</i>
CATI	Computer Assisted Telephone Interview Survey
IVR	Outbound Interactive Voice Response Survey
IVR.IN	Inbound Interactive Voice Response Survey
SIVR.P	Subordinate Voice Response Survey — Passive
CASI	Computer Assisted Self Interview

Table A.6: Structure of the CONFID data set.

<i>Variable<sup>a</sup></i>	<i>Type<sup>b</sup></i>	<i>Label</i>
ADDRESS	Char	Respondent's full address
ELIGIBLE	Num	Eligibility Flag
ID	Num	Respondent ID
FULLNAME	Char	Respondent's full name
POSTCODE	Char	Respondent postcode
SELECTED	Num	Selection Flag
STDPHONE	Char	Respondent telephone number

<sup>a</sup> These are the minimum variables. Additional variables could be present.

<sup>b</sup> Num=numeric variable. Char=character variable.

#### A.5.4 The CONFID data set

The CONFID data set was created by the user and contains the respondents' ID codes, names, telephone numbers, and so on. Its structure is shown in Table A.6. Respondent information was kept in CONFID to ensure that respondents' personal information were kept separate from their answers.

The respondents' ID codes were copied from the CONFID data set by the compiler to the CONTROL, ANSWERS, and PROTECT data sets. The CONFID data set was also commonly used as a location for external variables that were to be linked into the ANSWERS data set by means of a LINK item.

#### A.5.5 The CONTROL data set

The CONTROL data set was created by the compiler and used by GEIS to store information relating to the state of interviews. Its structure is shown in Table A.7.

The STATUS variable was called the monotonic call status. It contained a two-letter code that indicated the current state of each case. Codes used in the STATUS variable are shown in Table A.3 of Section A.3. A history of the consecutive call attempts was stored in STAT1–STAT5. Once a respondent had been contacted the STATUS variable could not be reset to a non-contacted code value as might occur if they were unavailable on a later contact. Instead, the STATUS variable remained unchanged. However, the actual outcome of each

Table A.7: Structure of the CONTROL data set.

<i>Variable</i> <sup>a</sup>	<i>Type</i> <sup>b</sup>	<i>Label</i>
ARESLT	Num	IVR Action result
ATTDATE	Num	Last attempt date
ATTTIME	Num	Last attempt time
AUXST	Char	Auxiliary call status
CALLCOST	Num	Cost of call (\$AU)
CALLLGTH	Num	Greeting length
CALLRSLT	Num	Call result
CBCKDATE	Num	Call back date
CBCKTIME	Num	Call back time
CMPDATE	Num	Completion date
CMPTIME	Num	Completion time
CNTDATE	Num	Contact date
CNTTIME	Num	Contact time
COMMENT	Char	Comment
DISTANCE	Num	Distance in Kilometres
DURAT1–DURAT5	Num	Duration of previous interviews
DURATION	Num	Accumulated interview duration
ELIGIBLE	Num	Eligibility flag
ENTRY	Num	Respondent IVR entry code
ID	Num	Respondent ID
INTVR	Char	Interviewer
INTVR1–INTVR5	Char	Previous interviewers
ISTATUS	Char	Status set by the interviewer
LASTQ	Char	Last question answered
LASTQ1–LASTQ5	Char	previous last questions
NC_DATE1–NC_TIME5	Num	Last noncontact dates
NC_TIME1–NC_TIME5	Num	Last noncontact times
NTIMECMP	Num	Number of calls until completion
NTIMECNT	Num	Number of calls to make contact
NTIMELST	Num	Number of calls since noncontact
NTIMES	Num	Number of call attempts
SCHSTR	Char	Search string
SELECTED	Num	Selection flag
START	Num	Earliest interview date/time
STATUS	Char	Monotonic call status
STATUS0	Char	Actual calls status
STAT1–STAT5	Char	Attempt status
STDPHONE	Char	Telephone number
STRATUM	Num	Stratum/Clustering ID
SURVTYPE	Char	Survey type

<sup>a</sup> Variables are listed in the actual physical order in which they appear in the data set.

<sup>b</sup> Num=numeric variable. Char=character variable.

Table A.8: Call codes used by the CALLSLT variable in the CONTROL data set.

<i>Code value</i>	<i>Acoustic Detection Outcome</i>
1	break in cadence
2	loop current drop
3	voice detection
4	answering machine

call attempt was stored in the STATUS0 variable.

The duration of successive call attempts was accumulated in the DURATION variable. A history of the duration of consecutive call attempts was stored in DURAT1–DURAT5.

The last question displayed by GEIS when an interview ended was stored in the LASTQ variable. A call history was stored in LASTQ1–LASTQ5.

The interviewer for each call attempt was stored in the INTVR variable. If the IVR system was conducting the interview then INTVR was set to the user's name with the voice card line number appended; e.g. 'Ross Corkrey3', 'Ross Corkrey4'.

The CBCKDATE and CBCKTIME variables stored the date and time to call back a respondent in SAS date and time values, respectively.

The total number of call attempts was stored in the NTIMES variable, the number to make contact was stored in NTIMECNT, the number of calls since the last non-contact in NTIMELST, and the number to finalize an interview in NTIMECMP.

The date and time of each call attempt were stored in ATTTIME and ATTDATE, respectively. A call history was maintained of non-contacted cases in NC\_DATE1–NC\_DATE5 and NC\_TIME1–NC\_TIME5. The date and time of contact were stored in CNTTIME and CNTDATE, while the date and time each case was finalized were stored in CMPTIME and CMPDATE.

The approximate accumulated cost of calls was CALLCOST. The call cost calculation used the distance from Newcastle, Australia to the respondent's telephone exchange, which the compiler stored in the DISTANCE variable (Section A.5.2). It was meant to be used as a guide only.

When the IVR system rang a respondent's number it attempted to identify whether the call was answered by a person or an answering machine by one of two methods: salutation length and the acoustic properties of the respondent's salutation. The outcome of the acoustic determination was stored in the variable CALLSLT and the length of the salutation in CALLLGTH. The outcome of other IVR actions was stored in the ARESULT variable. Code values used by the ARESULT and CALLSLT variables are listed in Tables A.9 and A.8, respectively.

The STDPHONE variable contained the respondent's telephone number.

The SELECTED variable might have two possible values: zero (0) or one (1). When set to one, a respondent could be interviewed; when set to zero, a respondent could not to be interviewed. Respondents who were de-selected would not appear in the pull-down list for selection by interviewers and were also unavailable for selection by the IVR system. This variable was principally used to de-select or select large groups of respondents.

Table A.9: Call codes used by the ARESULT variable in the CONTROL data set.

<i>Code</i>	<i>IVR Action Outcome</i>
-1	RuntimeError
0	OK
1	Stopped
2	TermDigit
3	MaxSilence
4	MaxNonSilence
5	MaxDigits
6	TimeOut
7	TimeOutInterDigit
8	Connected
9	Busy
10	NoRing
11	NoAnswer
12	OperatorIntercept
13	CallError
14	Fax
15	NoDialTone

The ELIGIBLE variable could contain any numeric value, but only a zero (0) value indicated a respondent could be interviewed; any non-zero value indicated that a respondent was not to be interviewed. The variable was used to remove individual respondents from selection during a survey. Suggested code values for the ELIGIBLE variable are shown in Supplementary Materials Section 3.2.

The START variable contained the earliest date and time a respondent could be interviewed.

The SCHSTR variable was used to hold a character string that was displayed to the interviewer when selecting respondents to interview. It was automatically generated by the compiler from either the respondent's ID code and name, or just the ID code, depending on whether names were to be suppressed or not.

The AUXST variable held any user-defined status codes.

#### A.5.6 The ANSWERS data set

The ANSWERS data set was created by the compiler and was used to store all response data. Its structure is shown in Table A.10. It was initially created with an ID variable and a series of other blank response variables. The formats, informats and labels of the ANSWERS data set response variables were set by the script. Each response variable corresponded to one of the items in the script, but there could be more than one variable per item. Some variables could be preset using LINK items. These items imported data from external data sets before the interview.

Table A.10: Structure of the ANSWERS data set<sup>a</sup>.

<i>Variable<sup>b</sup></i>	<i>Type<sup>c</sup></i>	<i>Label</i>
ID	Num	Respondent ID
T0	Num	Duration: start interview
IQ02	Num	Willing to answer questions?
CB2	Num	Convenient time or callback?
CBACK	Num	Ask if can callback
IQ03	Num	Persuade initial refusers
T6	Num	Duration: start technology
T_02	Num	Mobile phone
T_06	Num	Answering machine
T_07	Num	Experienced IVR before
H_04	Num	How often use HEROIN
T7	Num	Time end Heroin domain

<sup>a</sup> This is an example. The actual structure depends on the script.

<sup>b</sup> Variables are listed in the actual physical order in which they appear in the data set.

<sup>c</sup> Num=numeric variable. Char=character variable.

### A.5.7 The PROTECT data set

The PROTECT data set was created by the compiler and was used to store the outcomes of all Self-Protection Statements (SPSs) when an interview was stopped. The PROTECT data set was used to reconstruct the state of an interview when it was resumed. This was done by storing the values of all SPSs. Its structure is shown in Table A.11.

It had the same number of variables as the ANSWERS data set. Apart from the ID variable, all the variables stored a ‘Y’ to indicate an item was protected (could not be displayed) or ‘N’ to indicate it was unprotected (could be displayed).

### A.5.8 The CALLS data set

All call attempts made by interviewers and the IVR system were logged, in chronological sequence, to the CALLS data set. Its structure is shown in Table A.12. The date and time a call attempt ended, the status code before and after the call attempt, and the interviewer’s name were recorded.

### A.5.9 The COMBRESP data set

This was the data set that was used during an interview to store data before it was copied to the local data sets or the main data sets. Its structure is shown in Table A.13.

Table A.11: Structure of the PROTECT data set<sup>a</sup>.

<i>Variable</i> <sup>b</sup>	<i>Type</i> <sup>c</sup>	<i>Label</i>
ID	Num	Respondent ID
INTRO	Char	
IQ02	Char	
CB1	Char	
RESTART	Char	
CB2	Char	
CBACK	Char	
ENDCB	Char	
NEXTTERM	Char	
N3	Char	
THANKS	Char	
THANK_CB	Char	
TERM	Char	

<sup>a</sup> This is an example. The actual structure depends on the script.

<sup>b</sup> Variables are listed in the actual physical order in which they appear in the data set.

<sup>c</sup> Num=numeric variable. Char=character variable.

Table A.12: Structure of the CALLS data set.

<i>Variable</i> <sup>a</sup>	<i>Type</i> <sup>b</sup>	<i>Label</i>
DATETIME	Num	Date and time
ID	Num	ID
STATUS	Char	Call Status
LASTSTTS	Char	Last call Status
INTVR	Char	Interviewer

<sup>a</sup> Variables are listed in the actual physical order in which they appear in the data set.

<sup>b</sup> Num=numeric variable. Char=character variable.

Table A.13: Structure of the COMBRESP data set<sup>a</sup>.

<i>Variable<sup>b</sup></i>	<i>Type<sup>c</sup></i>	<i>Label</i>
ID	Num	Respondent ID
_INTVR_	Char	Interviewer
_TITLE_	Char	Title
_STATUS_	Char	STATUS
_PHONE_	Char	Phone
_RSTAT_	Char	Remote status
_RNAME_	Char	Respondent name
_DATE_	Num	Attempt date
_TIME_	Num	Attempt time
_WAIT_	Num	Wait time
_NRPT_	Num	Number of times to repeat question
T_START	Num	Start time
Q1	Char	Persons name
Q2	Char	Persons position in organization
Q3	Char	Name of the organization
⋮	⋮	⋮
T_END	Num	Finish time
STAT_CQ	Char	
END	Num	

<sup>a</sup> This is an example. The actual structure depends on the script.

<sup>b</sup> Variables are listed in the actual physical order in which they appear in the data set.

<sup>c</sup> Num=numeric variable. Char=character variable.

Table A.14: Structure of the ILOG data set.

<i>Variable</i> <sup>a</sup>	<i>Type</i> <sup>b</sup>	<i>Label</i>
WHEN	Num	When
WHERE	Char	Where
NOTE	Char	Note
MSG	Char	Message

<sup>a</sup> Variables are listed in the actual physical order in which they appear in the data set.

<sup>b</sup> Num=numeric variable. Char=character variable.

#### A.5.10 The local data sets

The main data sets ANSWERS, CONTROL, PROTECT, and CONFID stored interview data centrally. There were four local data sets, LOCALANS, LOCALCTL, LOCALCNF, and LOCALPRT, each with the same structures as the corresponding main data sets ANSWERS, CONTROL, CONFID, and PROTECT. These local data sets stored information locally on CATI interviewer's workstations or the IVR machine. When an interview was to be conducted the respondent's data was copied from the main data sets to the local data sets. Data were periodically saved to the local data sets during interviews. This provided a safeguard against interruption by power failure or other cause. After an interview was stopped the contents of the local data sets were copied back to the main data sets.

#### A.5.11 The ILOG data set

Events along with the date and time they occurred were logged to the ILOG data set. Its structure is shown in Table A.14. Examples of the events recorded in the log are: script compile results, interviewer and user log-ins, interview attempt details, interview outcome details, pressing the '**BACK**' key in an interview, data back-ups, and manual user access to data sets.

#### A.5.12 The INTRVS data set

The INTRVS data set had to be created by the user before compiling. Its structure is shown in Table A.15. It was used to store the accumulated interviewing duration and log-in times for each interviewer, and interviewer passwords.

#### A.5.13 The STANDARD and PHRASES data sets

The STANDARD and PHRASES data sets were used with IVR surveys. The STANDARD data set was part of GEIS rather than a project data set. It was used to define standard spoken phrases used by GEIS. Its structure is shown in Table A.16.



Table A.15: Structure of the INTRVS data set.

<i>Variable</i> <sup>a</sup>	<i>Type</i> <sup>b</sup>	<i>Label</i>
ID	Num	
USERNAME	Char	Username
SURNAME	Char	Surname
FIRST_NAME	Char	First name
PASSWORD	Char	Password
DURATION	Num	Interviewer log time
NON_TALK	Num	Interviewer non-talking time

<sup>a</sup> Variables are listed in the actual physical order in which they appear in the data set.

<sup>b</sup> Num=numeric variable. Char=character variable.

Examples of standard phrases are:

“To confirm that you want to complete the interview at a later date, press 9 again, or press any key to continue.”

and

“The survey will now resume at the point left off.”

These phrases were essential and had to be recorded for each project, preferably in the voice used to record the project’s interview questions.

The PHRASES data set was created by the compiler. The compiler inserted the contents of the STANDARD data set and the keywords found in the GEIS script. At a later stage the pathnames and filenames of the voice recordings were inserted into the PHRASES data set. The PHRASES data set was used during interviews to select which message recordings to play. Its structure is shown in Table A.17.

#### A.5.14 The AUTOCALL data set

The AUTOCALL data set was created by the compiler and was used to coordinate call transfers to and from the IVR system. Its structure is shown in Table A.18. It was principally used in the Hybrid method and its use is described in Section 4.6.3 of Chapter 4.

## A.6 The share server

The share server used SAS/SHARE software<sup>223</sup>. The program shown in Figure A.2 was written to run the share server. The program shown in Figure A.3 illustrates how data was accessed using the share server.

Table A.16: Structure of the STANDARD data set.

<i>Variable</i> <sup>a</sup>	<i>Type</i> <sup>b</sup>	<i>Label</i>
LINE1	Char	LINE1
LINE2	Char	LINE2
LINE3	Char	LINE3
LINE4	Char	LINE4
LINE5	Char	LINE5
LINE6	Char	LINE6
LINE7	Char	LINE7
LINE8	Char	LINE8
PHRASE	Char	PHRASE

<sup>a</sup> Variables are listed in the actual physical order in which they appear in the data set.

<sup>b</sup> Num=numeric variable. Char=character variable.

Table A.17: Structure of the PHRASES data set.

<i>Variable</i> <sup>a</sup>	<i>Type</i> <sup>b</sup>	<i>Label</i>
PHRASE	Char	Phrase keyword
IDX	Num	Index
FNAME	Char	Filename
USED	Num	Usage
ACTION	Num	Action
LINE1	Char	LINE1
LINE2	Char	LINE2
LINE3	Char	LINE3
LINE4	Char	LINE4
LINE5	Char	LINE5
LINE6	Char	LINE6
LINE7	Char	LINE7
LINE8	Char	LINE8
LINE9	Char	
LINE10	Char	
LINE11	Char	
LINE12	Char	

<sup>a</sup> Variables are listed in the actual physical order in which they appear in the data set.

<sup>b</sup> Num=numeric variable. Char=character variable.

Table A.18: Structure of the AUTOCALL data set.

<i>Variable<sup>a</sup></i>	<i>Type<sup>b</sup></i>	<i>Label</i>
INTERVW	Char	Interviewer
CALLNO	Char	Calling number
EXTNO	Char	Extension number
ACTION	Char	Action
IDX1	Num	Call result
IDX2	Num	Call result
STATUS	Char	Remote interview status

<sup>a</sup> Variables are listed in the actual physical order in which they appear in the data set.

<sup>b</sup> Num=numeric variable. Char=character variable.

```
* SAS 6 server;
* location of GEIS application formats;
libname core 'd:\GEIS\bin' ;
* communications mode;
options comamid=spx;
* search order for formats;
options FMTSEARCH=(core library work);
* Generate formats;
%include 'd:\GEIS\bin\formats.sas';
* runs the server;
proc server serverid=shrsvr ;
run;
```

Figure A.2: Program used to run the share server.

```
%let bin=bin1_14;
%let projname=project_name;
options comamid=spx;
%include "d:\GEIS\bin\info.sas"/source2;
libname common          "n:\data\GEIS\common" ;
libname cati             "d:\GEIS\&bin" ;
libname core             "d:\GEIS\&bin" ;
libname datacati remote "d:\GEIS\projects\&projname" server=shrsvr;
libname library remote "d:\GEIS\projects\&projname" server=shrsvr;
options fmtsearch=(core library work);
proc print data=datacati.answers;
run;
```

Figure A.3: Example of a program used to access data on the share server.

