A Technical Guide to Designing and Implementing Effective Web Surveys

Gregory Baatard

Edith Cowan University
A Technical Guide to Designing and Implementing Effective web Surveys

Greg Baatard
Edith Cowan University, Perth, Australia
g.baatard@ecu.edu.au

Abstract: The Internet is becoming an increasingly prominent medium for the administration of surveys. Although individual findings vary, the majority of the literature agrees that the appropriateness and response rates of web surveys is expected to rise in the future as we enter a generation of “digital natives” and mail-based communication becomes increasingly antiquated. Just about every aspect and tenet of traditional survey methodology has received attention in academic literature, positioning it as one of the most mature data collection techniques and a mainstay in all areas of research. While much of this accumulated knowledge is applicable and relevant to web surveys, there are numerous issues that arise specifically when surveys are delivered online. Such issues concern the overall design, delivery and administration of web surveys and the structure, presentation and layout of their questions. The decisions made in these areas can influence the efficacy of a web survey in a number of ways, including the rate, integrity and quality of responses. This paper discusses such issues, and makes a number of recommendations to assist researchers in manually developing an effective web survey and in evaluating survey creation products and services.

Keywords: web, online, survey, questionnaire, guide

1. Introduction

Dillman, Smyth and Christian (2009) describe the past two decades as “turbulent times” for the survey methodology. The telephone-based surveys popular throughout the 1970s and 1980s have become substantially less popular amongst researchers, marred by numerous factors including a surge in telemarketing and the distasteful phenomenon of receiving calls during the “dinner hour”. While still heavily used and indeed the most suitable medium in some demographics, mail-based surveys have begun to lose ground with advances in technology spurring web surveys to prominence. As Internet access and eMail become almost synonymous with the ownership of or access to a personal computer, delivering self-administered surveys via such technologies has become an increasingly common practice. Doing so offers the same benefits of mail surveys – allowing the researcher to reach potential respondents all over the world with minimal cost of data collection and processing. Furthermore, web surveys require less effort by respondents to complete and return than mail surveys (Couper 2000; Fricker and Schonlau 2002; Lyons, Cude et al. 2005; Deutskens, de Ruyter et al. 2006; Couper and Miller 2008; Malhotra 2008; Dillman, Smyth et al. 2009; Vicente and Reis 2010). The medium also affords greater control in terms of validation and the delivery of the survey items; for example, conditional questions can be presented only if required (Oppenheim 1992; Couper 2000; Fricker and Schonlau 2002; Shropshire, Hawdon et al. 2009; Vicente and Reis 2010).

The already low response rates typically seen in mail surveys are set to become lower still as more and more communication occurs electronically, making the completion and return of a printed survey a somewhat archaic process. Dillman, et al. (2009) echo this sentiment, stating that “the shift toward eMail as the communication mode of choice for significant sectors of the population is somewhat ironic as it is one of the very factors that make Internet surveys possible, but it is also making surveys by traditional modes more difficult to complete.” Web surveys normally require no more than eMail and/or a Web browser to complete and submit – both of which the large majority of today’s computer users are familiar with. Furthermore, the completion of trivial online quizzes, polls and tests has emerged as a social pastime for many Internet users, particularly those in teen and young adult demographics (Fricker and Schonlau 2002). Although done for amusement, these activities serve to familiarise Internet users with the mechanics of web surveys and potentially increase their willingness to respond to other surveys.

Several studies conducted in the past decade have found that web surveys can achieve similar response rates to mail surveys, particularly for younger respondents and those in demographics that regularly use the Internet. Some of the studies finding response rates to mail surveys to be higher than those of web surveys have also acknowledged the need to take demographics into account and the likelihood of response rates differing in the future. These trends suggest that web surveys will become increasingly prominent and result in higher response rates as the population becomes increasingly made up of “digital natives” (Jones and Pitt 1999; Cook, Heath et al. 2000; Couper 2000;
The issues of survey design are well established and mature in academic literature, covering topics such as question wording, question types and survey length. While such issues are entirely applicable and must be respected in web surveys, this paper focuses upon issues which arise specifically when surveys are delivered online. The issues discussed can have a significant impact upon the accessibility and usability of web surveys, potentially influencing both response rates and the quality of responses themselves—“Before writing a questionnaire for the web, it is important to remember that a poorly designed survey can discourage people from responding, and it can also give skewed results” (Gonzalez-Bañales and Adam 2007). Several of the issues and recommendations are of a technical nature, written with the assumption that interested readers will have a moderate level of technical knowledge or at least an understanding of the technical aspects of developing and implementing a web survey.

