A Technical Guide to Effective and Accessible Web Surveys

Gregory Baatard
A Technical Guide to Effective and Accessible 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 accessibility, 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 and accessible web survey and in evaluating survey creation products and services.

Keywords: web, online, survey, questionnaire, guide, accessibility

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 potential benefits of web surveys have been recognised for over a decade – long before the Internet rose to the prominent position it currently holds. Schmidt (1997) lists access to a large population, saving time and money and allowing for dynamic/interactive surveys as the benefits of publishing a survey on the Web. Potential problems are also listed; incomplete or unacceptable responses, multiple submissions, and the security and integrity of data. Fifteen years later, these issues have largely been resolved and the benefits can be readily observed.

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; Couper, Traugott et al. 2001; Fricker and Schonlau 2002; Kaplowitz, Hadlock et al. 2004; van Selm and Jankowski 2006; Lusk, DeClos et al. 2007; Converse, Wolfe et al. 2008; Couper 2008; Couper and Miller 2008; Shih and Fan 2008). Indeed, the emergence of the Internet as a major medium for the delivery of self-administered surveys is easily observable and cannot be denied.

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). Web accessibility is an increasingly prominent topic, with a number of governments recommending or requiring Web pages to meet accessibility standards. There are a number of considerations that should be made to ensure a web survey is accessible to those with impairments or disabilities. 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; List and Gosline 2007). Thus, response rates of eMail-based surveys may be hampered by both unfamiliarity and technical hurdles. An eMail used to contact potential respondents should be concise. It should introduce the survey, tell the reader how long it should take to complete, mention an incentive if one is offered, and present a clear link to the survey itself. EMail can also be used to follow up with respondents – for example to thank them for their response, remind them to complete the survey or advise them of any outcomes. Unless the survey is entirely anonymous, the researcher has the ability to ensure that eMails are not sent unnecessarily – do not disturb someone who has already responded with an eMail reminding them to complete the survey (Kaczmarek 2005; Gonzalez-Bañales and Adam 2007).

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 (Kaczmarek 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. Such technologies also hinder the accessibility of a web survey, hindering its completion by those with some form of impairment or disability. 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; Kaczmarek 2005; Lyons, Cude et al. 2005; Gonzalez-Bañales and Adam 2007). The survey should be hosted on a Web site or domain that can be recognised as trustworthy and legitimate, such as that of an academic institution, a Web site related to the field of the research, or a well-recognised survey hosting Web site. Avoid hosting the survey on unrelated or advertising supported domains, as these can be perceived as unprofessional and untrustworthy (Kaczmarek 2005; Gonzalez-Bañales and Adam 2007).

Access to and availability of web surveys must also be considered. An all too common pitfall when conducting a web survey is to simply aim for the largest number of respondents possible – an attractive option when the costs of distributing, administering and collecting the survey have been
minimised by the medium. However, “large samples are no substitute for accurate samples” (List and Goslin 2003) – a medium that makes it easier to reach more people is no excuse to ignore 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). In addition to adhering to appropriate sampling procedures, 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). Survey responses in a database can be queried using SQL (Structured Query Language), offering a means of rapidly extracting and visualising meaningful information from the data. While SQL is unlikely to be the sole tool used for analysis, it is an eminently readable and writable language that can easily be used to explore a data set. The purpose and potential value of the sample SQL statements in Figure 1 should be largely self-evident, even to those with no knowledge of the language.

| SELECT AVG(likert_1), AVG(likert_2) FROM responses WHERE study_mode = 'OC' AND age_bracket >= 2 | SELECT age_bracket, AVG(likert_3) FROM responses WHERE study_level = 'PG' GROUP BY age_bracket |

Figure 1: Sample SQL statements used to explore survey responses in a database

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 applies to all ordinal values. A database can
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 section focuses 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). Hence, the most important questions should be placed towards the beginning of a survey. 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. These two factors should be complimentary in a well-designed survey; important questions should be of interest to a respondent. 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 2).

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; Couper 2008; 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. A number of the recommendations concern the accessibility of web surveys. The need to make the Web accessibly to those with impairments or disabilities is an increasingly significant issue, and one which should be taken into account by researchers – particularly if the research concerns a demographic with a higher proportion of impairments, such as the elderly.

- Make use of space to clearly separate questions from one another. If utilising 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 3.

Figure 3: Likert scale with ambiguous spacing and no labels (A) and Likert scale with clear spacing and labels (B). Clickable regions have been highlighted.

- 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.

- Use of the "label" tag also bolsters the accessibility of a survey (Theofanos and Redish 2003; Thatcher, Burks et al. 2006). Forms are one of the largest hurdles faced on the Web by those with impairments or disabilities. A screen reader is the most common piece of assistive technology used by the visually impaired, reading the content of a Web page. When the user focuses upon a form element such as a text field or radio button, the screen reader will read out the label associated with the field. Should a label not exist, the "name" or "id" attribute of the form element is read. This is likely to be generic, such as "question_5", or abbreviated, such as "oa_sat_lquart" ("overall satisfaction in last quarter"). The former is of little to no help to the user, and the latter will not be comprehensible to the screen reader or user.

- Acronyms and abbreviations are common in surveys, however not all respondents may be familiar with or remember them and screen readers can struggle with their pronunciation. The "acronym" and "abbr" tags can be used to alleviate this, and the "title" attribute can be used to provide further information where needed (Theofanos and Redish 2003; Thatcher, Burks et al. 2006).

- Avoid relying solely upon colours, images or subtle visual cues (such as an asterisk to indicate a required field, "greying out" a disabled form element or using the size of a form element to indicate the desired input length) to convey important information about the survey. Such cues are easily missed or indeterminable by those with low vision, colour blindness, or using a screen reader (Thatcher, Burks et al. 2006).

- 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. For example, when nonsubstantive answers such as "don’t know" or "no opinion" were not clearly separated from the Likert scale itself, responses were influenced by the different visual midpoint that this created. 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. Kaczmarek (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 is 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. These facilities can also serve to make a form more accessible to those using assistive software such as a screen reader, allowing the form to be navigated effectively without the use of a mouse.

- 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 field for a phone number, it is not always advisable to do so – for example, the
respondent may want or need to include spaces, parentheses or a plus sign in a phone number. 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). Consider including an “other” option in closed-ended questions, and allow respondents to provide open-ended comments (List and Goslin 2002). 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 Web-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. A number of recommendations have been made concerning the accessibility of web surveys, to raise awareness of the importance of ensuring that a web survey can be completed by someone with an impairment or disability.

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, underlying code, presentation and layout of questions in web surveys. These focus upon increasing the clarity, convenience, usability and accessibility of web surveys, promoting higher response rates and higher quality data. Many potential respondents to a web survey are not overly motivated or determined to complete it, and hence it is imperative to structure and present a web survey in a way that encourages convenient completion and invests the respondent early to discourage discontinuation of the survey.

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


www.ejbrm.com