## A.7 GEIS script syntax

A GEIS script consisted of a series of discrete items that corresponded to interview questions or performed other functions, and which were related by a logic structure specified by the use of Self-Protection Statements (SPSs). A list of the script item types is given in Section A.7.4 and SPSs are described in Section A.7.3. All items could have texts (Section A.7.1) and the texts could use answer-quoting (Section A.7.2).

In the IVR mode, each question had to have a recording of the question specified within the item using a keyword. The keyword was inserted by the compiler into the PHRASES data set (Section A.5.13). The keyword could also use answer-quoting (Section A.7.2).

### A.7.1 Item texts

As far as possible the screen design for CATI remained constant regardless of the script contents. For maximum contrast the question text was displayed in a yellow-coloured fixed-font eight points in height against a dark blue background. Each line contained 75 characters and 8 lines could be displayed without scrolling.

In the IVR mode, a voice recording of the question text was played whenever the text would normally be read out by an interviewer in the CATI mode. The voice recording for the question text was specified by means of a keyword. The PHRASES data set was searched for the keyword in order to identify which message recording file was to be played.

### A.7.2 Answer-quoting in IVR

In the IVR mode, recorded messages could use answer-quoting within keywords. This was done by quoting another keyword already defined in another item, or by using a system quote.

In Figure A.4 is an example of a CHCE item with eight lines of question text. It has the keyword “HELLO\_C|F, ^NRPT\_~, TIMES|F”. The keyword consists of a concatenation of three parts separated by commas:

**HELLO\_C:** a keyword defined in the PHRASES data set for a voice recording of the phrase “This message will repeat”

**NRPT:** a system quote that equalled the number of times the current item had been played;

**TIMES:** another keyword defined in the PHRASES data set for the voice recording “times”.

The suffixes of the first and third parts, “|F”, indicate that those parts are not system quotes, but refer to voice recordings stored in files. The effect of this triple keyword is to play the voice recording for the keyword ‘HELLO\_C’, followed by the repetition count, and then the voice recording labelled ‘TIMES’.

The listener hears a message such as:

“This message will repeat 3 times.”

CHCE	1	11	HELLO_C	8		_MAKE_		LABEL
CALL				17	25	1	KEY9	
(HELLO>.)&(RNDINTRO=3)								
Hello, this is a free recorded voice information service from the Hunter Centre for Health Advancement. We are making this call to try to increase community awareness in the Cessnock area about Pap tests. We are pleased to offer you the chance to listen to an important message about Pap tests in the privacy of your own home.								
If you'd like to listen to this 5 minute message, please press key 5 on your phone now. This message will repeat ^_NRPT_^ times.								
HELLO_C F,^_NRPT_^,TIMES F								
1			1				key 1	
2			2				key 2	
3			3				key 3	
4			4				key 4	
5			5				key 5	
6			6				key 6	
7			7				key 7	
8			8				key 8	
9			9				key 9	
0			0				key 0	
1006			-1				Timeout	
3		10						
Intro C								
*****								

Figure A.4: Example of a question text and answer-quoting for an IVR interview.

### A.7.3 Self-protection statements

Self-Protection Statements (SPSs) stored the logic structure of a questionnaire. They were used to control which question was displayed at each step in the interview. A SPS consisted of one to five text lines, each of 200 characters in length. These were evaluated as if they formed a single statement 1000 characters long. The SPS in a script had to contain references to other earlier items in the survey.

As values were entered during an interview the SPSs of following items were evaluated in order to determine which should be displayed next. Not all the SPSs of subsequent items were examined since there could be very many possibilities. Only the SPS of those items that the compiler previously determined were logically related to the current item were examined. If an item was protected then all the other items that also depended on the protected item became protected as well. This thereby locked off an entire pathway through the script.

It was possible for a CATI interviewer or IVR respondent to return to a previously answered question and alter the answer. When they moved ahead again, depending on the logic structure specified in the script, it was possible that a completely different series of questions might be displayed by GEIS and the previously-entered answers would then be erased. By this means, GEIS ensured that all data collected were consistent with the logical content of the script.

### A.7.4 Names and functions of GEIS script items

GEIS script item types had names consisting of a maximum of four characters. They are listed below along with their function in the script. The functions were identical for CATI and IVR scripts except where indicated otherwise.

**CALC** CALC inserted a calculation into the script that was executed when the item became active. The calculation might involve the value of any previous answer or imported values. They might only be simple expressions such as `Q2+3`, but could also be very complex.

**CALL** When a CALL item became active in an interview the telephone number of the respondent stored in the variable `STDPHONE` was called.

**CBDT** This item stored the date that the respondent should be called back.

**CBTM** This item stored the time that the respondent should be called back.

**CHCE** This item was used to ask a question that only had a limited set of possible answers, such as ‘yes’ and ‘no’.

In CATI mode, GEIS displayed a list box with between 1 and 50 text options, one of which might be selected by the interviewer.

In IVR mode, GEIS allowed the respondent to choose one option by pressing one of the keys 1, 2, ... 9 on their touchphone.

Refusals could be entered by including an option with code value `.R`.

**DO** This item started a DO loop. A DO loop might contain a series of other items that were executed multiple times. For example, the following questions could be replaced by a single item (not shown) within a DO loop.

```

“Has your oldest child been immunized?”
“Has your second oldest child been immunized?”
“Has your third oldest child been immunized?”
:

```

**ENDD** This item ended a DO loop.

**HUP** In CATI mode or IVR mode this item hung up a telephone call and, optionally, returned the call to another extension.

**INFO** In CATI mode this item displayed up to 12 lines of text on the screen. Additional lines could be scrolled into view. It was used to give instructions to the interviewer or to display text to be read out to the respondent.

In IVR mode this item played a message. After the message had played GEIS either: (a) immediately activated the next item, (b) paused briefly before moving on, (c) paused indefinitely until a key was pressed, (d) paused until a particular key was pressed, or (e) paused until a certain time elapsed.

**LINK** This item imported one or more variables from another data set. The data were imported during compilation but could also be updated immediately before an interview.

**LIST** This item was identical to a CHCE item except that the number of options that could be selected was unlimited. The options had to be specified in an external data set. A typical question was:

“In which country were you born?”

for which there could be very many possible answers.

This item was not implemented in IVR mode.

**LVLC** This item counted how many of a range of CHCE or MULT items in the script had had a specified option selected. For example, it may be used to check whether all of a set of CHCE questions were answered with “YES”.

**MAIL** This item sent an e-mail after the interview was completed. The address of the recipient, CC address, subject, and attachment, were specified in other variables.

**MULT** This item was used to ask a question that could have zero, one or more possible answers. In CATI mode GEIS displayed a list box with between 1 and 50 text options, one or more of which could be selected by the interviewer. An example of a typical question

“Which of the following diseases have you had:

Influenza,

Measles,

Chicken pox?”

This item was not implemented in IVR mode.

**NEW** If this item was activated during an interview it appended a new respondent record to the main data sets.

**NULL** This item was used to link together logic paths in a script. It enabled scripts to be simplified by breaking up lengthy SPSs into several shorter ones that were shared between several NULL items.

**NUM** The NUM item was used to ask a question that required a numeric answer. The answer could be any valid SAS numeric value. Numbers (e.g. 1, 2, 3, -7.1, 0.03), dates (e.g. 23jan1990, 230199, OCT3), and times (e.g. 12:40PM, 17:00, or 9:00:40), could be entered.

Upper and lower reasonableness and absolute limits could be specified. If the interviewer entered a value outside the reasonableness limits GEIS displayed a check box that had to be ticked before another item would be displayed. If the interviewer entered a value outside the absolute limits GEIS displayed an error message and refused to accept the answer.

With IVR mode, dates and times had to be entered using digits only. For example, the date 23Jan1950 was entered as 23011950.

Upper and lower reasonableness limits could be specified within the IVR mode as well. If the respondent entered a value outside the reasonableness limits GEIS played a message asking the respondent to confirm the value by pressing the hash key before the next item would be played. If the

respondent entered a value outside the absolute limits GEIS played an error message and the question was repeated.

Refusals were entered using the code value **.R**.

**NUMM** This item was similar to the NUM field but allowed up to 5 entry fields each with their own formats, informat, absolute limits, and reasonableness limits. An algebraic expression was used to set the value of a summary variable. A constraint expression ensured the field values were consistent.

Refusals were entered using the code value **.R**.

This item was not implemented in IVR mode.

**OPEN** This item was used to ask a question that required an open-ended answer. In CATI mode, the interviewer typed the verbatim answer into a text box.

In IVR mode, the respondent was asked to speak their answer, which was then recorded by GEIS.

Refusals were entered using a string “REFUSED”.

**RST** If this item became active the interview was reset, all data were erased, and the first item then became active again.

**SCAL** This item was used to define one or more numeric or character values. Although a LINK item could be used for the same function, the SCAL item value directly appeared in the script. It was also used in combination with DO loops.

**SCRP** This item was used to import another script fragment, called a sub-script, into the current script. This allowed a script to be broken up into a series of smaller pieces or a standard question domain to be imported.

**SMSG** This item appeared at the beginning of the script and did not form part of the script logic. It allowed a message to be defined for a given STATUS code. When an interview was begun with a matching STATUS code, the message in the SMSG item was displayed to the interviewer.

In IVR mode, a specific voice recording was played instead.

**STAT** This item was used to set the STATUS variable code value on exit from an interview. There had to be at least one STAT item in a script to set the STATUS variable code value to ‘CQ’ indicating a Completed Questionnaire. Usually there would also be other STAT items for codes like ‘CB’, ‘RT’, etc.

**TABL** This item displayed a table of 2 or 3 columns. This allowed entry of either a column of numbers, or a series of short open-ended answers. In the table the first column in each row contained a label. The second column was for numeric entry and the third for character entry. Buttons could be displayed instead of specific rows. This item type was used for complex questions that required multiple answers.

Refusals were entered by assigning the code value **.R** to a button.

This item was not implemented in IVR mode.



**TIME** When this item became active it stored the duration since the interview began. It only did this once, so that if the item became active at another time the duration was not changed.

**TITL** This item specified the title of the project and also set global options.

**TRNF** This item transferred a call to another number. It was used when the IVR and CATI modes were combined such as in the Hybrid method as described in Section 4.6.3 of Chapter 4.

## Appendix B

# Appendix for Chapter 4

### B.1 Pretest One

#### B.1.1 Sampling program for Pretest One

```
* Remove invalid records;
* Remove external prefixes from telephone numbers;
* Translate email addresses to user names;
data frame;
set datacati.frame;
  if stdphone=' ' then delete;
  if address=' ' then delete;
  if fullname=' ' then delete;
  stdphone=tranwrd(stdphone,"49246","46");
  address=tranwrd(lowercase(address),"@wallsend.newcastle.edu.au","");
  address=tranwrd(lowercase(address),"@wallsend.newcatle.edu.au","");
  address=tranwrd(lowercase(address),"@wallsend.newcastle.edu.au","");
run;

* Set frame and sample sizes;
%let m=61;
%let n=20;

* Draw sample;
data confid (drop=i j count) ;
  count=0;
  array obsnum (&m) _temporary_;
  do i=1 to &n;
    redo:
      select=ceil(uniform(111016)*&m);
      set work.frame point=select nobs=n;
      do j=1 to count;
        if (obsnum(j)=select) then goto redo;
      end;
      position=select;
      count=count+1;
      obsnum(count)=select;
      output;
    end;
  stop;
run;

* Add volunteers;
data t1;
```

```

    set frame;
    where (fullname="Dr Libby Campbell");
    position=5;
run;
data t2;
    set frame(obs=0);
    where (fullname="Ms Judy Foulkes");
    position=22;
run;
data confid;
    set confid t1 t2;
    selected=1;
run;

* Set requested times for the system to ring for some individuals;
data datacati.confid;
    set datacati.confid;
    start=.;
    if fullname="Dr Afaf Girgis"      then start='24aug1999:10:00:00'dt;
    if fullname="Mrs Tracey Smith"    then start='24aug1999:10:00:00'dt;
    if fullname="Mrs Pauline Chiarelli" then start='23aug1999:15:00:00'dt;
    format start datetime.;
    informat start datetime.;
run;

* Print sample in name order;
proc sort data=confid out=datacati.confid;
    by fullname;
run;

proc print data=datacati.confid;run;

```

### B.1.2 Participant instructions

The e-mail sent to Pretest One participants is shown in Figure B.1, the attached instruction sheet is shown in Figure B.2, and the assessment sheet is shown in Figure B.3.

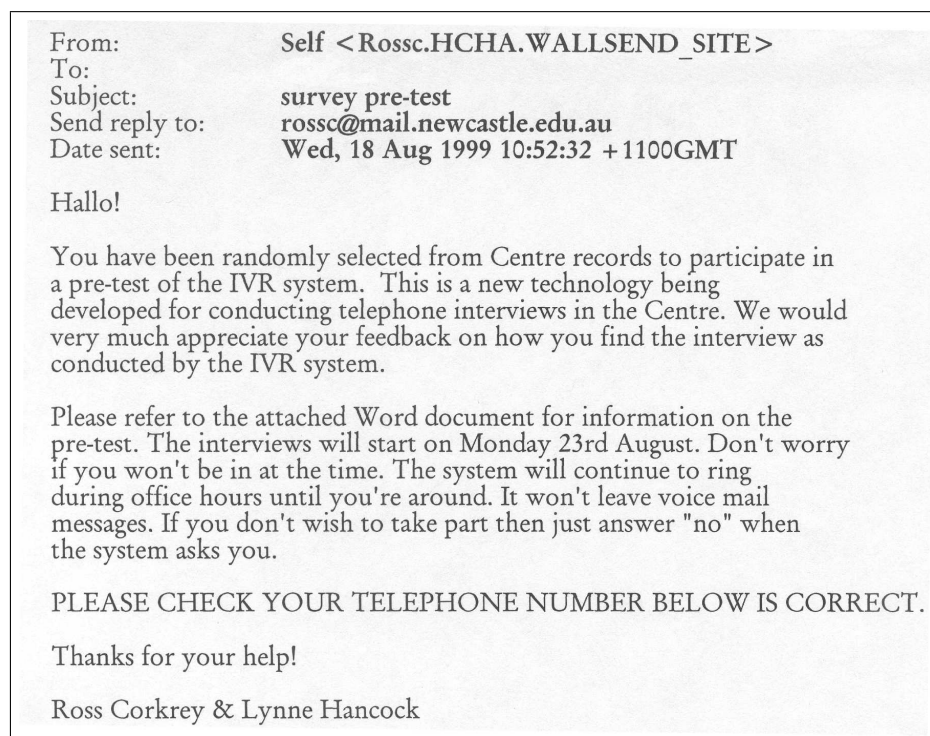


Figure B.1: E-mail sent to Pretest One participants.

## IVR Interface Pre-test

### The Project

An Interactive Voice Response (IVR) system to conduct telephone interviews has been developed in the Centre by Dr Lynne Hancock and Ross Corkrey. An IVR interview is very similar to a Computer-Assisted Telephone Interview (CATI), and in fact uses the same software, SAS, but the human interviewer is replaced by a high quality recorded voice. In IVR, the respondent provides answers by pressing the keys of a telephone. The study aims to compare IVR to conventional CATI as a tool to assess sensitive behaviours (illicit drug and alcohol use), and to assess cost-effectiveness and acceptability of IVR technology as a community survey tool. It is expected that the anonymity afforded by IVR technology will increase survey participation rates and increase accuracy of reporting. Pre-testing with a community-based sample is planned for September and will examine ease of use and user acceptability for the IVR interface. Following pre-test and modification, a pilot will be conducted in November with a sample of 400 randomly selected people throughout Australia. The pilot study will provide information on community participation rates, acceptability of IVR for different age groups, and completion rates for the survey and for individual items. The final survey developed by this process will be used to continue the study into Phase 2, which aims to compare IVR technology with existing and known acceptable CATI methods.

Note that the voice and script that will be used in the pilot will not necessarily be the one used in this pre-test.

### Your role

You have been randomly selected from Centre records to participate in a pre-test of the IVR system. We would like you to try out the system and feedback to us your experience. You don't have to do anything to prepare other than read this document. After you have used the system we would like you to complete the attached assessment form and return it to Ross Corkrey, room 20. Currently we are interested in the interface design, particularly the ease of responding to questions given the nature of the medium. Later on we will be conducting pre-testing on the script itself.

### Procedure

On Monday, 23rd August, at 10am the system will begin ringing numbers in the Centre. The exact time your call will be made will depend on how many interviews are ahead of yours, but we expect all numbers to have been called at least once within a few hours. If your line is busy the system will automatically ring again thirty minutes later. To avoid activating the voice mail system each number will only ring for 10 seconds. Once you have completed an interview the system will not ring you back again.

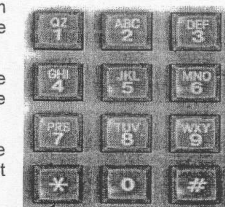
The first part of the interview allows you to select a more convenient time for the system to call back and conduct the interview.

During the interview you should assume that you have received a consent letter with information similar to this document and that you are the person nominated to answer the survey.

Each call plays a set of items consisting of spoken text and questions. When prompted you press keys on your phone. Which keys you should press will be specified within each item. The only keys you can press are the number keys (0-9), star key and hash key. Pressing the hash key repeats the current question. Pressing the star key moves to the previous question and lets you re-enter your answer.

**If you have any difficulty in using the system please contact Ross Corkrey on 46381.**

An example of a typical item is shown here:



Star key  
(backwards)

Hash key  
(repeat question)

"Were you born in a country other than Australia? Press 1 if yes. Press 2 if no. Press 3 if you don't wish to answer. Press 4 if you don't know."

Figure B.2: Instructions sent to Pretest One participants.

ID .....

**RETURN TO ROSS CORKREY, ROOM 20.**

**IVR Interface** **DATE:** \_\_\_\_\_

**Pre-test Assessment**

Q1. When you initially answered the call did you realise you were not talking to a person? ☐ Yes ☐ No

Comments: .....

Q2. Were there any questions with too many options so that it was difficult to remember them all. ☐ Yes ☐ No

Comments: .....

Q3. Do you think there should have been a greater space between questions? ☐ Yes ☐ No  
*This refers to the time delay between one question and the next.*

Comments: .....