2. Overall design, delivery and administration issues

This section discusses the primary issues related to the overall design, delivery and administration of web surveys as a whole. The first issue is that of the delivery and hosting of the survey. The author recommends against utilising eMail for anything other than contacting potential respondents, despite possibly appearing as a method of minimising the effort required to respond. While it is possible to embed or attach a survey to an eMail, this practice is relatively uncommon and is less likely to be supported in a uniform fashion by eMail client software (Fricker and Schonlau 2002; Lyons, Cude et al. 2005). Thus, response rates of eMail-based surveys may be hampered by both unfamiliarity and technical hurdles. Web surveys should be hosted on the Internet as standard Web pages. They should minimise any reliance on supporting software or technologies such as Flash or JavaScript (Kaczmirek 2005; Gonzalez-Bañales and Adam 2007). As discussed in the next section, if such technologies are used they should not be relied upon and the survey should function properly without their presence. Like all Web pages, all web surveys should be tested thoroughly before deployment to ensure that they appear and function correctly in all likely browsers, browser versions, screen resolutions, and so forth (Fricker and Schonlau 2002; Thompson, Surface et al. 2003; Kaczmirek 2005; Lyons, Cude et al. 2005; Gonzalez-Bañales and Adam 2007).

Access to and availability of web surveys must also be considered. In addition to adhering to appropriate sampling procedures (Simsek and Veiga 2001; Lyons, Cude et al. 2005; van Selm and Jankowski 2006; Couper and Miller 2008; Vicente and Reis 2010), web surveys should require some form of unique identifier such as an ID number, code, IP address or eMail address to deter automated responses or multiple responses by a single person (Couper, Traugott et al. 2001; Thompson, Surface et al. 2003; Lyons, Cude et al. 2005). Such measures can increase the quality of the data gathered by deterring behaviour which negatively influences response data. When using an ID number or access code, Couper, et al. (2001) recommend avoiding those containing potentially ambiguous characters such as the letter “l” and the number “1” or the letter “o” and the number “0”. Limiting the availability of a web survey has the potential to strengthen the data set. For example, if a survey is to be administered only to employees of a single organisation, it may be appropriate to deploy it in such a way that it can only be accessed via the organisation’s internal network (Simsek and Veiga 2001). Controlling and limiting the availability of a web survey also serve to prevent abuse. Though the likelihood of this is minimal for most surveys, this is of particular importance if the survey concerns topics considered to be controversial, sensitive or high-profile. The potential for abuse has been illustrated several times by the abuse of online polls—a different but definitely similar scenario. Time magazine’s 2009 “Time 100” poll was manipulated with enough sophistication to arrange the top 21 results (Schonfeld 2009). A similar hoax in 2010 targeted singer Justin Beiber; an online poll being abused in order to add North Korea as a destination in his first world tour (Emery 2010).

It is worthwhile at this point to acknowledge that numerous applications and Web-based services exist, which can be used to create and administer web surveys. Modern products offer high degrees of customisability and sophistication, allowing a web survey to be created and administered with minimal technical knowledge. While reviews or recommendations of these are not included in this paper (interested readers see for example, Thompson, Surface et al. 2003; Wright 2005 and http://websm.org/), the issues discussed here can serve as criteria by which to assess such
applications or services. In addition, one must consider factors such as the cost and the confidentiality, security and ownership of data when determining their suitability (Lyons, Cude et al. 2005; Wright 2005; Greenlaw and Brown-Welty 2009). Despite the availability of these products, various factors often make a purpose-built web survey a necessity.

A final consideration pertinent to the overall design of web surveys is that of the processing and storage of responses. The ability to store response data directly in a database or spreadsheet format is a definite advantage of administering surveys electronically, and this can be maximised with a well-designed web survey (Lyons, Cude et al. 2005; Gonzalez-Bañales and Adam 2007; Greenlaw and Brown-Welty 2009). A web survey is essentially a form, and as such must be processed like any other Web-based form once submitted. Although it is possible to create a form which simply eMails the response to the researcher, entering the response into a database is a much more effective method. This can be facilitated by ensuring that the names of the form elements such as text fields and radio button groups correspond to the database column names, allowing simple and generic form processing scripts to be created. Care should be taken to ensure that web surveys are submitted and stored in a secure manner, utilising appropriate encryption (Lyons, Cude et al. 2005). As well as being easily exportable to spreadsheet or statistical analysis software formats, survey responses in a database can be queried using SQL (Structured Query Language), offering sophisticated means of extracting and visualising meaningful information. To allow effective querying, the values submitted in a web survey should be numeric wherever appropriate. For example, a five-point Likert scale question ranging from Strongly Disagree to Strongly Agree and including a Neutral response should be stored as -2 to 2 with 0 representing Neutral, rather than storing responses textually.

