

# Web Survey Methodology

Mario Callegaro, Katja Lozar Manfreda, Vasja Vehovar

## Contents

Preface	xi
About the Web Survey Methodology book	xi
About the accompanying Web Survey Methodology website	xiii
Acknowledgements	xv
About the authors	xvii
List of acronyms	xix
<b>1 Survey research and web surveys</b>	<b>3</b>
<b>1.1 Definition and typology</b>	<b>4</b>
1.1.1 Surveys and web surveys	4
1.1.2 The users of web surveys	6
1.1.3 Sampling aspects of web surveys	7
<b>1.2 Web survey process</b>	<b>9</b>
<b>1.3 Evolution of web surveys, applications and related practices</b>	<b>11</b>
1.3.1 Web surveys vs Internet and online surveys	12
1.3.2 Precursors and alternatives to web surveys	13
1.3.3 Rise of modern web surveys	15
1.3.4 The context of survey modes	17
1.3.5 Overview of advantages and limitations	18

1.3.5.1 Costs	18
1.3.5.2 Speed of data collection	20
1.3.5.3 Ease of implementation	21
1.3.5.4 Computerization of the questionnaire	22
1.3.5.5 Multimedia	22
1.3.5.6 Time and geographic flexibility	23
1.3.5.7 Self-administration	23
1.3.6 Applications of web surveys	25
1.3.6.1 Web surveys of highly covered specific populations	25
1.3.6.2 Customer satisfaction surveys	26
1.3.6.3 General population surveys	27
1.3.6.4 Business surveys	28
1.3.6.5 Online panels	29
1.3.6.6 Online community surveys	29
1.3.6.7 Website evaluations	30
1.3.6.8 Entertainment polls, questionnaires and quizzes	31
1.3.7 Related practices of web surveys	31
1.3.7.1 Administrative forms on the web	31
1.3.7.2 Internet testing	31
• Open mode	
• Controlled mode	
• Supervised mode	
• Managed mode	
1.3.7.3 Psychological web experiments	33
<b>2 Pre-fielding</b>	<b>35</b>
<b>2.1 Mode elaboration</b>	<b>36</b>
2.1.1 Selecting the web survey mode	36
2.1.2 Mixed-mode setting	37

2.1.2.1 Mixed-mode surveys, multimode contacting and mixed-mode systems	37
2.1.2.2 Problems with mixed-mode surveys	38
2.1.2.3 Web surveys in mixed-mode setting	39
<b>2.2 Sampling</b>	<b>40</b>
2.2.1 Defining populations	41
2.2.2 Sampling design principles in probability web surveys	42
2.2.2.1 Sampling frames	42
2.2.2.1.1 Sampling Frames	44
2.2.2.1.2 Traditional Sampling Frames	45
2.2.2.1.3 Panels of Units	45
2.2.2.2 Sampling techniques	45
2.2.2.3 Sample size	46
2.2.3 Sampling design issues in non-probability web surveys	48
2.2.3.1 Recruitment for non-list-based web surveys	48
<ul style="list-style-type: none"> <li>• Crowdsourcing</li> <li>• Online social media</li> <li>• Network sampling</li> <li>• Guerrilla recruitment</li> </ul>	
2.2.3.2 Sampling design approximations in non-probability setting	51
2.2.3.2.1 Approximating Probability Sample Designs	51
2.2.3.2.2 Direct Incorporation of Probability Sampling Principles	52
2.2.4 Statistical inference in probability and non-probability surveys	52
2.2.4.1 Statistical inference with probability samples	53
2.2.4.2 Statistical inference with non-probability samples	54
2.2.4.3 Deciding on the probability –non-probability sampling dilemma	56
2.2.4.4 Concluding remarks on non-probability web surveys	57
2.2.5 Non-coverage problem	59