Q4. Did you have any problems in responding to questions by pressing the number keys? ☐ Yes ☐ No

Please describe: .....

Q5. Did you repeat any questions using the hash key? ☐ Yes ☐ No

----- If NO, skip to Q6 -----

Q5 (a) If you repeated questions, did you find the process confusing? ☐ Yes ☐ No

Please describe: .....

D:\ivr\pretest1\quest.doc 19/08/99 1

Figure B.3: Assessment sheet used by Pretest One participants, page one.

ID .....

Q6. Did you arrange a call-back to a later interview? ☐ Yes ☐ No  
*A call-back is an interview arranged by the respondent on a date and time convenient to him or her, and before any data have been collected.*

----- If NO, skip to Q7 -----

Q6 (a) Did you find arranging call-backs confusing? ☐ Yes ☐ No

Please describe: .....

Q6 (b) Was the process of setting a call-back too lengthy? ☐ Yes ☐ No

Comments: .....

Q6 (c) Did the call-back arrive on the date/time you wanted? ☐ Yes ☐ No

Comments: .....

----- If NO, skip to Q7 -----

Q6 (d) When the system call-back came, did you find that answering the call was confusing? ☐ Yes ☐ No

Please describe: .....

Q7. Did you hang-up after starting an interview or before pressing any keys? ☐ Yes ☐ No

----- If NO, skip to Q8 -----

Q7 (a) Did the system ring back again? ☐ Yes ☐ No

----- If NO, skip to Q8 -----

Q7 (b) When the system ring back came, were there any parts of the call that you found confusing? ☐ Yes ☐ No

Please describe: .....

Q8. Did the system do anything unexpected, such as fall silent or hang up suddenly? ☐ Yes ☐ No

Please describe: .....

Q9. Did you find that it would have been helpful if the system could have repeated each keystroke back to you? ☐ Yes ☐ No  
*eg if the keystroke pressed corresponded to YES then the system could say "YES".*

Comments: .....

D:\iv\pretest\quest.doc 19/08/99 1

Figure B.3: Assessment sheet used by Pretest One participants, page two.

Figure B.3: Assessment sheet used by Pretest One participants, page three.



## B.2 Pretest Two

### B.2.1 Program to generate dial-in code from a user network name

```

/*
The data step converts letters in the name
to digits as they appear on telephone key pads
Address contains a user network name.
ID2 and ID3 contains the numeric code equivalent.
eg address=ROSSC -> ID2=76772.
*/

libname datacati 'n:\data\rossc\ivr2';
* Reading is network userlist;
data WORK.users ;
%let _EFIERR_ = 0; /* clear ERROR detection macro variable */
infile 'N:\DATA\ROSSC\ivr2\Users.lis' TRUNCOVER lrecl=61;
format address $30. ;
format FULLNAME $30. ;
input
    @1 address $30.
    @32 FULLNAME $30.
;
if _n_ ge 7 & _n_ le 158 then /* starting record/number of records */
do;
output;
if _ERROR_ then
call symput('_EFIERR_',1);
end;
run;
* Remove facilities and create CONFID variables;
data work.users;
set work.users;
if fullname^=' ';
length stdphone $ 20;
eligible=0;
selected=1;
start=.;
run;
* recode username as ID code;
data work.confid;
set work.users;
* recode as number;
id2=translate(lowercase(address),'11','qz');
id2=translate(id2,'222','abc');
id2=translate(id2,'333','def');
id2=translate(id2,'444','ghi');
id2=translate(id2,'555','jkl');
id2=translate(id2,'666','mno');
id2=translate(id2,'777','prs');
id2=translate(id2,'888','tuv');
id2=translate(id2,'999','wxy');
* alternative recoding for some phones;
id3=translate(lowercase(address),'222','abc');
id3=translate(id3,'333','def');
id3=translate(id3,'444','ghi');
id3=translate(id3,'555','jkl');
id3=translate(id3,'666','mno');
id3=translate(id3,'7777','pqrs');
id3=translate(id3,'888','tuv');
id3=translate(id3,'9999','wxyz');

```

From: "Ross Corkrey" <Rossc.HCHA.WALLSEND\_SITE>  
 To: #centre  
 Date: Wed, 1 Sep 1999 10:00:27 +1000  
 Subject: chocolate frogs & surveys  
 Reply-to: rossc@mail.newcastle.edu.au

If you like chocolate frogs (or not) please keep reading.

Yes, we are running another IVR pretest. This time, though, we want YOU to call US. Right now, we have the technology. Anyone who wants to (please) can dial in to the IVR automatic interviewing system and pretest it for us.

Just dial the system up and listen for instructions. There are 2 lines in use at the moment: 46063 and 46064, so if one is busy then try the other. We'll keep it running 24 hours a day for several days so there's plenty of opportunity to try it. (You never know, it could be fun)

To access the IVR, you do need a network username (eg rossc) to use the system. Once in remember this: + The hash (#) key repeats the question. + The star (\*) key moves back to previous questions. You can go back as far as you like and change your answers.

So what have chocolate frogs to do with this? If you phone into the system and test it for us, we will give you a chance to win a chocolate frog by guessing a number! The frog (or frogs!) will go to the nearest guessers.

Ross  
 Ross Corkrey Room 24 Ph: 246381

Figure B.4: E-mail sent to Pretest Two participants.

```

if id2=id3 then same=1;else same=0;
id=input(id2,16.);
run;
* check for entry code differences;
proc fsvview data=confid;
  var id id2 id3 same address ;
  format same id 16. id2 id3 address $16.;
  where not same;
run;
* Sort into ID order and save;
proc sort data=work.confid out=atacati.confid;
  by id;
run;

```

## B.2.2 Participant instructions

The e-mail sent to Pretest Two participants is shown in Figure B.4, and the assessment sheet is shown in Figure B.5.

Name .....	ID .....
------------	----------

RETURN TO ROSS CORKREY, ROOM 20.

IVR Interface	DATE: .....
Pre-test 2 Assessment	

Q1. Were there any questions with too many options so that it was difficult to remember them all?      Yes      No

Comments: .....

Q2. Do you think there should have been a greater space between questions?      Yes      No  
*This refers to the time delay between one question and the next.*

Comments: .....

Q3. Did you have any problems in responding to questions by pressing the number keys?      Yes      No

Please describe: .....

Q4. Did you repeat any questions using the hash key?      Yes      No

---- If NO, skip to Q5 ----

Q4 (a) If you repeated questions, did you find the process confusing?      Yes      No

Please describe: .....

C:\WINDOWS\TEMP\SWPM3801.doc
07/09/99
1

Figure B.5: Assessment sheet sent to Pretest Two participants, page one.

Figure B.5: Assessment sheet sent to Pretest Two participants, page two.

## B.3 Pilot One

### B.3.1 Sampling program for Pilot One

```

* Subsample nos for ivr pretest pilot;

/*
1. find suburbs within 10 kilometres.
2. use these to find postcodes within same distance.
3. merge with nsw electronic directory.
4. only keep cases with tel. nos starting with "(02) 49" to
   eliminate mobile phones.
5. only keep cases with nonmissing address info.
6. SRS from within frame.
*/

libname phone 'd:\phone nos';
libname geog 'n:\data\sascati\callcost';
libname sgweb 'n:\statman1';

* get list of postcodes;
proc sort data=sgweb.postcode out=pc nodupkey;by locality;run;

* get list of suburbs within 10km;
data stats;
set geog.stats;
if distance<10 & geogtype='SUBURB' ;
run;
proc sort data=stats out=stats nodupkey;by locality;run;

* select postcodes based on closest suburbs;
data merged;
merge pc stats(in=a);
by locality;
if a and postcode>.;
keep postcode geogtype state locality;
run;

* get nonmonbile phones nos;
data phone;
set phone.nsw1998;
if index(stdphone,"(02) 49");
run;

* select frame;
proc sort data=merged;by postcode;run;
proc sort data=phone;by postcode;run;
data frame;
merge merged(in=a) phone;
by postcode;
if a & index(stdphone,"(02) 49") & strtnumb^=' '
  & strtname^=' ' & suburb^=' ' & name^=' ' & postcode^=.;
run;
proc sort data=frame out=f nodupkey;by suburb;run;

* get no of obs in frame;
data _null_;
dsid=open("work.frame");
nobs=attrn(dsid,"nobs");
call symput('m',nobs);
rc=close(dsid);
run;

```

```

data _null_;
put "number of cases in frame=" "&m";
run;

* required number of obs;
%let n=50;

* extract sample;
data sample (drop=i j count) ;
count=0;
array obsnum (&m) _temporary_;
do i=1 to &n;
redo:
select=ceil(uniform(587816)*&m);
set work.frame point=select nobs=n;
do j=1 to count;
if (obsnum(j)=select) then goto redo;
end;
position=select;
count=count+1;
obsnum(count)=select;
output;
end;
stop;
run;

data sample;
set sample;
rnd=uniform(-1);
run;

data sample;
set sample;
rename name=fullname;
address=trim(street)||" "||trim(suburb)||" "||put(postcode,4.);
format name $20. address $40.;
eligible=0;
selected=1;
ID=1000+_N_;
run;

proc sort data=sample;by rnd;run ;

data sample1 sample2;
set sample;
if _n_ le &n/2 then output sample1;else output sample2;
drop rnd position;
run;

* create confid datasets;
libname dc1 'n:\data\sascati\projects\ivr3_out';
proc sort data=sample1 out=dc1.confid;by suburb fullname strtnumb strtname ;run ;
libname dc2 'n:\data\sascati\projects\ivr3_in';
proc sort data=sample2 out=dc2.confid;by suburb fullname strtnumb strtname ;run ;

title 'Outbound sample';
proc print data=dc1.confid;
var fullname address stdphone;
run;

title 'Inbound sample';
proc print data=dc2.confid;

```

```
var fullname address stdphone;
run;
```

### B.3.2 Program to calculate distances from Wallsend

The program below uses longitudes and latitudes of placenames downloaded from <http://kaos.erin.gov.au/database/MAN200R.html> where they had been categorized as suburbs, towns, urban, or localities. Distances from Newcastle were calculated using Equation 3.1 of Chapter 3.

```
* Loads geographic data and calculates distances from Wallsend;

libname callcost 'n:\data\sascati\callcost';

data suburb ;
    length locality $50 str dum1 dum2 $ 200 state $10 geogtype $ 8;
    infile 'n:\data\sascati\callcost\new.txt' trunccover;
    input locality 5-55;
    input state 5-15;
    input geogtype 5-55;
    input str 5-200;
    input dum1 5-200;
    input dum2 5-200;
    geogtype='SUBURB';
    output;
    drop dum1 dum2 ;
run;

proc sort data=suburb nodupkey;
    by locality geogtype state;
run;

data towns ;
    length locality $50 str dum1 dum2 $ 200 state $10 geogtype $ 8;
    infile 'n:\data\sascati\callcost\new_t.txt' trunccover;
    input locality 5-55;
    input state 5-15;
    input geogtype 5-55;
    input str 5-200;
    input dum1 5-200;
    input dum2 5-200;
    geogtype='TOWN';
    output;
    drop dum1 dum2 ;
run;

proc sort data=towns nodupkey;
    by locality geogtype state;
run;

data urban ;
    length locality $50 str dum1 dum2 $ 200 state $10 geogtype $ 8;
    infile 'n:\data\sascati\callcost\urban.txt' trunccover;
    input locality 5-55;
    input state 5-15;
    input geogtype 5-55;
    input str 5-200;
    input dum1 5-200;
    input dum2 5-200;
    geogtype='URBAN';
    output;
```

```

    drop dum1 dum2 ;
run;

proc sort data=urban nodupkey;
    by locality geogtype state;
run;

data local ;
    length locality $50 str dum1 dum2 $ 200 state $10 geogtype $ 8;
    infile 'n:\data\sascati\callcost\local2.txt' trunccover;
    input locality 5-55;
    input state 5-15;
    input geogtype 5-55;
    input str 5-200;
    input dum1 5-200;
    input dum2 5-200;
    geogtype='LOCALITY';
    output;
    drop dum1 dum2 ;
run;

proc sort data=local nodupkey;
    by locality geogtype state;
run;

data comb;
    set suburb towns local urban ;
    locality=tranwrd(locality,'[ EnviroMaps Link ]','');
    lat=input(scan(str,2,"()"),12.);
    long=input(scan(str,3,"()"),12.);
    label lat='Latitude (degrees)';
    label long='Longitude (degrees)';
    locality=upcase(locality);
    drop str;
run;


* Distance from Wallsend to each location;
data nsw;
    set comb;
    PI=3.141592654;
    * Wallsend - first point latitude;
    L1=-32.89996 * PI / 180;
    * Wallsend - first point longitude;
    G1=151.66664 * PI / 180;
    * Latitude on second point;
    L2=lat * PI / 180;
    * Longitude on second point;
    G2=long * PI / 180;
    * Differnce in long;
    DG = G2 - G1;
    * Differnce in lat;
    DL = L2 - L1;
    * Distance;
    tmp = arcos ( SIN(L1) * SIN(L2) + COS(L1) * COS(L2) * COS(DG)) ;
    Distance = 1.852 * 60 * tmp * 180 / PI;
    drop DG DL L1 L2 G1 G2 tmp pi;
    label distance='Distance from Wallsend in Kilometres';
run;
proc sort data=nsw out=callcost.nsw nodupkey;
    by locality geogtype state;
run;

```



### **B.3.3 Information letters and instruction sheets**

The letter sent to participants in the outbound mode is shown in Figure B.6. The letter sent to participants in the inbound mode is shown in Figure B.7. The instruction sheet sent to participants in both modes is shown in Figure B.8.



## Hunter Centre for Health Advancement

**Locked Bag No 10**  
**WALLSEND NSW 2287**  
 Phone: (02) 49246 326

The household  
 Street number and name  
 SUBURB  
 NSW Post code

Dear Household Members


The Hunter Centre for Health Advancement has developed a new method for conducting telephone interviews that uses a computer to ask questions. This project is part of the doctoral studies being undertaken by Mr Ross Corkrey, under the supervision of Dr Lynne Hancock. We are currently conducting a pre-test of a survey of alcohol and illicit drug use in the Australian community developed using this new method. The advantage of this new method over traditional interviewing is that it is completely anonymous since only a computer knows what people answer. As part of this process, we would like to invite community members to be involved in helping us to develop the survey.

**How was your household chosen?** It is important that we get the views of as wide a range of people as possible. So we would like to talk to people from 50 randomly chosen households throughout the Hunter. We obtained your household address at random from the Telstra White Pages telephone listings.

**Do you have to talk to us?** While your views are important to us, participation in this study is voluntary, and you do not have to answer our questions if you do not want to. If you don't wish to be contacted then please ring Ross Corkrey on (02) 49246381. To ensure that our results represent those of the general community we need to select a person in your household at random. So the person the system will want to talk to is the person in your household who is 18 or over and who had the last birthday. We'd like you to decide who this person is before the system phones you.


**What is the survey about?** The survey will ask questions about alcohol use and drug use, and some background information. After the system has rung you, the Project Manager (Mr Ross Corkrey) will ring again within 2 weeks. Mr Corkrey will ask the person who did the interview what they thought of it and how we can improve how the questions are asked.

**What will happen to the information you give us?** All the answers to our questions will be treated in STRICT CONFIDENCE. All the data will be entered onto computer and stored securely at the Hunter Centre for Health Advancement. As this is a pre-test of the survey only, your answers to the survey won't be analysed. Your ideas about how to improve the survey will be used to improve the survey for next time. Your name, phone number or address will not be located near your other information. Once the survey is complete all names, addresses, and telephone numbers will be destroyed. Since the system is run by computer the only time your telephone number will be accessible to anyone at all is when Mr Corkrey rings you to ask what you thought of the first interview. This person will not know the answers you gave during the first interview. In fact, no person will ever be able to connect your telephone number to your answers since this is locked inside the computer. The



Hunter Health

A commitment to excellence in health promotion by Hunter Health and  
the Faculty of Medicine and Health Sciences University of Newcastle



The UNIVERSITY  
of NEWCASTLE

Figure B.6: Information letter sent to Pilot One participants in the outbound calling mode, page one.

report of our findings will be for group results only. Individual people or their answers will not be identified in the survey results.

**What happens now?** Within 2 weeks of receiving this letter, the interviewing system at the Hunter Centre for Health Advancement will contact your household. If your line is busy the system will automatically try again thirty minutes later. When you answer, you will hear a recorded voice asking questions. It will ask you to press keys on your phone in order to answer the questions.

**Phones.** To participate in the survey you must have the right kind of telephone. So the first thing the system will do is ask you to press key "5" on your telephone. If you don't have a telephone with keys, just wait and the system will thank you and then hang up. You will not be called again.

If you do have the right type of keyed phone the system will ask you if you want your household to take part in the survey. You can say "no" if you wish and then the system will not ring you back. If you say "yes", the system will ask to speak to the person who is 18 or over and who had the last birthday. The first thing the system asks this person is whether he or she actually wants to do the interview. Remember, that person doesn't have to do the interview at the time the system rings, since he or she can arrange a better time. Since this is a computer system and not a person, we would like to ask you not to hang up before this point, so that you can let the system know not to phone again if that is what you choose. You can still hang up at any point. If you do the system will not ring back again.

The attached page tells you more about how the phone survey system will work and what telephone keys will mean when it phones you. Please keep this page near the phone.

The survey should take about 10 minutes. Once you have completed an interview the system will not ring you back again. However, Mr Ross Corkrey will contact you to see if you'd like to talk to him about how you found the survey and how to make it better.

If you have any questions about this study before then, please do not hesitate to contact Ross Corkrey on 49246 381. Thank you for your time and consideration.


Yours sincerely  
HUNTER CENTRE FOR HEALTH ADVANCEMENT

Dr Lynne Hancock  
Principal Investigator

Mr Ross Corkrey  
Project Manager

If you have any complaint concerning the manner in which a research project is conducted, it may be given to the researcher, or if an independent person is preferred, to the Professional Officer, HAREC, C/- HAHS, Locked Bag 1, New Lambton Heights 2305, telephone (02) 4921 4950. Alternatively, you can contact The Human Research Ethics Officer, Research Branch, The Chancellery, University of Newcastle, University Drive, Callaghan NSW 2308, telephone (02) 4921 6333.