This section has mainly discussed issues that aim to improve the quality, correctness, usability and reliability of data gathered via web surveys. The recommendations are overarching – benefiting the researcher directly, rather than via the improvement of the respondent’s experience. The following sections focus upon issues that improve the respondent’s experience.

3. Structure, presentation and question layout issues

This section discusses issues of structure, presentation and question layout within web surveys that can influence the respondent’s experience. Regardless of the medium, a survey that is clear, convenient and usable is of more appeal to potential respondents, resulting in higher response rates and potentially better quality data. The author feels it is important to make all relevant information available with the survey itself. While eMail or other methods may be used to introduce the research and request participation, all pertinent information should also be available with the web survey. While this may result in some redundancy, it ensures that respondents do not need to search outside of the survey Web page for any instructions or information they may need. While minor, inconveniences such as these have the potential to deter a respondent – particularly if they encounter the inconvenience before beginning the survey, when they have not yet invested any time or effort into its completion.

The issue of survey length has been addressed in numerous pieces of academic literature, often finding that longer surveys are not only less likely to be completed, but also that “questions asked later in the questionnaire bear the risk of producing lower quality data, especially if they are in open format or in long grids” (Galesic and Bosnjak 2009). Based on such findings (Herzog and Bachman 1981; Gonzalez-Bañales and Adam 2007), two recommendations that are applicable to surveys in any medium are worth repeating. Lengthy surveys should be avoided if possible, and the ordering of questions should be considered in longer surveys – acknowledging that the quality of responses to questions towards the end may be lower than that of those near the beginning (Vicente and Reis 2010). A further consideration regarding the ordering of questions arises from Shropshire, Hawdon and Witte (2009), which found respondent interest to be a significant factor in early termination of a questionnaire. Hence, questions likely to be of highest interest to respondents should appear early in the survey. Two further recommendations can be made for lengthy web surveys in particular. Firstly, the design and presentation of the survey can be tailored to make a long survey more manageable for respondents. One method is to divide the survey into sections and display one section at a time (Figure 1).
Greg Baatard

Figure 1: Example of a long survey divided into sections

This serves to avoid overwhelming the respondent with a large number of questions on one page and the need to scroll, which have been found to influence item non-response and survey non-completion rates (Couper, Traugott et al. 2001; Schonlau, Fricker et al. 2002; Toepoel, Das et al. 2008; Vicente and Reis 2010). Research (see for example, Tourangeau, Couper et al. 2004; Dillman, Smyth et al. 2009; Toepoel, Das et al. 2009) has found that respondents see questions on the same page as being more closely related, having a slight impact upon the intercorrelation of responses. This should be taken into account when determine if and how to divide a survey into multiple pages. An overall progress bar and section-by-section validation are advisable if a survey is divided into sections. Yan, Conrad, Tourangeau and Couper (2010) examine the use of a progress indicator in web surveys, concluding that it encourages completion if it accurately reflects the expected duration of the survey, particularly if the survey is short. This further emphasises the need to minimise the length of web surveys and to provide information such as its expected duration. Other research regarding the use of progress indicators in surveys divided into sections is generally supportive, however findings vary and the effects of progress indicators are often minor (Couper, Traugott et al. 2001; Crawford, Couper et al. 2001; Vicente and Reis 2010).

The second recommendation specific to lengthy web surveys is giving respondents the ability to save their progress and resume the survey at a later time. This is recommended in Kaczmirek (2005), who states “Do not introduce problems in your online questionnaire which would not occur in a paper and pencil questionnaire.” While implementing the ability to save and resume survey progress is likely to be beyond the technical capabilities of a person otherwise capable of creating a web survey, the feature is common amongst applications and Web-based services that allow surveys to be created.