<b>2.3 Questionnaire preparation</b>	<b>61</b>
2.3.1 General issues	61
2.3.1.1 Typologies of survey questions	61
• Nominal measurement	
• Ordinal measurement	
• Interval measurement	
• Ration measurement	
2.3.1.2 Typologies of survey questions	61
2.3.1.3 Question development process	62
2.3.1.4 Question and questionnaire design principles	63
2.3.1.5 Measurement process and measurement errors	64
2.3.1.6 Cognitive aspect of the response process	64
2.3.1.7 Specifics of non-substantive responses	64
2.3.1.8 Question banks	65
2.3.1.9 Specifics of web questionnaires	65
2.3.2 Question types	66
2.3.2.1 Single item questions	67
• Questions with multiple items	
• Other question types	
2.3.2.1.1 Questions With Nominal Measurement	68
• Open-ended text entries	
• Radio buttons	
• Drop-down menu	
• Numeric entry	
• Advanced graphical presentation	
2.3.2.1.2 Tables with Questions at Ordinal Measurement	81
• Radio buttons	

<ul style="list-style-type: none"> <li>• Drop-down menu</li> <li>• Numeric entry</li> <li>• Advanced graphical presentation</li> </ul>	
2.3.2.1.3 Tables with Questions at Interval and Ratio Measurement	83
<ul style="list-style-type: none"> <li>• Open-ended numeric entry</li> <li>• Radio buttons and drop-down menus</li> <li>• Continuous scale</li> <li>• Advanced graphical presentation</li> </ul>	
2.3.2.2 Questions with multiple items	77
2.3.2.2.1 Tables with Questions at Nominal Measurement	78
<ul style="list-style-type: none"> <li>• Tables of open-ended text entries</li> <li>• Tables of radio buttons</li> <li>• Tables of drop-down menus</li> <li>• Tables of numeric entries</li> <li>• Series of dichotomous questions</li> <li>• Tables of checkboxes and multiple selection boxes</li> </ul>	
2.3.2.2.1 Tables with Questions at Ordinal Measurement	81
<ul style="list-style-type: none"> <li>• Tables of rating scales with radio buttons</li> <li>• Unfolded tables – horizontal scrolling matrix</li> <li>• Other tables with questions at nominal measurement</li> </ul>	
2.3.2.2.2 Tables with Questions at Interval and Ratio Measurement	83
2.3.2.2.3 Combined tables	83
2.3.2.3 Other question types	84
2.3.2.3.1 Ranking	84
2.3.2.3.2 Constant sum	84
2.3.2.3.3 Extensive use of graphics	85
<ul style="list-style-type: none"> <li>• Graphics for illustrating survey questions</li> </ul>	

<ul style="list-style-type: none"> <li>• Graphics for simulating social presence</li> <li>• Graphics for forming new question types <ul style="list-style-type: none"> <li>◆ Calendar layout</li> <li>◆ Virtual reality environment</li> <li>◆ Heat-map</li> </ul> </li> </ul>	
2.3.2.3.4 Content-specific implementations	86
<ul style="list-style-type: none"> <li>• Ego-centric social network questions</li> <li>• Conjoint analysis questions</li> <li>• 360° or multi-rater feedback</li> </ul>	
2.3.2.3.5 Questions including observations and external data	87
2.3.3 Questionnaire structure, computerization and lay out	87
2.3.3.1 Structure of the questionnaire	87
2.3.3.1.1 Question order and context effects	87
2.3.3.1.2 Page breaks and paging vs scrolling	88
<ul style="list-style-type: none"> <li>• Non-question pages and sections</li> <li>• An introduction page (splash page, welcome screen)</li> <li>• Transition page</li> <li>• Instruction pages</li> <li>• An incentive or raffle page</li> <li>• File upload page or section</li> <li>• Thank you page</li> </ul>	
2.3.3.1.3 Questionnaire navigation	90
2.3.3.1.4 Blocks of questions	90
2.3.3.1.5 Special Layouts	91
2.3.3.2 Computerization, interactivity and dynamics	91
2.3.3.2.1 Branching	91
2.3.3.2.2 Randomization	92

2.3.3.2.3 Real-time validations and prompt messages	92
2.3.3.2.4 Item nonresponse prompts	93
2.3.3.2.5 Data-piping	94
2.3.3.2.6 Progress indicator	94
2.3.3.2.7 Other interactive features	95
2.3.3.3 Visual layout	95
2.3.3.3.1 Basic layout	96
2.3.3.3.2 Logotypes	96
2.3.3.3.3 Position and Navigation of other action buttons	96
2.3.3.3.4 Position and formatting of additional instructions	97
2.3.3.3.5 Position and formatting of definitions	97
2.3.3.3.6 Visual layout of help and other survey-related information	97
2.3.3.3.7 Adaptability to various screen resolutions	97
2.3.4 Ensuring survey engagement	98
2.3.4.1 Engagement, flow and user experience	98
2.3.4.2 Survey gamification	99
2.3.4.3 Questionnaire length	101
2.3.4.4 Survey satisficing	102
• Pre-fielding	
• Fielding	
• Post-fielding	
2.3.5 Web questionnaire testing	104
2.3.5.1 Traditional testing approaches	105
• Conventional pre-testing	
• Expert evaluation	
• Cognitive interviews	
• Respondents' debriefings	