Figure B.6: Information letter sent to Pilot One participants in the outbound calling mode, page two.



## Hunter Centre for Health Advancement

**Locked Bag No 10**  
**WALLSEND NSW 2287**  
 Phone: (02) 49246 326

The household  
**Street number and name**  
**SUBURB**  
**NSW Post code**

Dear Household Members


The Hunter Centre for Health Advancement has developed a new method for conducting telephone interviews that uses a computer to ask questions. This project is part of the doctoral studies being undertaken by Mr Ross Corkrey, under the supervision of Dr Lynne Hancock. We are currently conducting a pre-test of a survey of alcohol and illicit drug use in the Australian community developed using this new method. The advantage of this new method over traditional interviewing is that it is completely anonymous since only a computer knows what people answer. As part of this process, we would like to invite community members to be involved in helping us to develop the survey.

**How was your household chosen?** It is important that we get the views of as wide a range of people as possible. So we would like to talk to people from 50 randomly chosen households throughout the Hunter. We obtained your household address at random from the Telstra White Pages telephone listings.

**Do you have to talk to us?** While your views are important to us, participation in this study is voluntary, and you do not have to answer our questions if you do not want to. If you don't wish to be contacted then please ring Ross Corkrey on (02) 49246381. To ensure that our results represent those of the general community we need to select a person in your household at random. So we would like to invite the person in your household who is 18 or over and who had the last birthday to phone into our system to complete a telephone computer survey.


**What is the survey about?** The survey will ask questions about alcohol use and drug use, and some background information. After you have rung into the system and completed a survey, the Project Manager (Mr Ross Corkrey) will ring again within 2 weeks. Mr Corkrey will ask the person who did the interview what they thought of it and how we can improve how the questions are asked.

**What will happen to the information you give us?** All the answers to our questions will be treated in STRICT CONFIDENCE. All the data will be entered onto computer and stored securely at the Hunter Centre for Health Advancement. As this is a pre-test of the survey only, your answers to the survey won't be analysed. Your ideas about how to improve the survey will be used to improve the survey for next time. Your name, phone number or address will not be located near your other information. Once the survey is complete all names, addresses, and telephone numbers will be destroyed. Since the system is run by computer the only time your telephone number will be accessible to anyone at all is when Mr Corkrey rings you to ask what you thought of the first interview. This person will not know the answers you gave during the first interview. In fact, no person will ever be able to connect your telephone number to your answers since this is locked inside the computer. The



HUNTER HEALTH

A commitment to excellence in health promotion by Hunter Health and  
the Faculty of Medicine and Health Sciences University of Newcastle



The UNIVERSITY  
of NEWCASTLE

Figure B.7: Information letter sent to Pilot One participants in the inbound calling mode, page one.

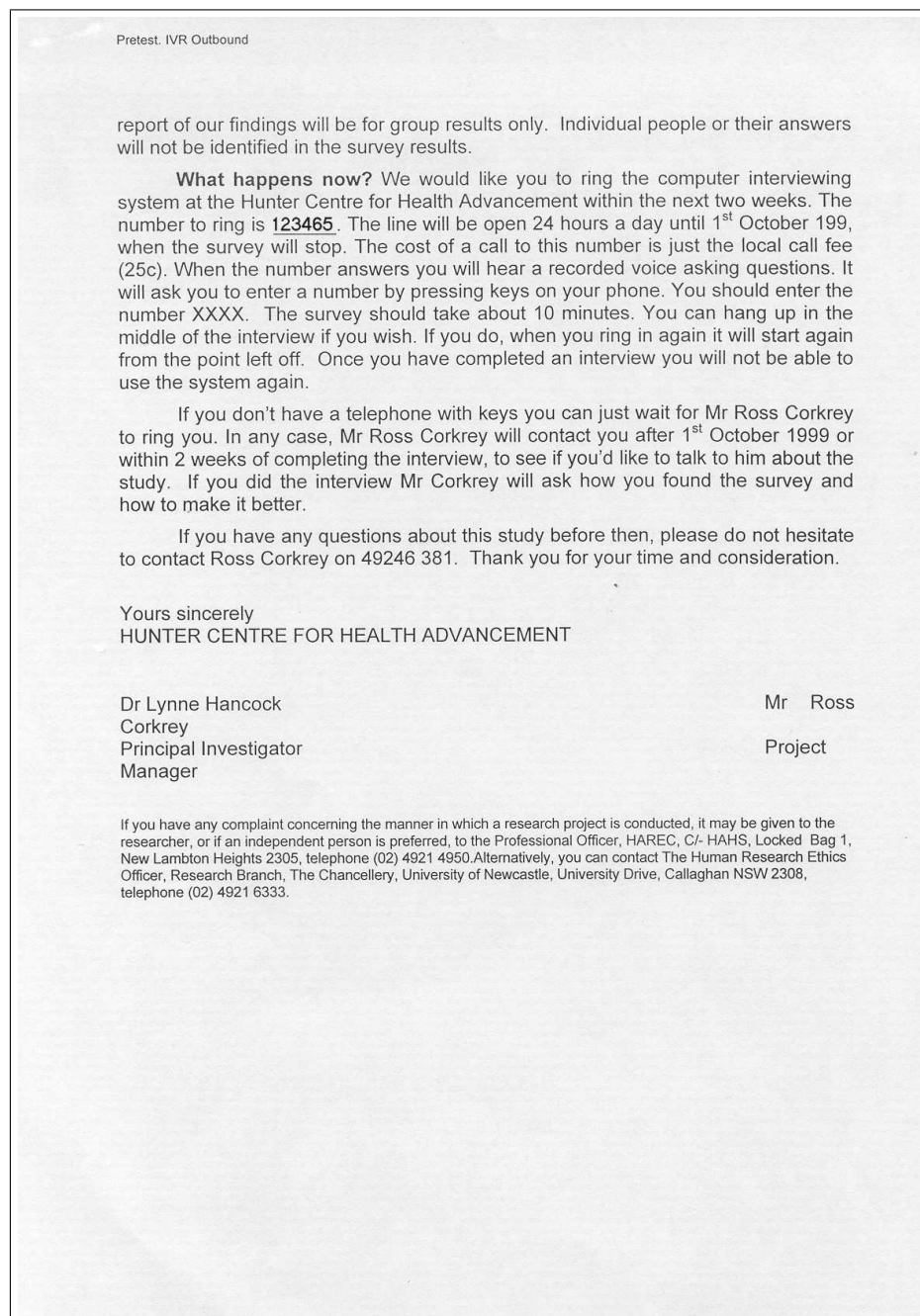



Figure B.7: Information letter sent to Pilot One participants in the inbound calling mode, page two.

**Please keep this page near your phone**

- The only keys you can press are the number keys (0, 1, 2, 3, 4, 5, 6, 7, 8, 9), the star key (\*) and the hash key (#).
- Press the hash key (#) to repeat the question. You can do this at any time.
- Press the star key (\*) to return to the previous question.
- To stop the interview and continue it later on press key 9.
- If you know the answer to a question and would rather not listen to all the options to be read out to you, you can go ahead and press the key you want.



Star key (backwards)
Hash key (repeat question)

Here is an example of a typical question:

"Were you born in a country other than Australia? Press 1 if yes. Press 2 if no. Press 3 if you don't wish to answer. Press 4 if you don't know."

If you have any questions about this study please do not hesitate to contact Ross Corkrey on 49246 381.

If you have any complaint concerning the manner in which a research project is conducted, it may be given to the researcher, or if an independent person is preferred, to the Professional Officer, HAREC, C/- HAHS, Locked Bag 1, New Lambton Heights 2305, telephone (02) 4921 4950.

Figure B.8: Instruction sheet sent to all Pilot One participants.

## B.4 Pilot Two

### B.4.1 Sampling program for Pilot Two.

```

libname phone 'd:\phone';
libname aria 'n:\data\statsgrp\data\ariarura';
options compress=yes;

* Sort ARIA codes by postcode;
proc sort data=aria.ariapost out=aria;
by postalar;
run;

data aria;
set aria;
rename postalar=post;
run;

/*-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
DO NOT RUN THIS BIT UNLESS RECREATING THE FRAME

* Sort phone nos by postcode;
proc sort data=phone.nswpost out=phone;by postcode;run;
data phone;
    set phone;
    post=input(trim(left(postcode)),18.);
run;

* merge in aria codes;
data x2;
merge aria (in=a) phone (in=b drop=ariascor ariacate);
by post;
if b;
run;

* remove mobiles ;
data phone.frame;
    set x2;
    if address^=' ' & name^=' ' & postcode^=' ';
    if area_co in ('02','03','07');
run;

-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+*/

* remove cases with missing addresses, telephone nos, or zero street numbers;
data frame;
    set phone.frame;
    if ariacate="HA" & var13^=' ' & var12^=' ' &
        suburb^=' ' & ac_phon^=' ' & var13^='0';
run;

* get no of obs in frame;
data _null_;
dsid=open("frame");
nobs=attrn(dsid,"nobs");
call symput('m',nobs);
rc=close(dsid);
run;
data _null_;
put "number of cases in frame=" "&m";
run;

```

```

proc univariate data=frame plot;
var ariascor;
run;

* required number of obs;
%let n=30;

* extract sample;
data sample (drop=i j count) ;
  count=0;
  array obsnum (&m) _temporary_;
  do i=1 to &n;
    redo:
    select=ceil(uniform(38299)*&m);
    set work.frame point=select nobs=n;
    do j=1 to count;
      if (obsnum(j)=select) then goto redo;
    end;
    position=select;
    count=count+1;
    obsnum(count)=select;
    output;
  end;
stop;
run;
data sample;
  set sample;
  rename var12=strtrname;
  rename var13=strtrnumb;
run;
proc sort data=sample ;
by suburb name strtrnumb strtrname ;
run ;
* create confid datasets;
title 'Hybrid Pilot Sample';
libname dc2 'n:\data\sascati\projects\ivr\ivr4\hybcati';
data dc2.confid;
  set sample;
  rename name=fullname;
  rename ac_phon=stdphone;
  address=trim(address)||" "||trim(suburb)||" "||put(postcode,4.);
  format name $20. address $40.;
* All are eligible;
  eligible=0;
* Make ID variable;
  ID=_N_;
* Create some negative IDs for practice cases;
  if _n_ ge 27 then ID=-_N_;
run;
%include 'n:\data\sascati\bin\distance.sas';
data conf;set dc1.confid;run;
%distance(dsin=conf);
proc print data=conf;
var id fullname address stdphone distance selected eligible;
format selected eligible 5.;
run;
proc freq data=conf;
table suburb;
where selected=1;
run;


```



```
title 'IVR4 - HYBIVR Sample';  
proc print data=dc2.confid;  
var fullname address stdphone selected eligible;  
id id;  
run;
```

### B.4.2 Information letter sent to participants

The letter sent to Pilot Two participants is shown in Figure B.9. The instruction sheet sent to Pilot Two participants is shown in Figure B.10.



## Hunter Centre for Health Advancement

**Locked Bag No 10**  
**WALLSEND NSW 2287**  
 Phone: (02) 49246 326

The household  
**Street number and name**  
**Suburb**  
 NSW 2036

Dear Household Members

The Hunter Centre for Health Advancement is doing a survey of technology, alcohol and drug use. We will be ringing households chosen at random and asking one person some simple questions. Since your household was one of those chosen we will ring you within weeks of getting this letter. We would still talk to you even if you think this survey does not apply to your household since we have to ask the same questions of everyone. The whole interview will only take around 10 minutes to do.

**Who are we?** This project is part of the doctoral studies being undertaken by Mr Ross Corkrey, under the supervision of Dr Lynne Hancock. The Hunter Centre for Health Advancement is a unit of the University of Newcastle and the Hunter Area Health Service.

**How was your household chosen?** It is important that we get the views of as wide a range of people as possible. Your household address was chosen at random from the Telstra White Pages telephone listings.

**Do you have to talk to us?** While your views are important to us, participation in this study is voluntary, and you do not have to answer our questions if you do not want to.

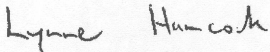
**What is the survey about?** The survey will ask questions about technology, alcohol and drug use. We are testing different ways of doing telephone surveys of health matters.

**What will happen to the information you give us?** The results of the survey will be used in health planning and research. All the answers to our questions will be treated in **STRICT CONFIDENCE**. Your answers will be entered onto a computer and stored securely at the Hunter Centre for Health Advancement. We will not record anyone's name. Your phone number and address will not be located near your other information. Once the survey is complete everyone's addresses and telephone numbers will be destroyed.

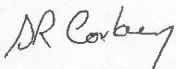
**What happens now?** An interviewer from the Centre for Health Advancement will ring you within the next two weeks. To make sure we interview someone in your household at random the interviewer will ask for the person in your household who is 18 or over and who had the last birthday. A few questions in the survey will be asked using a recorded voice. The interviewer will explain how this works.

**What if you need more information?** If you have any questions about this study before then, please do not hesitate to contact Ross Corkrey on (02) 4924 6381. If you want to check this is a genuine survey please ring the Projects Info Line at the Hunter Centre for Health Advancement on (02) 4924 6166. Thank you for your time and consideration.

Yours sincerely




**HUNTER CENTRE FOR HEALTH ADVANCEMENT**  
 Dr Lynne Hancock  
 Principal Investigator




**Mr Ross Corkrey**  
 Project Manager

If you have any complaint concerning the manner in which a research project is conducted, it may be given to the researcher, or if an independent person is preferred, to the Professional Officer, HAREC, C/- HAHS, Locked Bag 1, New Lambton Heights 2305, telephone (02) 4921 4950. Alternatively, you can contact The Human Research Ethics Officer, Research Branch, The Chancellery, University of Newcastle, University Drive, Callaghan NSW 2308, telephone (02) 4921 6333.



HUNTER HEALTH

A commitment to excellence in health promotion by Hunter Health and  
 the Faculty of Medicine and Health Sciences University of Newcastle



The UNIVERSITY  
 of NEWCASTLE

Figure B.9: Information letter sent to Pilot Two participants.

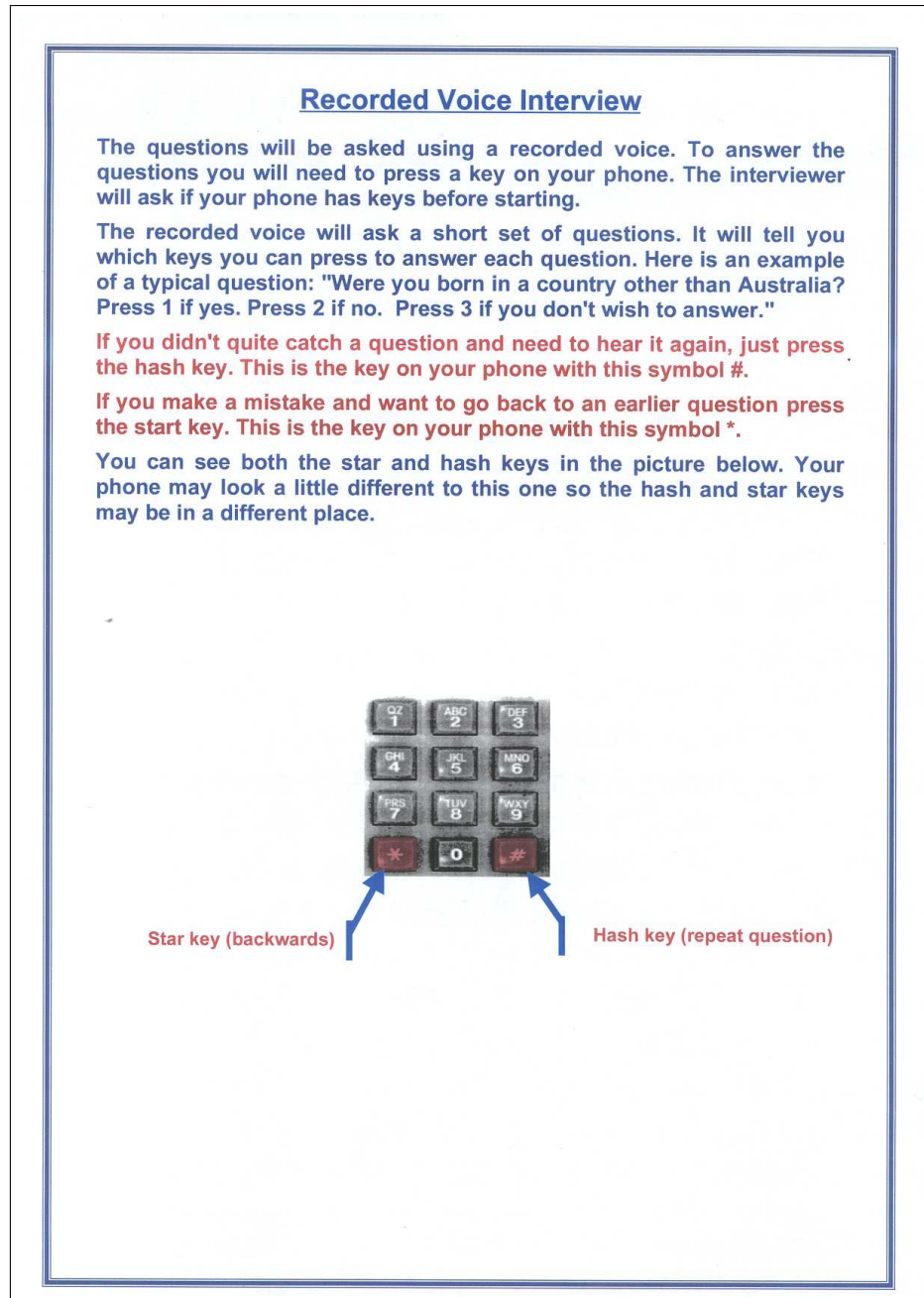


Figure B.10: Instruction sheet sent to Pilot Two participants.

Step 1. Below, *GEIS* reports running on line 1: ‘SYSTEM READY’. It is waiting for an incoming call.