Unlike mail surveys, web surveys can adapt to a respondent’s answers to questions on the fly. In a mail survey, conditional questions are typically implemented via instructions to the respondent – for example, “If you answered ‘No’ to this question, continue to Question 12”. Using Web-based scripting languages such as JavaScript, web surveys can implement conditional questions in a more effective manner, showing or hiding parts of the survey based on the answers to questions (Oppenheim 1992; Fricker and Schonlau 2002; Gonzalez-Bañales and Adam 2007). In addition to helping to ensure the correctness of data, this serves to “reduce the length of a survey to the individual minimum” (Kaczmirek 2005). Though JavaScript can improve the implementation of web surveys, it is inadvisable to rely upon it heavily. Ideally it should be possible and convenient to complete the survey without requiring the respondent’s browser to have JavaScript enabled. If the survey cannot be completed without JavaScript, potential respondents with JavaScript disabled should be informed of its need when they attempt to access it.

The following recommendations concern the presentation, layout, accessibility and formatting of web survey questions. While some of the recommendations are relatively minor, they are all worthy of consideration when creating a web survey or deciding which survey-generation product to use. Research such as that of Christian, Dillman and Smyth (2007) underpin the potential effects of seemingly minor issues in web survey presentation, finding that the relative size of text fields resulted in an eight percent increase in correctly formatted responses and the use of symbols rather than words as labels resulted in a seven percent increase. In order to avoid excess technical detail, the recommendations are presented in dot-point form. The term “form elements” refers to text fields, radio
buttons, checkboxes and other form components used to allow people to respond to questions in web surveys.

- Make use of space to clearly separate questions from one another. If utilizing a grid-based layout, often used to present series of Likert scale questions, use alternating background colours to clearly align question text with the appropriate form elements.

- Use spacing to ensure that radio buttons and checkboxes are clearly associated with their labels, preventing ambiguity and confusion. Furthermore, use the “label” tag to increase the clickable region of form elements. These points are illustrated in Figure 2.

![Figure 2: Likert scale with ambiguous spacing and no labels (A) and Likert scale with clear spacing and labels (B). Clickable regions have been highlighted](image)

- If there is insufficient space to repeat labels next to each radio button or checkbox, consider using the “title” attribute in the “<input>” or “<label>” tag. The content of this attribute will appear as a tooltip when the mouse is over the radio button, allowing respondents to select a response without needing to refer back to other labels.

- The layout of Likert scale responses is explored in Tourangeau, Couper and Conrad (2004), who find that respondents make inferences about the meaning of survey items based on visual cues such as the perceived midpoint of responses, their spacing and order. These findings should be taken into consideration when designing Likert scales in web surveys.

- If a question requires participants to select from a number of options, use radio buttons rather than drop down lists as radio buttons present the options in a more convenient and accessible manner. Healey (2007) found that drop-down lists result in higher item non-response rates, longer response times and increased accidental changing of responses. Kaczmirek (2005) recommends only using drop-down lists “if respondents know the answer without having to look at all entries”, and to make the default selection a placeholder such as “select here” so that non-response can be identified.

- Users can navigate form elements via the tab key, with each press bringing the focus to the next element from top-left to bottom-right. The layout of form elements in a survey may cause the default tab order to traverse form elements in an incorrect or unintuitive manner. If this is the case, the ordering can be changed via the “tabindex” in form element tags. The “accesskey” attribute can also be used to allow keyboard shortcuts to form elements if deemed necessary for accessibility reasons.

- On the spot validation of responses is possible in web surveys, via scripting languages such as JavaScript. While it is possible to enforce strict validation criteria such as only allowing digits to appear in a text field for a phone number, it is not always advisable to do so. When faced with overly strict, inappropriate or frustrating validation requirements, respondents may invent data, attempt to circumvent the validation, or discontinue the survey entirely (Best and Krueger 2004; Kaczmirek 2005; Christian, Dillman et al. 2007; Gonzalez-Bañales and Adam 2007; Vicente and Reis 2010). It is also worthwhile reiterating that reliance upon JavaScript is not recommended. Failing to take heed of issues such as these can lead to web surveys which are unclear, confusing and frustrating, often resulting in discontinuation or by respondent, or lower quality data. By taking these recommendations into consideration, researchers can design, develop and implement an effective web survey, or select a product which will allow them to create one.