- Focus groups
- Experiments
- Pilot studies
- Statistical models
- Questionnaire assessment tools

#### 2.3.5.2 Specific testing approaches for web questionnaires

107

- Questionnaire technical testing
  - Checking for technical correctness and debugging
  - Stability across browsers, operating systems, devices
  - Compliance with standards
  - Other technical aspects
- Electronic observations
  - Analysis of automatically collected paradata
  - Screen recording and respondents' video sessions
  - Eye tracking
  - Brain activity
- Online commenting
- Usability evaluation
- Integrated automated diagnostics
  - The questionnaire preview
  - The test questionnaire
  - Real-time programming controls
  - Response simulations
  - Paradata
  - The online commenting system
  - Automated diagnostics

▪	Integration of the support for capturing and processing screenshots or videos	
2.3.6	Integration of questionnaire development, computerization and testing	112
<b>2.4</b>	<b>Technical preparations</b>	<b>115</b>
2.4.1	Preparations for capturing and organizing data	115
2.4.1.1	Database preparations	115
	<ul style="list-style-type: none"> <li>• Questionnaire data</li> <li>• Paradata</li> <li>• Auxiliary data</li> <li>• Metadata</li> <li>• Capability</li> <li>• Flexibility</li> <li>• Speed</li> </ul>	
2.4.1.2	Respondent tracing and unique identifiers	116
	<ul style="list-style-type: none"> <li>• List-based web surveys</li> <li>• Non-list-based surveys</li> </ul>	
2.4.1.3	Email settings	118
2.4.1.4	Sample management	120
2.4.2	Paradata collection	121
2.4.2.1	Contact-info paradata	121
	<ul style="list-style-type: none"> <li>• Outcomes of an email invitation</li> <li>• Access to the questionnaire introduction page</li> <li>• Last question answered before breakoff</li> </ul>	
2.4.2.2	Device-type paradata	122
	<ul style="list-style-type: none"> <li>• User agent string</li> <li>• Detection of JavaScript and Flash</li> <li>• Cookie and IP recordings</li> </ul>	

2.4.2.3 Questionnaire navigation paradata	122
• Time spent per screen or response latency	
• Keystrokes and mouse-clicks	
• Change of answers	
• Real-time validation messages	
2.4.2.4 Software for collecting paradata	123
2.4.2.5 Analysis of paradata	124
• The first level paradata	
• The second level paradata	
• The third level paradata	
• The fourth level paradata	
2.4.2.6 Paradata privacy and ethics	124
2.4.3 Preparations to support security and privacy	125
2.4.3.1 Securing the server	126
2.4.3.2 Securing the web survey response process	127
2.4.3.3 Technical aspects of anonymity	127
2.4.3.4 Securing physical and electronic access to the datafile	128
2.4.3.5 Access auditing	128
2.4.3.6 Relation to legislation and professional standards	129
2.4.4 Technical pre-testing of web survey software	129
<b>2.5 Nonresponse strategy</b>	<b>130</b>
2.5.1 General issues in web survey nonresponse	130
2.5.1.1 Nonresponse theories	131
2.5.1.2 Nonresponse process in web surveys	132
• The contacting phase	
• The cooperation phase	
2.5.1.3 Factors affecting web survey nonresponse	133