	<i>INTERVW</i>	<i>CALLNO</i>	<i>ACTION</i>	<i>IDX1</i>	<i>STATUS</i>
<b>Row 1:</b>	LINE1		SYSTEM READY		
<b>Row 2:</b>					

Step 2. As shown in row 2, the interviewer transfers respondent with ID 1162 from extension ‘46381’ to the RVS. *GEIS* is interviewing the respondent on line 1.

	<i>INTERVW</i>	<i>CALLNO</i>	<i>ACTION</i>	<i>IDX1</i>	<i>STATUS</i>
<b>Row 1:</b>	LINE1		SYSTEM READY		
<b>Row 2:</b>		46381	SUB	1162	

Step 3. The respondent is transferred back to interviewer on extension ‘46381’. The IVR interview was completed as indicated by status code ‘CQ’ in row 2.

	<i>INTERVW</i>	<i>CALLNO</i>	<i>ACTION</i>	<i>IDX1</i>	<i>STATUS</i>
<b>Row 1:</b>	LINE1		SYSTEM READY		
<b>Row 2:</b>		46381	DONE	1162	‘CQ’

Figure B.11: Example of states of the AUTOCALL data set during the stages of the Hybrid method.

### B.4.3 Use of the AUTOCALL data set within the Hybrid method.

The AUTOCALL data set was used to control the transfer of calls between a CATI interviewer and *GEIS* running on the RVS. The CATI interviewers and *GEIS* simultaneously accessed and updated individual rows of the AUTOCALL data set by using the share server (see Section 3.5.3 of Chapter 3).

An example showing the progress of a call is shown in Appendix Figure B.11. In this example, *GEIS* initially reports that it is running on row one (“SYSTEM READY”). If this line is not present the CATI interviewer will see a message indicating that the call may not be transferred. The CATI interviewer presses the transfer button on the CATI screen, which causes the transfer to be registered in the AUTOCALL data set by updating row two (“SUB”). From this point *GEIS* will expect the call. When the call is transferred, *GEIS* runs the IVR script and conducts the interview. Once the IVR interview is complete, the call is transferred back to the CATI interviewer, and row two of the AUTOCALL data set is updated to indicate this (“INTERVIEW DONE”).

While the IVR interview is occurring, the interviewer can query the AUTOCALL data set using the CATI system to determine if the IVR interview has been completed, partly-completed, or terminated.

## Appendix C

# Appendix for Chapter 5

### C.1 The five-item AUDIT instrument

The 5-item Alcohol Use Disorders Identification Test (AUDIT)<sup>193</sup> was used in this study. It is shown below.

How often do you have a drink containing alcohol?	
Never	0
Monthly or less	1
Two to four times a month	2
Two to three times a week	3
Four or more times a week	4
How many drinks containing alcohol do you have on a typical day when you are drinking?	
1 to 2	0
3 to 4	1
5 to 6	2
7 to 9	3
10 or more	4
How often during the last year have you found you were not able to stop drinking once you had started?	
Never	0
Less than monthly	1
Monthly	2
Weekly	3
Daily or almost daily	4
How often during the last year have you failed to do what was normally expected from you because of drinking?	
Never	0
Less than monthly	1
Monthly	2
Weekly	3
Daily or almost daily	4
Has a relative or friend, or a doctor or other health worker been concerned about your drinking or suggested you cut down?	
No	0

Yes, but not in the last year	2
Yes, during the last year	4

The overall score is obtained by summing the scores of all items. The maximum score is 20. A score of five or more indicates hazardous or harmful consumption.

## C.2 Sampling program

```
* Sample from Australia using SRS;

libname phone 'd:\phone'; /* NSW white pages */
libname aria 'n:\data\statsgrp\data\ariarura';
options compress=yes;
options comamid=tcp remote=apc112;

options compress=yes;
libname rphone 'd:\phone';
* remove cases with missing addresses, telephone nos, or zero street numbers;
data frame(compress=yes);
    set rphone.aust;
    if strtrnumb=' ' & strtrname=' ' &
        suburb=' ' & stdphone=' ' & fullname='0' &
        ariascor>.;
run;
* get no of obs in frame;
data _null_;
dsid=open("work.frame");
nobs=attrn(dsid,"nobs");
call symput('m',nobs);
;rc=close(dsid);
run;
data _null_;
put "number of cases in frame=" "&m";
run;
* required number of obs including dummies;
%let n=1500;
* extract sample;
data sample (drop=i j count) ;
    count=0;
    array obsnum (&m) _temporary_;
    do i=1 to &n;
        redo:
        select=ceil(uniform(3111119)*&m);
        set work.frame point=select nobs=n;
        do j=1 to count;
            if (obsnum(j)=select) then goto redo;
        end;
        position=select;
        count=count+1;
        obsnum(count)=select;
        output;
    end;
    stop;
run;
data sample;
    set sample;
    rand=uniform(-1);
run;
```

```

proc sort data=sample out=sample(drop=rand) ;
by rand;
run ;
data sample2;
    length id selected eligible 8;
    set sample;
    format fullname $20. address $40.;
* All are eligible;
    eligible=0;
* Make ID variable;
    ID=_N_+2500;
    rename area_code=areacode;
    drop longitude latitude;
    postcode=trim(left(postcode));
run;
proc sort data=sample2 ;
by postcode suburb fullname strtnumb strtname ;
run ;
*CATI/Hybrid1;
* create confid datasets;
/*libname dc1 'n:\data\sascati\projects\ivr\ivr5\hybcati';
data dc1.confid;
    set sample2;
run;
*/

*IVR;
* required number of obs including dummies;
%let n=1300;
* extract sample;
data sample (drop=i j count) ;
    count=0;
    array obsnum (&m) _temporary_;
    do i=1 to &n;
        redo:
        select=ceil(uniform(31241519)*&m);
        set work.frame point=select nobs=n;
        do j=1 to count;
            if (obsnum(j)=select) then goto redo;
        end;
        position=select;
        count=count+1;
        obsnum(count)=select;
        output;
    end;
    stop;
run;
data sample;
    set sample;
    rand=uniform(-1);
run;
proc sort data=sample out=sample(drop=rand) ;
by rand;
run ;
data sample2;
    length id selected eligible 8;
    set sample;
    format fullname $20. address $40.;
* All are eligible;
    eligible=0;
* Make ID variable;
    ID=_N_;

```

```

* Select away the last 300 buffer;
  if _n_ le 1200 then selected=1;else selected=0;
* Create some negative IDs for practice cases;
  if _n_ gt 1495 then ID=-1501+_N_;
  rename area_code=areacode;
  drop longitude latitude;
  postcode=trim(left(postcode));
run;
proc sort data=sample2 ;
by postcode suburb fullname strtnumb strtname ;
run ;
* create confid datasets;
/*libname dc2 'n:\data\sascati\projects\ivr\ivr5\ivr';
data dc2.confid;
  set sample2;
run;
*/

*Hybrid II;
options comamid=tcp remote=apc127;
filename rlink
  'd:\program files\sas institute\sas\v8\connect\saslink\tcpwin.scr';
signon apc127;
rsubmit;
libname rphone 'd:\phone';
* remove cases with missing addresses, telephone nos, or zero street numbers;
data frame(compress=yes);
  length areacode $10;
  set rphone.aust;
  if strtnumb^=' ' & strtname^=' ' &
    suburb^=' ' & stdphone^=' ' & fullname^='0' ;
  areacode=scan(stdphone,1,'()');
  * remove mobiles etc ;
  if indexc(stdphone,'()')=0 then delete;
  if left(areacode)='01' then delete;
  else if left(areacode)='04' then delete;
  else if left(areacode)='05' then delete;
  else if areacode='1300' then delete;
  else if areacode='1800' then delete;
  else if areacode='1900' then delete;
run;
endrsubmit;
rsubmit;
* get no of obs in frame;
data _null_;
dsid=open("work.frame");
nobs=attrn(dsid,"nobs");
call symput('m',nobs);
rc=close(dsid);
run;
data _null_;
put "number of cases in frame=" "&m";
run;
* required number of obs including dummies;
%let n=750;
* extract sample;
data sample (drop=i j count) ;
  count=0;
  array obsnum (&m) _temporary_;
  do i=1 to &n;
    redo:
      select=ceil(uniform(6782568)*&m);

```



```


        set work.frame point=select nobs=n;
        do j=1 to count;
            if (obsnum(j)=select) then goto redo;
        end;
        position=select;
        count=count+1;
        obsnum(count)=select;
        output;
    end;
    stop;
run;
proc download data=work.sample out=work.sample;
run;
endrsubmit;
data sample;
    set sample;
    rand=uniform(-1);
    if state='WA' then wa=1;else wa=0;
run;
* Sort into random order but leave WA till last;
proc sort data=sample out=sample1;
by rand;
where state^='WA';
run ;
proc sort data=sample out=sample2;
by rand;
where state='WA';
run ;
data sample3;
    set sample1 sample2;
run;
data sample3;
    length id 8;
    set sample3;
    format fullname $20. ;
    drop longitude latitude;
    * postcodes as numeric;
    format pc z4.;
    pc=input(trim(left(postcode)),4.);
    drop postcode;
    rename pc=postcode;
    * time differences;
    if state='SA' then timediff=-0.5;
    if state='QLD' then timediff=0;
    if state='NT' then timediff=-0.5;
    if state='TAS' then timediff=0;
    if state='VIC' then timediff=0;
    if state='WA' then timediff=-2;
    if state='NSW' then timediff=0;
    if state='ACT' then timediff=0;
    if (index(suburb,'BROKEN HILL')>0) then timediff=-0.5;
    if (index(suburb,'LINDEMAN')>0)&(STATE='QLD') then timediff=0;
    if (index(suburb,'LORD HOWE')>0)&(STATE='QLD') then timediff=+0.5;
    * make address;
    length address $70;
    address=trim(strtxtra) || ' ' ||
        trim(strtnumb) || ' ' ||
        trim(strtname) || ' ' ||
        trim(suburb) || ' ' ||
        trim(state) || ' ' ||
        trim(postcode);
ID=_N_+10000;

```

```
eligible=0;
selected=1;
start='09oct2000:09:00:00'dt;
format start datetime.;
informat start datetime.;
arm=1;
cbtm1='11:00:00't;
drop position rand;
run;
libname dc 'n:\data\sascati\projects\ivr\ivr7\hybcati';
/*
data dc.confid(compress=yes);
    set sample3;
    run;
*/
signoff apc127;
```

### C.3 Letters to participants

The letters sent to households assigned to the CATI, Hybrid, and IVR methods are shown in Figures C.1, C.2, and C.4, respectively. The instruction sheets for the households assigned to the Hybrid and IVR methods are shown in Figures C.3 and C.5, respectively.



## Hunter Centre for Health Advancement

**Locked Bag No 10**  
**WALLSEND NSW 2287**  
 Phone: (02) 49246 326  
 Fax: (02) 49246 212

The Household  
 Street address  
 Suburb state postcode

Dear Household Members

The Hunter Centre for Health Advancement is doing a survey of technology, alcohol and drug use. We will be ringing households chosen at random and asking one person some simple questions. Since your household was one of those chosen we will ring you within two weeks of getting this letter. We would still talk to you even if you think this survey does not apply to your household since we have to ask the same questions of everyone. The whole interview will only take around 10 minutes to do.

**Who are we?** This project is part of the doctoral studies being undertaken by Mr Ross Corkrey, under the supervision of Dr Lynne Hancock. The Hunter Centre for Health Advancement is a unit of the University of Newcastle and the Hunter Area Health Service.

**How was your household chosen?** It is important that we get the views of as wide a range of people as possible. Your household address was chosen at random from the Telstra White Pages telephone listings.

**Do you have to talk to us?** While your views are important to us, participation in this study is voluntary, and you do not have to answer our questions if you do not want to.

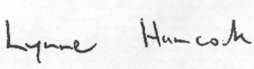
**What is the survey about?** The survey will ask questions about technology, alcohol and drug use. We are testing different ways of doing telephone surveys of health matters.

**What will happen to the information you give us?** The results of the survey will be used in health planning and research. All the answers to our questions will be treated in **STRICT CONFIDENCE**. Your answers will be entered onto a computer and stored securely at the Hunter Centre for Health Advancement. We will not record anyone's name. Your phone number and address will not be located near your other information. Once the survey is complete everyone's addresses and telephone numbers will be destroyed.

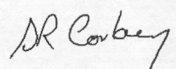
**What happens now?** An interviewer from the Centre for Health Advancement will ring you within the next two weeks. To make sure we interview someone in your household at random the interviewer will ask for the person in your household who is 18 or over and who had the last birthday.

**What if you need more information?** If you have any questions about this study before then, please do not hesitate to contact Ross Corkrey on (02) 4924 6381. If you want to check this is a genuine survey please ring the Projects Info Line at the Hunter Centre for Health Advancement on (02) 4924 6166. Thank you for your time and consideration.

Yours sincerely  
 HUNTER CENTRE FOR HEALTH ADVANCEMENT




Dr Lynne Hancock  
Principal Investigator




Mr Ross Corkrey  
Project Manager

If you have any complaint concerning the manner in which a research project is conducted, it may be given to the researcher, or if an independent person is preferred, to the Professional Officer, HAREC, C/- HAHS, Locked Bag 1, New Lambton Heights 2305, telephone (02) 4921 4950. Alternatively, you can contact The Human Research Ethics Officer, Research Branch, The Chancellery, University of Newcastle, University Drive, Callaghan NSW 2308, telephone (02) 4921 6333.




HUNTER HEALTH

A commitment to excellence in health promotion by Hunter Health and  
the Faculty of Medicine and Health Sciences University of Newcastle



The UNIVERSITY  
of NEWCASTLE

Figure C.1: Letter sent to households assigned to the CATI method.

[10080] 

**Hunter Centre for Health Advancement**  
 Locked Bag No 10  
 WALLSEND NSW 2287  
 Phone: (02) 4924 6499

The Household  
 Street address  
 Suburb state  
 postcode

Dear Household Members

The Hunter Centre for Health Advancement is doing a survey of technology, alcohol and drug use. We will be ringing households chosen at random and asking one person some simple questions. Since your household was one of those chosen we will ring you within weeks of getting this letter. We would still talk to you even if you think this survey does not apply to your household since we have to ask the same questions of everyone. The whole interview will only take around 10 minutes to do.

**Who are we?** This project is part of the doctoral studies being undertaken by Mr Ross Corkrey, under the supervision of Dr Lynne Hancock. The Hunter Centre for Health Advancement is a unit of the University of Newcastle and the Hunter Area Health Service.

**How was your household chosen?** It is important that we get the views of as wide a range of people as possible. Your household address was chosen at random from the Telstra White Pages telephone listings.

**Do you have to talk to us?** While your views are important to us, participation in this study is voluntary, and you do not have to answer our questions if you do not want to.

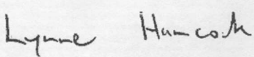
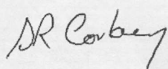
**What is the survey about?** The survey will ask questions about technology, alcohol and drug use. We are testing different ways of doing telephone surveys of health matters.

**What will happen to the information you give us?** The results of the survey will be used in health planning and research. All the answers to our questions will be treated in STRICT CONFIDENCE. Your answers will be entered onto a computer and stored securely at the Hunter Centre for Health Advancement. We will not record anyone's name. Your phone number and address will not be located near your other information. Once the survey is complete everyone's addresses and telephone numbers will be destroyed.

**What happens now?** An interviewer from the Centre for Health Advancement will ring you within the next two weeks. To make sure we interview someone in your household at random the interviewer will ask for the person in your household who is 18 or over and who had the last birthday. Some questions in the survey will be asked using a recorded voice. The interviewer will explain how this works.

**What if you need more information?** If you have any questions about this study before then, please do not hesitate to contact Ross Corkrey on (02) 4924 6381. If you want to check this is a genuine survey please ring the Projects Info Line at the Hunter Centre for Health Advancement on (02) 4924 6166. Thank you for your time and consideration.

Yours sincerely

HUNTER CENTRE FOR HEALTH ADVANCEMENT  
 Dr Lynne Hancock  
 Principal Investigator

Mr Ross Corkrey  
 Project Manager

If you have any complaint concerning the manner in which a research project is conducted, it may be given to the researcher, or if an independent person is preferred, to the Professional Officer, HAREC, C/- HAHS, Locked Bag 1, New Lambton Heights 2305, telephone (02) 4921 4950. Alternatively, you can contact The Human Research Ethics Officer, Research Branch, The Chancellor, University of Newcastle, University Drive, Callaghan NSW 2308, telephone (02) 4921 6333.



 *A commitment to excellence in health promotion by Hunter Health and the Faculty of Medicine and Health Sciences University of Newcastle*  **The UNIVERSITY of NEWCASTLE**

Figure C.2: Letter sent to households assigned to the Hybrid methods.