4. Conclusion

There is much academic literature which addresses the design and implementation of effective surveys. Of that, a relatively small proportion discusses issues specific to the increasingly prominent medium of Internet-based surveys. Web surveys present unique opportunities and challenges not applicable to mail-based surveys, and care must be taken to ensure that web surveys are developed and implemented in an effective manner. This paper has discussed numerous issues, primarily technical, that should be taken into consideration in regards to web surveys.
Issues pertaining to the overall design, delivery and administration of web surveys were covered first. EMail-based surveys are advised against, and the importance of appropriate advertising, unique identifiers and data format were discussed. The recommendations in this section seek to give researchers overarching guidance towards creating web surveys that are more likely to produce data of high quality, correctness, usability and reliability. Next, a number of recommendations were made regarding the structure, presentation and layout of questions in web surveys. These focus upon increasing the clarity, convenience and usability of web surveys, promoting higher response rates and higher quality data.

As the world’s population becomes increasingly made up of "digital natives", the Internet has emerged as a key medium for self-administered surveys. The sophistication and availability of modern applications and Web-based services for the creation of web surveys has also made the medium more accessible to researchers in all domains, furthering their use. The recommendations made in this paper assist researchers both in manually developing a web survey and evaluating survey creation products. A well-developed web survey provides both the researcher and respondents with a clear, accessible and highly usable mechanism, minimising the confusing and frustrating elements that may lower response rates or the quality of data.

References


Ozlu Azakli works as a research assistant at Istanbul University at the Department of Organizational Behavior. She has studied management at Bilkent University in Ankara. Then, she has studied human resources management at Yildiz Technical University Graduate School. Now, she is a Ph.D. student at Marmara University at the Department of Organizational Behavior.

Dr Greg Baatard teaches computer science at Edith Cowan University in Western Australia, specialising in web development, programming and databases. His research interests include web-based applications, e-learning, Computer Supported Collaborative Work (CSCW) and computer-supported research methodologies.

Matthew Bell is an Integrated Project Team (IPT) Leader with experience in developing and deploying aircraft maintenance, training and capability insertion programmes through concept, design to in-service deployment. He has worked for British Aerospace, BAE SYSTEMS and completed a PhD by research in maintenance optimizing programmes for the Typhoon, Tornado, Hawk, Harrier and Nimrod aircraft.

Martin Bergaus is a PhD candidate at Leeds Metropolitan University, UK. Beside 17 years of professional experiences in the ICT, he holds 3 master degrees. He teaches as an associate lecturer at the University of Applied Science in Vienna. His current research interests focus on a qualitative research using Grounded Theory in the field of Ubiquitous Computing.

Doctor Bronwyn Betts is a Senior Lecturer at Anglia Ruskin University. She has lectured extensively on Human Resource Management programmes over a period of 26 years. She is an external examiner for CIPD programmes and has served on validation panels as an external advisor. Research interests include workplace learning and international HRM.

John Lorimer Campbell is a Senior Lecturer in nursing studies at the University of Cumbria, teaching bioscience and other clinical topics at pre and post registration levels. In addition to nursing and educational training he also holds academic qualifications in biology and health science. John produces educational materials for international distance learning.

Dr Aileen Cater-Steel is Associate Dean (Research) in the Faculty of Business and Law, and Associate Professor of Information Systems at USQ, Toowoomba, Australia. She leads two industry funded research projects in IT Service Management. Aileen’s work has been published in many international IS journals and conference proceedings and she has compiled three edited books.

Panayiotis Constanti teaches in the fields of human resources and organizational behaviour with his main research interests focusing on people at work and learning and teaching. He has published in a number of academic journals including the Service Industries Journal; Education and Training; the International Journal of Education Management and the International Journal of Contemporary Hospitality Management.

Barbara Crump’s research involves evaluation of digital divide projects and research projects investigating the culture of the computing tertiary and work environments and collaborated with research colleagues from Japan, Malaysia and the UK. She is a Senior Lecturer in the information systems group in the School of Management, Massey University, Wellington, New Zealand.

Ke Cui is a PhD student specializing in sociology and social policy. She employed life-history and narrative research methods to explore the relationship between parental expectation and their children’s status attainment. She got a bachelor’s degree of economics and management in China Agricultural University in 2009.

Leslie Dinauer is Professor and Program Director for Research Methods in the Doctor of Management Program at University of Maryland University College. She has been developing and teaching research methods online for over eight years. Her secondary research program investigates the cognitive processes of attitude formation and change, particularly within the context of consumer behavior.

Elena Fitkov-Norris is Principal Lecturer in Information Management at the Department of Informatics and Operations Management, Kingston Business School, London, UK. She received a BSc in