<ul style="list-style-type: none"> <li>• The socio-technical context <ul style="list-style-type: none"> <li>▪ Over-surveying</li> <li>▪ Spam and viruses</li> <li>▪ Internet privacy</li> </ul> </li> <li>• Information-communication infrastructure</li> <li>• Respondent's characteristics</li> <li>• Web survey design</li> </ul>	
2.5.2 Defining nonresponse, breakoffs and completeness	135
2.5.2.1 Indicators of unit nonresponse	136
<ul style="list-style-type: none"> <li>• Probability list-based samples with mail invitations</li> <li>• Repeated probability surveys</li> <li>• Non-probability list-based web surveys</li> <li>• Non probability non-list web surveys</li> <li>• Probability non-list-based samples</li> </ul>	
2.5.2.2 Breakoffs	138
2.5.2.3 Item nonresponse at the question level	140
2.5.2.4 Item nonresponse at unit level	140
2.5.2.5 Completes, partials and unusable units	140
2.5.3 Relation between nonresponse rates and nonresponse biases	141
2.5.4 General web survey design characteristics and nonresponse	143
2.5.4.1 Survey sponsor	146
2.5.4.2 Topic salience	146
2.5.4.3 Questionnaire design	148
2.5.5 Incentives in web surveys	148
2.5.6 The role of the contacting process: mode, timing and frequency	149
2.5.6.1 Mode of contact	151
2.5.6.2 Day and time of sending invitations	152

2.5.6.3	Number of contacts and their scheduling	152
2.5.6.4	Length of the fieldwork period	153
2.5.7	Specifics of the invitation format	154
2.5.7.1	Personalization of the invitations	155
2.5.7.2	Providing motivation and tone of request	156
2.5.7.3	Format and visual design of email invitations	156
•	Visual appearance	
•	Length	
•	Spam	
•	Header of the email invitation – ‘From’ line	
•	Header of the email invitation – ‘Subject’ line	
2.5.8	Towards a nonresponse strategy	158
<b>1.1</b>	<b>General management</b>	
<b>3</b>	<b>Fielding</b>	<b>165</b>
<b>3.1</b>	<b>Recruiting</b>	<b>166</b>
<b>3.2</b>	<b>Measurement</b>	<b>167</b>
<b>3.3</b>	<b>Processing and monitoring</b>	<b>168</b>
3.3.1	Activation, soft launch, full launch and deactivation	168
3.3.2	Providing help	169
3.3.3	Fieldwork interventions	170
3.3.4	Monitoring of the fieldwork process	171
<b>4</b>	<b>Post-fielding</b>	<b>175</b>
<b>4.1</b>	<b>Data preparation</b>	<b>176</b>
4.1.1	Establishing response statuses	176
4.1.1.1	Eligibility, completeness and quality status of the units	177
4.1.1.2	Missing values and non-substantive responses	178

4.1.1.3 Reporting on response statuses	178
4.1.2 Validation and editing	179
• Basic validations	
• Item nonresponse validation	
• Advanced validations	
• Substantive validations	
4.1.3 Imputations	180
4.1.4 Weighting	182
4.1.5 Coding and recoding	183
<b>4.2 Preliminary results</b>	<b>185</b>
4.2.1 Reporting	185
4.2.2 Dissemination	186
<b>4.3 Data exporting and documentation</b>	<b>187</b>
4.3.1 Data exporting	187
4.3.2 Statistical disclosure	187
4.3.3 Documentation	188
4.3.4 Data archiving	189
<b>5 Selected topics in web survey implementation</b>	<b>191</b>
<b>5.1 Smartphones, tablets and other devices</b>	<b>192</b>
5.1.1 Introduction to multiple devices	192
5.1.2 Survey strategies for handling mobile devices	195
• Doing nothing	
• Explicitly discouraging the use of mobile device for surveying	
• Optimizing web questionnaire for mobile browsers	
• Providing a questionnaire as a mobile application, developed for a particular survey (custom mobile survey application)	

<ul style="list-style-type: none"> <li>• Providing a questionnaire through a general application for surveying on mobile devices (general mobile survey application)</li> </ul>	
5.1.3 Device effects	196
<ul style="list-style-type: none"> <li>• Faster reaction, but longer questionnaire completion times on smart phones</li> <li>• Lower participation rates on smartphones</li> <li>• Higher breakoff rates on smartphones</li> <li>• Shorter length of answers to open-ended text questions on smartphones</li> <li>• Mixed findings on social desirability</li> <li>• Unclear evidence on differential satisficing behavior on smartphones</li> </ul>	
5.1.4 Designing web questionnaires for smart phones	199
5.1.5 Browser-based mobile surveys vs mobile survey apps	201
<ul style="list-style-type: none"> <li>• Displaying the questionnaire</li> <li>• Data management</li> </ul>	
5.1.6 Software for multi-device web surveys and mobile survey apps	203
5.1.7 Challenges and conclusions	204
<b>5.2 Online panels</b>	<b>205</b>
5.2.1 Terminology and typology	206
5.2.2 Recruitment, sampling and selection	206
<ul style="list-style-type: none"> <li>• Non-probability online panels</li> <li>• Probability online panels</li> </ul>	
5.2.2.1 Recruitment methods for non-probability online panels	207
5.2.2.2 Recruitment methods for probability online panels	207
5.2.2.3 Sampling and selecting online panel members for individual surveys	208
5.2.3 Maintaining an online panel	209
5.2.4 Online panel study designs	210
5.2.5 Online panel data quality	210