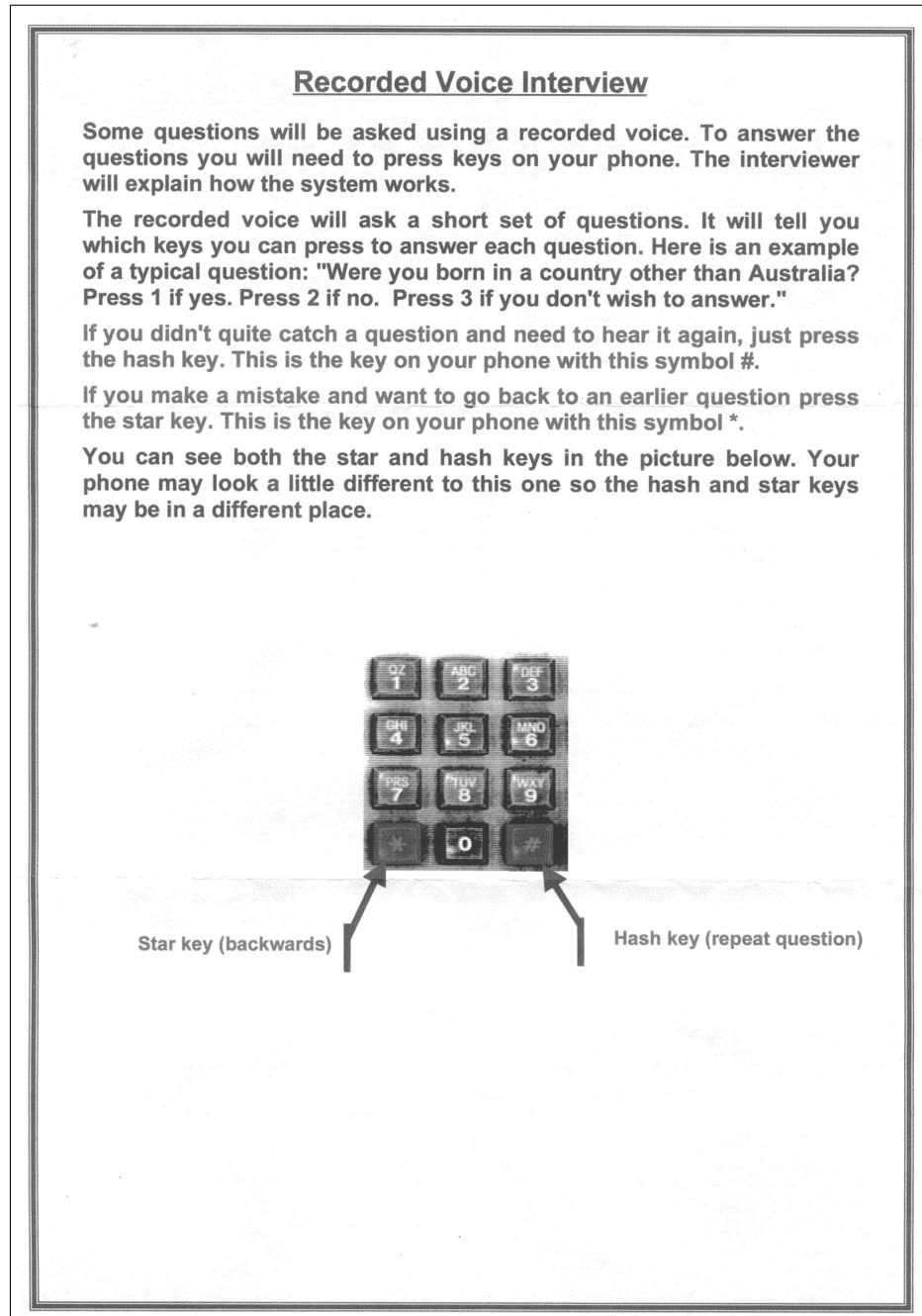



Figure C.3: Instruction sheet included with letter sent to households assigned to the Hybrid methods.



## Hunter Centre for Health Advancement

**Locked Bag No 10**  
**WALLSEND NSW 2287**  
 Phone: (02) 49246 326  
 Fax: (02) 49246 212

**The Household**  
**Street address**  
**Suburb state**  
**postcode**

Dear Household Members

The Hunter Centre for Health Advancement is doing a survey of technology, alcohol and drug use. We will be ringing households chosen at random and asking one person some simple questions. Since your household was one of those chosen we will ring you within two weeks of getting this letter. We would still talk to you even if you think this survey does not apply to your household since we have to ask the same questions of everyone. The whole interview will be done using a recorded voice and only take around 10 minutes to do.

**Who are we?** This project is part of the doctoral studies being undertaken by Mr Ross Corkrey, under the supervision of Dr Lynne Hancock. The Hunter Centre for Health Advancement is a unit of the University of Newcastle and the Hunter Area Health Service.

**How was your household chosen?** It is important that we get the views of as wide a range of people as possible. Your household address was chosen at random from the Telstra White Pages telephone listings.

**Do you have to talk to us?** While your views are important to us, participation in this study is voluntary, and you do not have to answer our questions if you do not want to.

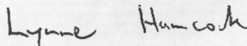
**What is the survey about?** The survey will ask questions about technology, alcohol and drug use. We are testing different ways of doing telephone surveys of health matters.

**What will happen to the information you give us?** The results of the survey will be used in health planning and research. All the answers to our questions will be treated in **STRICT CONFIDENCE**. Your answers will be entered onto a computer and stored securely at the Hunter Centre for Health Advancement. We will not record anyone's name. Your phone number and address will not be located near your other information. Once the survey is complete everyone's addresses and telephone numbers will be destroyed.

**What happens now?** To make sure we interview someone in your household at random we would only like to talk to the person in your household who is 18 or over and who had the last birthday. You will receive a phone call from the Centre for Health Advancement within the next two weeks. This will be a recorded voice interview. You will be able to answer questions by pressing keys on your phone. I've attached a sheet to help you with this part of the interview.

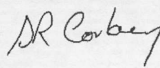
**What if you need more information?** If you have any questions about this study before then, please do not hesitate to contact Ross Corkrey on (02) 4924 6381. If you want to check this is a genuine survey please ring the Projects Info Line at the Hunter Centre for Health Advancement on (02) 4924 6166. Thank you for your time and consideration.

Yours sincerely  
**HUNTER CENTRE FOR HEALTH ADVANCEMENT**  
 Dr Lynne Hancock




Principal Investigator

Mr Ross Corkrey




Project Manager

If you have any complaint concerning the manner in which a research project is conducted, it may be given to the researcher, or if an independent person is preferred, to the Professional Officer, HAREC, C/- HAHS, Locked Bag 1, New Lambton Heights 2305, telephone (02) 4921 4950. Alternatively, you can contact The Human Research Ethics Officer, Research Branch, The Chancellery, University of Newcastle, University Drive, Callaghan NSW 2308, telephone (02) 4921 6333.



HUNTER HEALTH

A commitment to excellence in health promotion by Hunter Health and  
 the Faculty of Medicine and Health Sciences University of Newcastle



The UNIVERSITY  
 of NEWCASTLE

Figure C.4: Letter sent to households assigned to the IVR method.

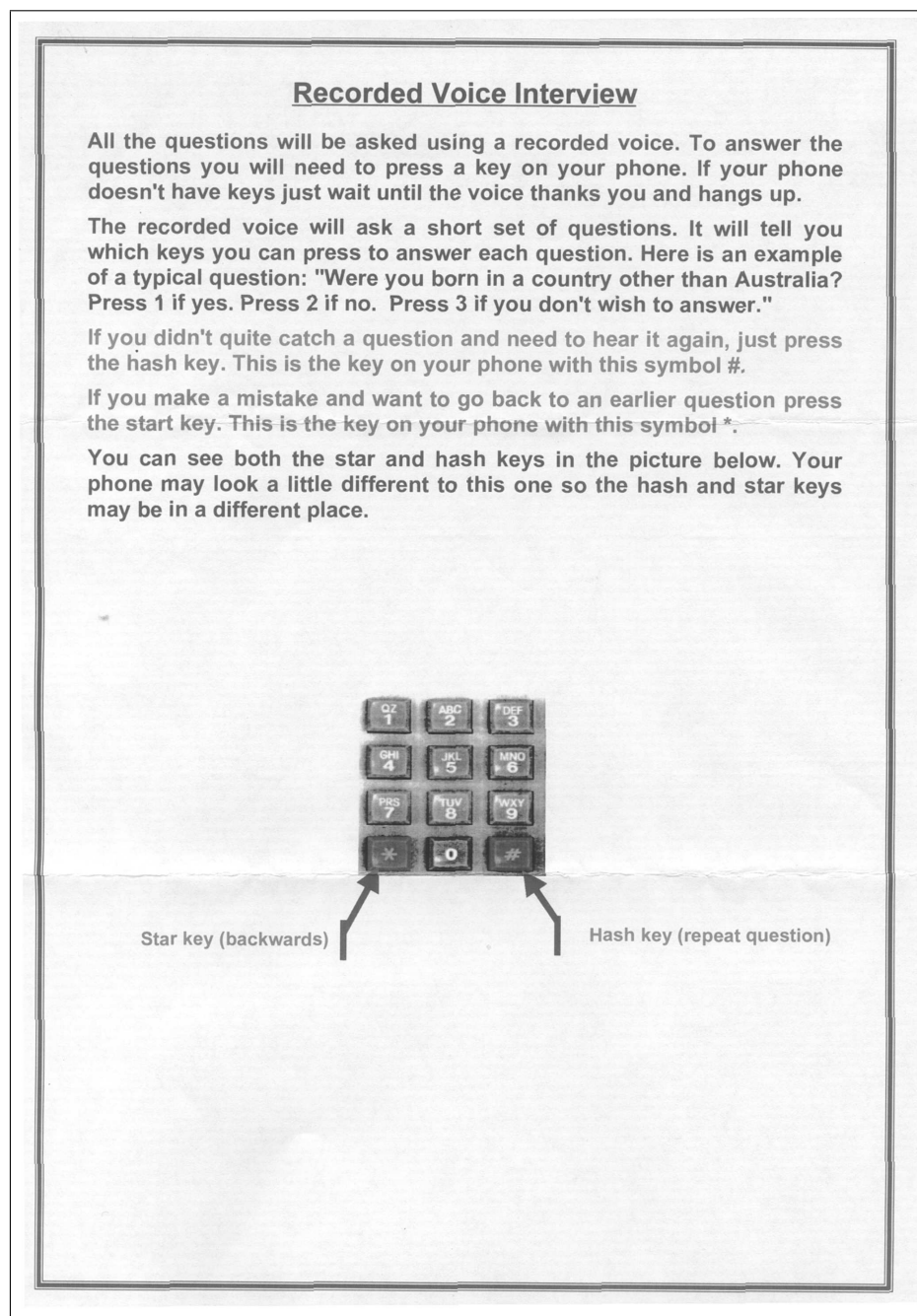


Figure C.5: Instruction sheet included with letter sent to households assigned to the IVR method.

## Appendix D

# Appendix for Chapter 6

### D.1 Cessnock pilot sampling programs

#### D.1.1 IVR phase sampling program

```
* sample program for IVR6 CSP;
* Ross Corkrey;
* Sample from Cessnock using SRS;

libname phone 'd:\phone';/* NSW white pages */
libname aria 'n:\data\statsgrp\data\ariarura';
options compress=yes;
options comamid=tcp remote=apc112;

options compress=yes;
libname rphone 'd:\phone';
* remove cases with missing addresses, telephone nos, or zero street numbers;
data frame(compress=yes);
  set rphone.hunter;
  where (
    index(SUBURB, 'BELLBIRD') or
    index(SUBURB, 'CESSNOCK') or
    index(SUBURB, 'ABERDARE') or
    index(SUBURB, 'KEARSLEY') or
    index(SUBURB, 'KURRI KURRI') or
    index(SUBURB, 'WESTON') or
    index(SUBURB, 'PELAW MAIN') or
    index(SUBURB, 'STANFORD MERTHYR') or
    index(SUBURB, 'HEDDON GRETA') or
    index(SUBURB, 'ABERMAIN') or
    index(SUBURB, 'NEATH') or
    index(SUBURB, 'BRANXTON') or
    index(SUBURB, 'NORTH ROTHBURY') or
    index(SUBURB, 'GRETA') or
    index(SUBURB, 'BLACK HILL') or
    index(SUBURB, 'BUCHANAN') or
    index(SUBURB, 'MULBRING') or
    index(SUBURB, 'KITCHENER') or
    index(SUBURB, 'ELLALONG') or
    index(SUBURB, 'MILLFIELD') or
    index(SUBURB, 'RICHMOND VALE') or
    index(SUBURB, 'TOMALPIN') or
    index(SUBURB, 'BIG YANGO') or
    index(SUBURB, 'PAXTON') or
```



```

        index(SUBURB, 'WOLLOMBI') or
        index(SUBURB, 'LAGUNA') or
        index(SUBURB, 'CEDAR CREEK') or
        index(SUBURB, 'KEINBAH') or
        index(SUBURB, 'SAWYERS GULLY') or
        index(SUBURB, 'POKOLBIN') or
        index(SUBURB, 'ROTHBURY') or
        index(SUBURB, 'LOXFORD')
    ) &
    ( state='NSW' )&
    (stdphone=:'(02) 49');

run;
proc freq data=frame;
table suburb;
run;
* get no of obs in frame;
data _null_;
dsid=open("work.frame");
nobs=attrn(dsid,"nobs");
call symput('m',nobs);
rc=close(dsid);
run;
data _null_;
put "number of cases in frame=" "&m";
run;
* required number of obs;
%let n=5000;
* extract sample;
data sample (drop=i j count) ;
    count=0;
    array obsnum (&m) _temporary_;
    do i=1 to &n;
        redo:
        select=ceil(uniform(33434219)*&m);
        set work.frame point=select nobs=n;
        do j=1 to count;
            if (obsnum(j)=select) then goto redo;
        end;
        position=select;
        count=count+1;
        obsnum(count)=select;
        output;
    end;
    stop;
run;
data sample;
    set sample;
    rand=uniform(-1);
run;
proc sort data=sample out=sample(drop=rand) ;
by rand;
run ;
data sample2;
    length id selected eligible 8;
    set sample;
    format fullname $20. address $40.;
* All are eligible;
    eligible=0;
    selected=1;
* Make ID variable;
    ID=_N_+5000;
    rename area_code=areacode;

```

```

drop longitude latitude;
  postcode=trim(left(postcode));
mailto='LBates@doh.health.nsw.gov.au';
copyto= ' ' ;
attach= ' ' ;
subject= ' ' ;
  rndintro=mod(_n_,5)+1;
run;
* create confid datasets;
/*libname dc1 'n:\data\sascati\projects\csp\';
data dc1.confid;
  set sample2;
run;
*/

```

### D.1.2 CATI phase sampling program

```

* sample program for IVR6 CSP FUP;
* Ross Corkrey;
* Sample from Cessnock IVR6 using SRS;

libname cd 'n:\data\statsgrp\data\cati_bu\csp\final';
libname lib 'n:\data\statsgrp\data\cati_bu\csp\final';
libname fmt 'n:\data\sascati\projects\csp';
libname core 'n:\data\sascati\bin';
option fmtsearch=(work fmt lib core);

data frame;
  merge cd.control(keep=id status cmpdate attdate selected eligible)
        cd.answers(keep=id bus age1 age2 age3)
        cd.confid(keep=id STDPHONE ID SELECTED ELIGIBLE ADDRESS STRTNUM INITIALS
                  FULLNAME POSTCODE STATE STRTXTRA STRTNAME STRTNUMB
                  SUBURB SURNAME TITLE AREACODE RNDINTRO START);

  by id;
  drop bus;
  if selected=1 & eligible=0 & status in ('CQ','OS','DR') & bus in (.,1);
run;

data cq;
  set frame;
  if status='CQ';
run;
data osdr;
  set frame;
  if status in ('OS','DR');
run;

%macro sample(dsin,dsout,n,seed);

* get no of obs in frame;
data _null_;
  dsid=open("&dsin");
  nobs=attrn(dsid,"nobs");
  call symput('m',nobs);
  rc=close(dsid);
run;
data _null_;
  put "number of cases in frame=" "&m";
run;

* extract sample;
data &dsout (drop=i j count) ;

```

```

count=0;
array obsnum (&m) _temporary_;
do i=1 to &n;
  redo:
  select=ceil(uniform(&seed)*&m);
  set &dsin point=select nobs=n;
  do j=1 to count;
    if (obsnum(j)=select) then goto redo;
  end;
  position=select;
  count=count+1;
  obsnum(count)=select;
  output;
end;
stop;
run;

%mend;

* Keep all CQs;
* Sample DR and OS;
%sample(dsin=work.osdr,dsout=work.osdrs,n=500,seed=78236529);

proc sort data=work.cq nodupkey;
  by id;
run;
proc sort data=work.osdrs nodupkey;
  by id;
run;
data sample;
  merge work.cq work.osdrs;
  by id;
  stdphone='(02) '||trim(left(substr(stdphone,2)));
* All are eligible;
  eligible=0;
  selected=1;
  format cmpdate weekdate17.;
  format attdate weekdate17.;
  cbtm1='14:00't;
run;

* create confid datasets;
/*libname dc1 'n:\data\sascati\projects\csp\fup';
data dc1.confid(label='Sample of cases from CSP');
  set sample;

run;
*/

```


## D.2 Cessnock pilot information letter

The letter sent to households before the IVR call is shown in Figure D.1 and the accompanying instruction sheet is shown in Figure D.2.

## D.3 Cessnock pilot script technical details

Towards the end of the IVR script, the CHCE item was used to ask the woman if she wanted someone to ring her back. She was also asked if she wanted

[5005]



## Hunter Centre for Health Advancement

Locked Bag No 10  
WALLSEND NSW 2287  
Phone: (02) 49246499  
Fax: (02) 49246212

The household  
Street number and name  
CESSNOCK NSW 2325

Dear Household Members

The Hunter Cervical Screening Program at the Hunter Centre for Health Advancement is launching a community awareness campaign about Pap tests and cervical cancer in the Cessnock area. During this campaign, 5000 households in the Cessnock area will be randomly contacted by telephone using a Recorded Voice message to prompt women between 18-70 years to have Pap tests.

**Who are we?** The project is part of the doctoral studies being undertaken by Mr Ross Corkrey and the master studies being undertaken by Dr. Aye T. Htun under the supervision of Dr. Lynne Hancock.

**How was your household chosen?** It is important that we get the views of a wide range of women (18-70 years) in the Cessnock area as possible. Your household address was chosen at random from the Telstra White Pages telephone listings.

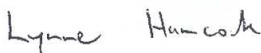
**Do you have to talk to us?** While your views are important to us, participation in this study is voluntary, and you do not have to answer our questions if you do not want to.


**What will happen to the information you give us?** The results of the survey will be used in health planning and research. All the answers to our questions will be treated in STRICT CONFIDENCE. Your answers will be entered onto a computer and stored securely at the Hunter Centre for Health Advancement. We will not record anyone's name. Your phone number and address will not be located near your information. Once the survey is complete everyone's addresses and telephone numbers will be destroyed.

**What happens now?** The Recorded Voice system from the Hunter Centre for Health Advancement will ring you within the next two weeks. The Recorded Voice system can automatically make calls and play high quality recordings of questions and messages. To ensure that we get information for all the different age groups within Cessnock area, the person we would like to talk to when we call is the woman aged between 18-70 years, who had the last birthday. This telephone prompt will take about 5 minutes.


**What if you need more information?** If you have any questions about this study before then, please do not hesitate to contact Ms Lucy Bates, Coordinator, Hunter Cervical Screening Program on 2924 6499. If you want to check if this is a genuine study please ring the Projects Info Line at the Hunter Centre for Health Advancement on (02) 4924 6166. Thank you for your time and consideration.

Yours sincerely  
HUNTER CENTRE FOR HEALTH ADVANCEMENT


  
Dr. Lynne Hancock  
Program Manager  
Women's Health

  
Ms Lucy Bates  
Area Coordinator  
Hunter Cervical Screening Program

If you have any complaint concerning the manner in which a research project is conducted, it may be given to the researcher, or if an independent person is preferred, to the Professional Officer, HAREC, C/- HAHS, Locked Bag 1, New Lambton Heights 2305, telephone (02) 4921 4950. Alternately, you can contact The Human Research Ethics Officer, Research Branch, The Chancellory, University of Newcastle, University Drive, Callaghan NSW 2308, telephone (02) 4921 6333.



*A commitment to excellence in health promotion by Hunter Health and the Faculty of Medicine and Health Sciences University of Newcastle*



The UNIVERSITY  
of NEWCASTLE

Figure D.1: Information letter sent to households in the Cessnock pilot.

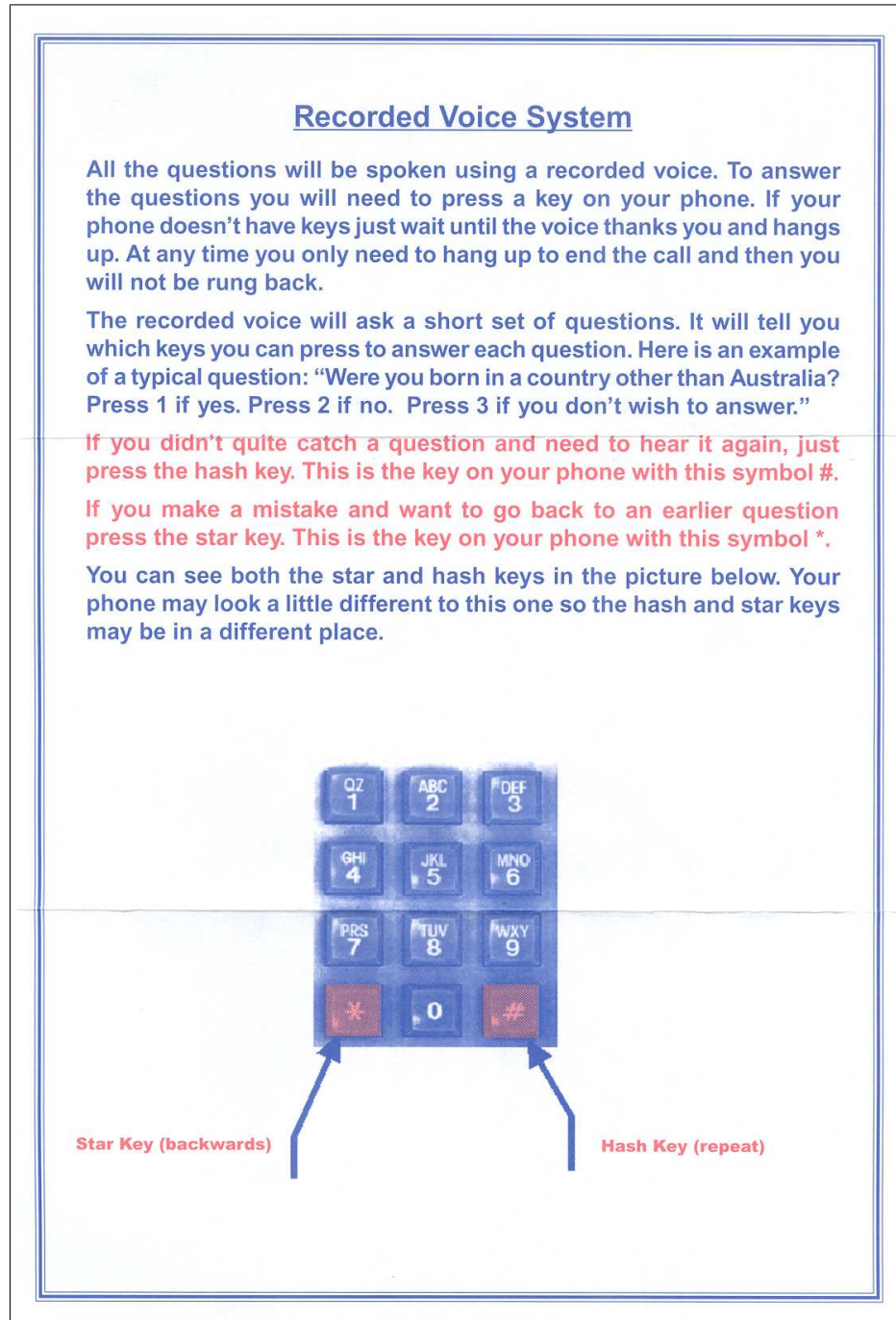


Figure D.2: Instruction sheet sent to households in the Cessnock pilot.

to record a question beforehand. This was done using the OPEN script item EMAIL3, which allowed a message of up to thirty seconds to be recorded. After the interview terminated, GEIS then generated an e-mail using script items EMAIL4 and EMAIL5. The e-mail was sent to an expert in cervical screening. If a question had been recorded, this was included with the e-mail as a WAV file attachment. The expert was a local staff member responsible for cervical screening promotion in the Hunter region.

In the follow-up script, women who had not listened to the IVR message were offered a second opportunity to listen to the message. This required sending information from the CATI interview system to the IVR system. It was done using external updating as described in Appendix Section A.5.3. The items IVR\_4 and IVR\_5 in the script (Section 6.1) externally updated the IVR CONTROL data set with the date and time for GEIS to call back. If the woman reported that another telephone number was more convenient then this was also updated. The script items IVR\_7, IVR\_8, IVR\_9, IVR\_10, and IVR\_11 updated the variables STDPHONE, SELECTED, ELIGIBLE, STATUS, and IVRLINE, respectively. The CALC item IVR\_8 updated the STATUS variable to code 'CB', which had the effect of causing GEIS to call back at the preset date and time.

## Appendix E

# Appendix for Chapter 7

### E.1 Sampling program for the IVR message

```
* Sample selection for Hunter IVR CSP;
* Read in hunter female postcode populations
  to get postcodes only;
* these are ABS Census 1996 - from library;
option nocenter nofmterr;
option validvarname=v6;
PROC IMPORT OUT= WORK.FEMHUNT2 DATAFILE=
"f:\officialdata\Hunter CLIB96\postcodes\female_hunter_postcodes.xls"
  DBMS=EXCEL2000 REPLACE;
  GETNAMES=YES;
RUN;
data femhunt2;
  set femhunt2;
  if agegroup in ('age0_4','age5_9','age10_14','age15_19',
    'age70_74','age75_79','age80_84',
    'age85_89','age90_94','age95_98',
    'age99+', 'overseas', 'total') then delete;

run;
* List of eligible postcodes;
proc transpose data=femhunt2 out=femhunt2;
  var PC2259 PC2282 PC2281 PC2280 PC2267 PC2265 PC2264
    PC2337 PC2336 PC2335 PC2334 PC2333 PC2330
    PC2329 PC2328 PC2327 PC2326 PC2325 PC2324 PC2323
    PC2322 PC2321 PC2320 PC2315 PC2314 PC2311 PC2308
    PC2307 PC2306 PC2305 PC2304 PC2303 PC2302 PC2300
    PC2299 PC2298 PC2297 PC2296 PC2295 PC2294 PC2293
    PC2292 PC2291 PC2290 PC2289 PC2287 PC2286 PC2285
    PC2284 PC2283 PC2421 PC2420 PC2343 PC2340 PC2339
    PC2338 PC2850 PC2787 PC2423 PC2422;
  by agegroup;
run;
data femhunt2;
  set femhunt2;
  _name_=tranwrd(_name_,'PC','');
  pc=input(_name_,5.0);
  drop _name_ _label_;
  rename col1=NumWomen;
run;
proc means data=femhunt2 nway sum;
  class pc;
  var numwomen;
```

```

        output out=femhunt3(drop=_freq_ _type_) sum(numwomen)=;
    run;
    * calc SE to postcode level;
    libname seifa 'f:\officialdata\seifa';
    proc means data=seifa.aust nway noprint;
        class pc;
        weight pop;
        output out=seifa mean(se)= mean(ier)= mean(ieo)=;
    run;
    * merge in SEIFA to get socioeconomic data;
    proc sort in=seifa;
        by pc;
    run;
    proc sort in=femhunt3 ;
        by pc;
    run;
    data seifa;
        merge seifa(in=a) femhunt3(in=b);
        by pc;
        if b;
        drop _type_ _freq_;
    run;
    data seifa;
        set seifa;
        format pc 4.;
        length pc2 $4;
        pc2=put(pc, 4.);
    run;
    /*
    remove postcodes that are in Cessnock;
    Remove Newcastle Uni;
    Remove Newcastle West due to small pop size;
    Remove Williamtown airforce base;
    */
    data exclude;
        length r $2;
        input pc r;
        datalines;
2320 c
2321 c
2322 c
2323 c
2324 c
2325 c
2326 c
2327 c
2330 c
2334 c
2335 c
2308 nu
2302 nw
2314 w
;
    run;
    proc sort data=exclude;by pc;run;
    data seifa;
        merge seifa(in=A) exclude(in=B);
        by pc;
        if B then sv='Cessnock etc';
            else sv='In survey  ';
    run;
    *Hunter General Practice study postcodes - from Lucy Bates;

```



```

data gp;
  length pc2 $4;
  input pc2;
  cards;
2265
2280
2281
2282
2283
2284
2285
2287
2289
2291
2300
2302
2304
2303
2316
2337
2422
2428
2427
2430
2330
2429
;
  run;
proc sort data=gp;
  by pc2;
  run;
* Remove MURRINDINDI postcodes - used by Womens Health program;
data wh;
  length pc2 $4;
  input pc2;
  cards;
2337
2338
2339
2343
;
run;
data seifa;
  length gp $15 sv $15 wh $30;
  merge gp(in=a) seifa(in=b) wh(in=c);
  by pc2;
  if a then gp='In GP';else gp='Not in GP';
  if c then wh='In Womens Health';
  else wh='Not in Womens Health';
  run;
data seifa;
  set seifa;
  if WH='In Womens Health' and GP='Not in GP' and r=' ';
  run;
* principal components;
proc princomp data=seifa cov out=Prin standard;
  var se numwomen;
run;
proc sort data=prin;
  by prin1 prin2;
  run;
* break into control and intervention sets;

```

```

data prin2b prin2c;
  length trt $4;
  set prin;
  if mod(_n_,2) then do;
    k1+1;
  k=k1;
  trt='INTV';
    output prin2b;end;
  else do;
    k2+1;
  k=k2;
  trt='CTRL';
    output prin2c;
  end;
  label k='Group';
  label trt='Treatment';
run;
*re-integrate;
data seifa;
  set prin2b(in=a) prin2c(in=b);
  by k;
run;
proc sort in=seifa out=postcodes;
  by pc;
run;
* get SLA codes;
proc sort data=regloc.locality out=loc;
  by post1;
run;
data loc2;
  merge postcodes(in=a) loc(rename=post1=pc);
  by pc;
  if a;
run;
proc sort data=loc2 nodupkey;
  by locname;
run;
* merge postcodes with telephone data;
data postcodes2;
  set postcodes;
  length postcode $4.;
  postcode=put(pc,4.);
  keep postcode trt ;
run;
proc sort data=postcodes2;
  by postcode;
run;
option validvarname=v7;
data phone(compress=yes);
/* takes 3-4 minutes - skip 2nd time */
  set phone.aust;
  where state='NSW';
run;
proc sort data=phone tagsort;
/* takes 5-6 minutes - skip 2nd time*/
  by postcode;
run;
data frame;
/* takes 1 minute */
  merge postcodes2(in=a) phone;
  by postcode;
  if a;

```

```

run;
* Remove addresses that lay in cessnock;
data frame;
  set frame;
  if (
    index(SUBURB, 'ABERDARE') or
    index(SUBURB, 'ABERMAIN') or
    index(SUBURB, 'BELLBIRD') or
    index(SUBURB, 'BIG YANGO') or
    index(SUBURB, 'BLACK HILL') or
    index(SUBURB, 'BRANXTON') or
    index(SUBURB, 'BUCHANAN') or
    index(SUBURB, 'CEDAR CREEK') or
    index(SUBURB, 'CESSNOCK') or
    index(SUBURB, 'ELLALONG') or
    index(SUBURB, 'GRETA') or
    index(SUBURB, 'HEDDON GRETA') or
    index(SUBURB, 'KEARSLEY') or
    index(SUBURB, 'KEINBAH') or
    index(SUBURB, 'KITCHENER') or
    index(SUBURB, 'KURRI KURRI') or
    index(SUBURB, 'LAGUNA') or
    index(SUBURB, 'LOXFORD') or
    index(SUBURB, 'MILLFIELD') or
    index(SUBURB, 'MULBRING') or
    index(SUBURB, 'NEATH') or
    index(SUBURB, 'NORTH ROTHBURY') or
    index(SUBURB, 'PAXTON') or
    index(SUBURB, 'PELAW MAIN') or
    index(SUBURB, 'POKOLBIN') or
    index(SUBURB, 'RICHMOND VALE') or
    index(SUBURB, 'ROTHBURY') or
    index(SUBURB, 'SAWYERS GULLY') or
    index(SUBURB, 'STANFORD MERTHYR') or
    index(SUBURB, 'TOMALPIN') or
    index(SUBURB, 'WESTON') or
    index(SUBURB, 'WOLLOMBI')
  ) then delete;
run;
* Get NSW yellow pages ;
data bus ;
/* takes 1.5 minutes - skip 2nd time*/
  set phone.bus;
  where state='NSW';
  strtextra=upcase(street_extra);
  strtnumb=upcase(street_number);
  strtname=upcase(street_name);
  suburb=upcase(suburb);
  if street_name=' ' or street_number=' '
    or suburb=' ' then delete;
  Busid=_n_;
  drop street_name street_number street_extra;
run;
* sort by address;
proc sort data=bus tagsort;
  by suburb strtname strtnumb ;
/* takes 5-7 minutes - skip 2nd time*/
run;
* sort white pages by address;
data frame;
  set frame;
  strtextra=upcase(strtextra);

```

```

    strtnumb=upcase(strtnumb);
    strtname=upcase(strtname);
    suburb=upcase(suburb);
run;
proc sort data=frame tagsort; /* takes 1 minutes */
  by suburb strtname strtnumb ;
run;
* merge in yellow pages - drop any businesses;
data frame2;
  merge frame(in=a) bus(drop= postcode state in=b);
  by suburb strtname strtnumb ;
  if a ;
  id=_n_;
run;
data frame3;
  set frame2;
  if strtnumb=lag(strtnumb) and
    strtname=lag(strtname) and
    suburb= lag(suburb) then lag=1;else lag=0;
run;
proc sort data=frame3 tagsort; /* takes 1 minutes */
  by descending suburb descending strtname
    descending strtnumb descending lag;
run;
data frame4;
  set frame3;
  if strtnumb=lag(strtnumb) and
    strtname=lag(strtname) and
    suburb= lag(suburb) then lag=1;
run;
* Eliminate cases that have multiple entries and are listed in
  the yellow pages, excluding caravan parks and apartments;
data frame5;
  set frame4;
  if lag=1 and busid>. and strtnumb=' ' and strtname=' ' and
    anzsic_code_1='5710' and
    index(company_name,'Caravan Park')=0 and
    index(company_name,'Apartment')=0 then delete;
run;
* Exclude ineligible cases;
data frame6 ;
  set frame5;
  * exclude some titles;
  if index(upcase(title),'SISTER')>0 then delete;
  if index(upcase(title),'DRS ')>0 then delete;
  * exclude general practices;
  if anzsic_code_1='8621' then delete;
  * Exclude businesses;
  if stdphone=ac_phone then delete;
  * exclude companies;
  if
    index(upcase(company_name),'ASSISTED LIVING')>0 OR
    index(upcase(company_name),'CLUB')>0 OR
    index(upcase(company_name),'SHOP')>0 OR
    index(upcase(company_name),'P/L')>0 OR
    index(upcase(company_name),'POLICE')>0
  then delete;
  * exclude businesses and retirees;
  if
    index(upcase(strtxtra),'AERODROME')>0 OR
    index(upcase(strtxtra),'AMBULANCE')>0 OR
    index(upcase(strtxtra),'ARABIANS')>0 OR