5.2.5.1 Comparison of online panel data quality and results	210
5.2.5.2 Effect of weighting on online panel data quality	212
5.2.5.3 Other data quality issues in online panel research	212
5.2.5.4 Conclusions on online panel data quality	213
5.2.6 Buying online panel research	214
5.2.7 The future of online panels	214
<b>5.3 Web survey software</b>	<b>215</b>
5.3.1 Integrated ICT support for the web survey process	217
5.3.2 The web surveys software market	219
5.3.3 Characteristics of web survey software	220
5.3.4 Discussion of two selected features	223
5.3.4.1 Interface speed	223
5.3.4.2 Critical features of the GUI	224
5.3.5 Finding the right tool	225
• Preliminary list of specifications	
• Shortlist of specifications and criteria	
• Shortlisting of the tools	
• Selection process	
5.3.6 Trends in web survey software	226
<b>6 Broader context of web surveys</b>	<b>231</b>
<b>6.1 Broader methodological context</b>	<b>232</b>
6.1.1 Web survey mode revised	232
• Interviewer involvement	
• Usage of computer technology	
• Information transmission medium	
• Question presentation input channel	

• Closeness of the interviewer interaction	
6.1.2 Total survey error and survey data quality	233
6.1.3 Evaluation of the web survey mode effect	235
6.1.4 Cost-error optimization	236
6.1.5 Beyond the core web survey process	237
6.1.5.1 The step of preliminary research activities	237
6.1.5.2 The step of advanced analysis, processing and valorization	238
6.1.5.3 Mixed-modes	238
6.1.5.4 Mixed methods	240
6.1.6 Towards interactive fieldwork design	240
<b>6.2 Web surveys within the project management framework</b>	<b>242</b>
<b>6.3 The web survey profession</b>	<b>244</b>
6.3.1 Related disciplines	245
6.3.2 Legal issues	246
• Unsolicited email invitations	
• Cookies	
• Spam	
• Incentive	
• Lotteries or sweepstakes	
• Privacy, confidentiality, anonymity	
• Paradata	
• Minors	
• Intellectual property rights	
• Archiving	
• Software	
• Illegal content	
• Disclosure	

• Security	
6.3.3 General ethical concerns	246
6.3.4 Formalized certificates, codes and standards	247
6.3.5 Informal guides, recommendations and other material	248
<b>6.4 Web survey bibliography</b>	<b>249</b>
<b>7 Future of web surveys</b>	<b>255</b>
<b>7.1 General technological developments</b>	<b>256</b>
• Voice input	
• Intelligent data analysis	
• Online collaboration tools	
• Multiple devices	
• Body as a language	
• Development of electronic micropayments	
• Direct shift to smartphone web surveys in developing countries	
<b>7.2 Web survey software</b>	<b>257</b>
• Pre-fielding	
• Fielding	
• Post-fielding	
<b>7.3 Methodology</b>	<b>258</b>
• Rise and decline of mixed-mode data collection	
• Increased use of paradata	
• Increased integration with qualitative research	
• Technology-enhanced complementary data collection	
• Technology-enhanced survey questionnaires	
• Augmentation and integration of survey data with big data	
• The rise of customer panels and online communities	

- Social media recruitment, real-time recruitment and river samples
- Inference from non-probability samples

**7.4 Broader business and societal issues** **259**

- Further integration with administrative and business processes
- Expansion of Do-It-Yourself (DIY) research
- Democratization of research
- Decreasing participation in surveys
- Rise and decline of the web survey profession
- Decline of the survey industry?

**8 Conclusions** **262**

Bibliography 267

Author Index 301

Subject Index 313