```

```

index(uppercase(strtxtra),'ARMY')>0 OR
index(uppercase(strtxtra),'BANK')>0 OR
index(uppercase(strtxtra),'BARACKS')>0 OR
index(uppercase(strtxtra),'BARRACKS')>0 OR
index(uppercase(strtxtra),'BIBLE')>0 OR
index(uppercase(strtxtra),'CAMP')>0 OR
index(uppercase(strtxtra),'CENTER')>0 OR
index(uppercase(strtxtra),'CENTRE')>0 OR
index(uppercase(strtxtra),'CINEMA')>0 OR
index(uppercase(strtxtra),'CLINIC')>0 OR
index(uppercase(strtxtra),'CLUB')>0 OR
index(uppercase(strtxtra),'CNTR')>0 OR
index(uppercase(strtxtra),'COLLEGE')>0 OR
index(uppercase(strtxtra),'DAIRY')>0 OR
index(uppercase(strtxtra),'DENTURE')>0 OR
index(uppercase(strtxtra),'ESTATE')>0 OR
index(uppercase(strtxtra),'EST ')>0 OR
index(uppercase(strtxtra),'EXCHANGE')>0 OR
index(uppercase(strtxtra),'FARM')>0 OR
index(uppercase(strtxtra),'FLOOR')>0 OR
index(uppercase(strtxtra),'GARDEN CITY')>0 OR
index(uppercase(strtxtra),'GOLF')>0 OR
index(uppercase(strtxtra),'HOSPITAL')>0 OR
index(uppercase(strtxtra),'HOSPTL')>0 OR
index(uppercase(strtxtra),'HOSTEL')>0 OR
index(uppercase(strtxtra),'HOUSE')>0 OR
index(uppercase(strtxtra),'HSE ')>0 OR
index(uppercase(strtxtra),'JOINERY' )>0 OR
index(uppercase(strtxtra),'KENNEL')>0 OR
index(uppercase(strtxtra),'K-MART')>0 OR
index(uppercase(strtxtra),'LODGE')>0 OR
index(uppercase(strtxtra),'LVL ')>0 OR
index(uppercase(strtxtra),'LEVEL ')>0 OR
index(uppercase(strtxtra),'MALL')>0 OR
index(uppercase(strtxtra),'MARINA')>0 OR
index(uppercase(strtxtra),'MARKET')>0 OR
index(uppercase(strtxtra),'MASONIC')>0 OR
index(uppercase(strtxtra),'MEDCL')>0 OR
index(uppercase(strtxtra),'MEDICAL')>0 OR
index(uppercase(strtxtra),'H M A S')>0 OR
index(uppercase(strtxtra),'MINE')>0 OR
index(uppercase(strtxtra),'NRSNG')>0 OR
index(uppercase(strtxtra),'NURSING')>0 OR
index(uppercase(strtxtra),'PARISH')>0 OR
index(uppercase(strtxtra),'PIGGERIES')>0 OR
index(uppercase(strtxtra),'PLZ')>0 OR
index(uppercase(strtxtra),'PLAZA')>0 OR
index(uppercase(strtxtra),'POST OFFICE')>0 OR
index(uppercase(strtxtra),'PRAYER')>0 OR
index(uppercase(strtxtra),'RESORT')>0 OR
index(uppercase(strtxtra),'RETIREMENT')>0 OR
index(uppercase(strtxtra),'RETIREMNT')>0 OR
index(uppercase(strtxtra),'RTREMENT')>0 OR
index(uppercase(strtxtra),'RETIR ')>0 OR
index(uppercase(strtxtra),'RETIRE ')>0 OR
index(uppercase(strtxtra),'RETIREM')>0 OR
index(uppercase(strtxtra),'RETRMNT')>0 OR
index(uppercase(strtxtra),'RTMNT VILLAGE')>0 OR
index(uppercase(strtxtra),'RETIR ')>0 OR
index(uppercase(strtxtra),'SARGENTS MESS')>0 OR
index(uppercase(strtxtra),'SCHOOL')>0 OR
index(uppercase(strtxtra),'SERVICE')>0 OR

```

```

index(upcase(strtxtra),'SHOP')>0 OR
index(upcase(strtxtra),'SHOPNG')>0 OR
index(upcase(strtxtra),'SHP ')>0 OR
index(upcase(strtxtra),'SNGLE CHAN RADIO FR RNVL')>0 OR
index(upcase(strtxtra),'SPEC 13')>0 OR
index(upcase(strtxtra),'SPECIALIST')>0 OR
index(upcase(strtxtra),'ST ')>0 OR
index(upcase(strtxtra),'STUD ')>0 OR
index(upcase(strtxtra),'SUITE')>0 OR
index(upcase(strtxtra),'THOROUGHbred')>0 OR
index(upcase(strtxtra),'VETERINARY')>0 OR
index(upcase(strtxtra),'VINEYARD')>0 OR
index(upcase(strtxtra),'WARDENS RES-INTERNATIONAL HOUSE')>0 OR
index(upcase(strtxtra),'WARIALDA RAIL')>0 OR
index(upcase(strtxtra),'WHARF')>0
then delete;
* Exclude control postcodes;
if trt=0 then delete;
run;
* Only residential numbers;
data frame7 ;
set frame6;
if left(stdphone)='(02)' ;
run;
proc sort data=frame7;
by postcode;
run;
* Draw sample of size 20000 from frame ;
proc freq data=frame7 noprint;
table postcode/out=sampsize ;
run;
* get stratum sample size;
data sampsize2;
set sampsize ;
_nsize_=floor(20000*percent/100);
run;
* Draw stratified sample;
proc surveyselect data=frame7 method=srs
sampsiz=sampsize2
seed=1248702 out=sample;
strata postcode;
run;
* Create CONFID;
data sample2;
length id selected eligible 8 comment $200;
length stdphone $16;
length attach $200;
set sample;
drop trt latitude longitude
ac_phone anzsic_code_1--emp_size_desc
lag ;
rename fax_ac_phone=fax;
label fax_ac_phone='Fax number';
rename company_name=Co_name;
label company_name='Company name';
rename selectionprob=Sel_prob;
label selectionprob='Unit selection probability';
rename SamplingWeight=sampwght;
label SamplingWeight='Unit sampling weight';
format fullname $20. ;
* All are eligible;
eligible=0;

```

```


    selected=1;
* Make ID variable;
    ID=_N_+20000;
    * remove area code since all are 02;
    areacode=scan(stdphone,1,'()');
    substr(stdphone,1,length(areacode)+2)='';
    * insert leading zero;
    stdphone='0 '||trim(left(stdphone));
    format stdphone $16.;
    drop longitude latitude;
    postcode=trim(left(postcode));
    mailto='Lucy.Bates@hunter.health.nsw.gov.au';
    copyto='Ross.Corkrey@hunter.health.nsw.gov.au';
    attach='d:\ivr\ivr8\recloc\r'||trim(left(put(id,5.))||
        '_email3.wav';
    noattach='';
    subject='Question from '||title||' '||fullname;
    address=trim(strtxtra)||' '||
        trim(left(strtnumb))||' '||
        trim(left(strtname))||' '||
        trim(left(suburb));
run;
* deselect duplicate tel nos and addresses;
proc sort data=sample2;
by stdphone;
run;
data sample2;
set sample2;
if stdphone=lag(stdphone) then do;
    eligible=15;
    comment='Duplicated number';
end;
run;
proc sort data=sample2;
by strtxtra strtnumb strtname suburb;
run;
data sample2;
set sample2;
if strtxtra=lag(strtxtra) & strtnumb=lag(strtnumb) &
    strtname=lag(strtname) & suburb=lag(suburb)
    & strtxtra^=' ' & strtnumb^=' '
    & strtname^=' ' & suburb^=' '
    then do;
    eligible=16;
    if comment='' then comment='Duplicated address';
    else comment=trim(comment)||
        ' Duplicated address';
end;
run;
* Sort by postcode to minimise post costs;
proc sort data=sample2;
by postcode id;
run;
* create confid datasets;
libname datacati v6 'n:\data\sascati\projects\ivr\ivr8';
data datacati.oldconf(compress=yes label=Hunter IVR CSP);
length attach $200;
set sample2;
attach='d:\ivr\ivr8\recloc\r'||trim(left(put(id,5.))||
    '_email3.wav';
run;

```

## **E.2 Information letter**

The information letter sent to households is shown in Figure E.1 and the accompanying instruction sheet is shown in Figure E.2.





**Hunter Centre for Health Advancement**

30649

**Locked Bag No 10**  
**WALLSEND NSW 2287**  
 Phone: (02) 49246499  
 Fax: (02) 49246209

R Corkrey  
 Street number and name  
 SUBURB  
 NSW 2298

Dear Household Members

The Hunter Cervical Screening Program at the Hunter Centre for Health Advancement is launching a community awareness campaign about Pap tests and cervical cancer in the Hunter area. During this campaign, 20,000 households will be randomly contacted by telephone using a Recorded Voice message to prompt women between 18-70 years to have Pap tests. A Pap test is a simple test that can be done in a doctor's office, clinic, or hospital that can detect early changes in cells before they turn cancerous.

**Who are we?** The project is part of the doctoral studies being undertaken by Mr Ross Corkrey under the supervision of Dr Lynne Hancock.

**How was your household chosen?** It is important that we get the views of as wide a range of women (18-70 years) in the Hunter area as possible. Your household address was chosen at random from the Telstra White Pages telephone listings.


**Do you have to talk to us?** While your views are important to us, participation in this study is voluntary, and you do not have to answer our questions if you do not want to. You can simply listen to the message or hang up at any point. If you hang up you will not be rung back.


**What will happen to the information you give us?** The results of the project will be used in health planning and research. All the answers to our questions will be treated in STRICT CONFIDENCE. Your answers will be entered onto a computer and stored securely at the Hunter Centre for Health Advancement. We will not record anyone's name. Your phone number and address will not be located near your information. Once the project is complete everyone's addresses and telephone numbers will be destroyed.

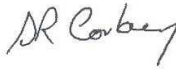
**What happens now?** The Recorded Voice System from the Hunter Centre for Health Advancement will ring you within the next two weeks. The Recorded Voice System can automatically make calls and play high quality recordings of questions and messages. To ensure that we get information for all the different age groups within Hunter area, the person we would like to talk to when we call is the woman aged between 18-70 years, who had the last birthday. This telephone call will take about 5 minutes.

**What if you need more information?** If you have any questions about this study before then, please do not hesitate to contact Ms Lucy Bates, Coordinator, Hunter Cervical Screening Program on (02) 4924 6499. If you want to check if this is a genuine study please ring the Projects Info Line at the Hunter Centre for Health Advancement on (02) 4924 6166. Thank you for your time and consideration.


Yours sincerely  
 HUNTER CENTRE FOR HEALTH ADVANCEMENT

  
 Dr Lynne Hancock  
 Program Manager  
 Women's Health

  
 Ms Lucy Bates  
 Area Coordinator  
 Hunter Cervical Screening Program

  
 Mr Ross Corkrey  
 PhD Student  
 University of Newcastle

If you have any complaint concerning the manner in which a research project is conducted, it may be given to the researcher, or if an independent person is preferred, to the Professional Officer, HAREC, C/- HAHS, Locked Bag 1, New Lambton Heights 2305, telephone (02) 4921 4950. Alternately, you can contact The Human Research Ethics Officer, Research Branch, The Chancellery, University of Newcastle, University Drive, Callaghan NSW 2308, telephone (02) 4921 6333.

  
 HUNTER HEALTH

A commitment to excellence in health promotion by Hunter Health and  
 the Faculty of Medicine and Health Sciences University of Newcastle


  
 The UNIVERSITY  
 of NEWCASTLE  
 AUSTRALIA

Figure E.1: Information letter sent to households.

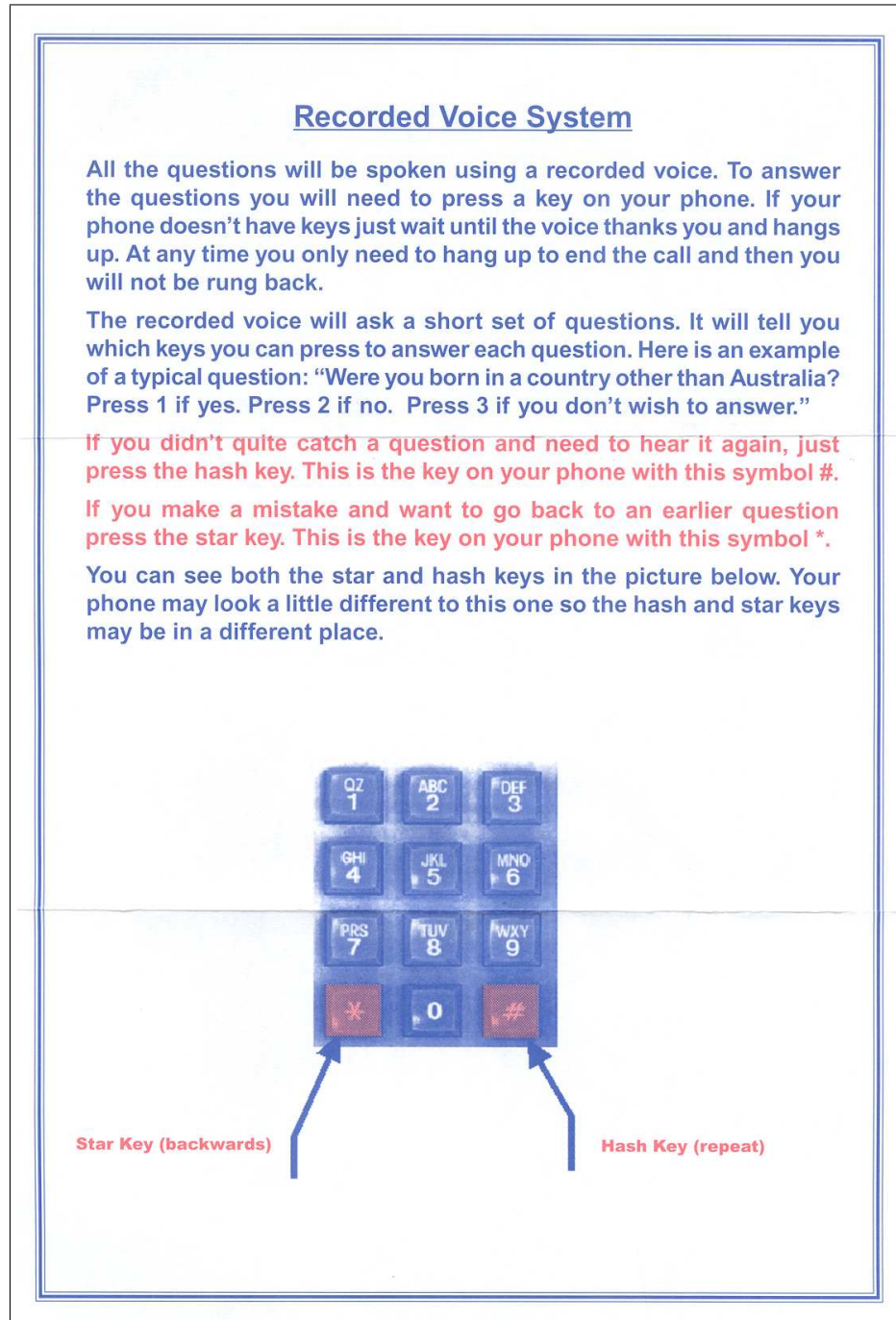


Figure E.2: Instruction sheet sent to households.

### E.3 Data request to the Health Insurance Commission

#### Request for Pap test data summary.

##### Contacts

Dr Lynne Hancock, (02) 4924 6365, Lynne.Hancock@hunter.health.nsw.gov.au  
Mr Ross Corkrey, (02) 4924 6381, Ross.Corkrey@hunter.health.nsw.gov.au

##### Postal address:

Hunter Centre for Health Advancement  
Locked bag 10  
WALLSEND NSW 2287

##### Rationale

We recently conducted a campaign aimed at increasing Pap test rates in the Hunter area. This consisted of an innovative Interactive Voice Response (IVR) computerised system that rang 20,000 households in selected postcodes in the Hunter area and played high quality voice recordings over the telephone. Women responded to the system by pressing keys of their touchphones. It provided tailored advice to women on the importance of having regular Pap tests and encouraged underscreened women to speak to their health care provider about a Pap test. In order to assess the efficacy of the IVR intervention, the system only rang households in randomly selected postcodes. The Pap test rates within the selected postcodes will be compared to rates in non-selected postcodes before and after the IVR intervention. In order to obtain sufficient statistical power it is essential that the Pap test rates be obtained at the postcode level. The project forms part of Mr Ross Corkrey's PhD thesis work and the supervisor for the project is Dr Lynne Hancock.

##### Data coverage

The data should be counts of patients with the following restrictions:

#### 1. MBS item numbers.

The following codes should be included:

73053 [Cytology of smears from cervix: (a) for detection of pre-cancerous or cancerous changes in women with no symptoms, signs or recent history suggestive of cervical neoplasia; or (b) due to an unsatisfactory smear taken in the circumstances defined in para (a) above; or (c) if there is inadequate information provided to use item 73055]

73055 [Cytology not associated with item 73053, of smears from cervix in association with: (a) the management of previously detected abnormalities including precancerous or cancerous conditions; or (b) the investigation of women with symptoms, signs or recent history suggestive of cervical neoplasia; each test]

73057 [Cytology of smears from vagina, not associated with item 73053 or 73055 nor to monitor hormone replacement therapy - each test]

#### 2. Time period.

Only services carried out in the following period should be included:

Starting: 8 October 2000

Ending: 6 January 2002

### 3. Area.

Only services provided for patients who reside in the postcodes:

2259, 2264, 2267, 2286, 2290, 2292, 2293, 2295, 2296, 2297, 2298, 2299, 2305, 2307, 2311, 2315, 2328, 2329, 2333, 2336, 2340, 2420, 2421, 2423, 2787, 2850.

### 4. Reports

We would like a final report be generated as soon possible after the end of the data period, preferably by late January 2002. The report should include details of the data extraction methodology used and any adjustments needed to ensure confidentiality of the patients. The report should consist of counts of services by patient's residential postcode.

If privacy issues arise where it may be possible to identify individuals, counts should be combined for consecutive months, not postcodes. This is because combining postcodes would invalidate our data analysis.

#### Data format:

The data should be produced as a comma-separated file. The first line of the file should include variable names, or a separate key should be provided.

The file should contain the following variables:

Suggested variable name	Variable meaning
Date	Date of service
AgeGrp	Patient's age group
RPost	Patient's residential postcode
PrevServ	Previous similar service
Count	Count of patients

Notes:

1. **Date** should be the date the Pap test was carried out. The format may be "monyyyy" or similar. For example "FEB2001".

2. **AgeGrp** should be grouped in to the groups 20-49 and 50-69. Other ages can be ignored.

3. **PrevServ** should contain a code indicating whether the patients had a similar service less than one year previous, between one and up to two years previous, or more than two years previous, has not had a previous similar service, or the date of a previous service is unknown. A similar service is any of the item codes 73053, 73055, 73057.

It is suggested that the coding system to be used is: 0='<1year', 1='1-2 years', 2='>2 years', 3='None previous', 4='Unknown'.

Where a woman has had Pap tests in more than one time period (<1 year, 1-2 years, >2 years), she should be counted in the most recent period in which she last screened.

We are aware that data are only available for the past 5 years.

4. **Count** should be the total of women who had a service classified by the above variables. It should *not* be the number of procedures.

**Example of the data format:**

```
DATE, AGEGRP, RPOST, PREVSERV, COUNT
APR2000, 20_49, 2259, 0, 9
APR2000, 50_69, 2259, 0, 17
MAY2000, 20_49, 2259, 0, 12
MAY2000, 50_69, 2259, 0, 19
```