IDEA GROUP PUBLISHING

701 E. Chocolate Avenue, Suite 200, Hershey PA 17033, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

ITP5213

This paper appears in *Managing Modern Organizations Through Information Technology*, Proceedings of the 2005 Information Resources Management Association International Conference, edited by Mehdi Khosrow-Pour.

Online-Questionnaire Design: Establishing Guidelines and Evaluating Existing Support

Joanna Lumsden

NRC IIT e-Business, 46 Dineen Dr., Fredericton, NB, Canada, E3B 9W4, jo.lumsden@nrc-cnrc.gc.ca

Wendy Morgan

University of New Brunswick, PO Box 4400, Fredericton, NB, Canada, E3B 5A3, f995g@unb.ca

INTRODUCTION

As a new medium for questionnaire delivery, the internet has the potential to revolutionise the survey process. Online (web-based) questionnaires provide several advantages over traditional survey methods in terms of cost, speed, appearance, flexibility, functionality, and usability [1, 2]. For instance, delivery is faster, responses are received more quickly, and data collection can be automated or accelerated [1-3]. Online-questionnaires can also provide many capabilities not found in traditional paper-based questionnaires: they can include pop-up instructions and error messages; they can incorporate links; and it is possible to encode difficult skip patterns making such patterns virtually invisible to respondents.

Like many new technologies, however, online-questionnaires face criticism despite their advantages. Typically, such criticisms focus on the vulnerability of online-questionnaires to the four standard survey error types: namely, coverage, non-response, sampling, and measurement errors. Although, like all survey errors, coverage error ("the result of not allowing all members of the survey population to have an equal or nonzero chance of being sampled for participation in a survey" [2, pg. 9]) also affects traditional survey methods, it is currently exacerbated in online-questionnaires as a result of the digital divide. That said, many developed countries have reported substantial increases in computer and internet access and/or are targeting this as part of their immediate infrastructural development [4, 5]. Indicating that familiarity with information technologies is increasing, these trends suggest that coverage error will rapidly diminish to an acceptable level (for the developed world at least) in the near future, and in so doing, positively reinforce the advantages of online-questionnaire delivery.

The second error type – the non-response error – occurs when individuals fail to respond to the invitation to participate in a survey or abandon a questionnaire before it is completed. Given today's societal trend towards self-administration [2] the former is inevitable, irrespective of delivery mechanism. Conversely, non-response as a consequence of questionnaire abandonment *can* be relatively easily addressed. Unlike traditional questionnaires, the delivery mechanism for online-questionnaires makes estimation of questionnaire length and time required for completion difficult1, thus increasing the likelihood of abandonment. By incorporating a range of features into the design of an online-questionnaire, it is possible to facilitate such estimation – and indeed, to provide respondents with context sensitive assistance during the response process – and thereby reduce abandonment while eliciting feelings of accomplishment [6].

For online-questionnaires, sampling error ("the result of attempting to survey only some, and not all, of the units in the survey population" [2, pg. 9]) can arise when all but a small portion of the anticipated respondent set is alienated (and so fails to respond) as a result of, for example, disregard for varying connection speeds, bandwidth limitations, browser configurations, monitors, hardware, and user requirements during the questionnaire design process. Similarly, measurement errors ("the result of poor question wording or questions being presented in such a way that inaccurate or uninterpretable answers are obtained" [2, pg. 11]) will lead to respondents becoming confused and frustrated.

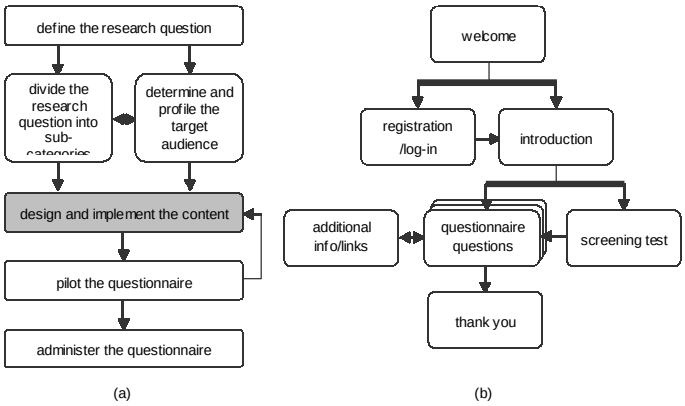
Sampling, measurement, and non-response errors are likely to occur when an online-questionnaire is poorly designed. Individuals will answer questions incorrectly, abandon questionnaires, and may ultimately refuse to participate in future surveys; thus, the benefit of online-questionnaire delivery will not be fully realized. To prevent errors of this kind2, and their consequences, it is extremely important that practical, comprehensive guidelines exist for the design of online-questionnaires. Many design guidelines exist for paper-based questionnaire design (e.g. [7-14]); the same is not true for the design of online-questionnaires [2, 15, 16]. The research presented in this paper is a first attempt to address this discrepancy. Section 2 describes the derivation of a comprehensive set of guidelines for the design of online-questionnaires and briefly (given space restrictions) outlines the essence of the guidelines themselves.

Although online-questionnaires reduce traditional delivery costs (e.g. paper, mail out, and data entry), set up costs can be high given the need to either adopt and acquire training in questionnaire development software or secure the services of a web developer. Neither approach, however, guarantees a *good* questionnaire (often because the person designing the questionnaire lacks relevant knowledge in questionnaire design). Drawing on existing software evaluation techniques [17, 18], we assessed the extent to which current questionnaire development applications support our guidelines; Section 3 describes the framework used for the evaluation, and Section 4 discusses our findings. Finally, Section 5 concludes with a discussion of further work.

COMPREHENSIVE DESIGN GUIDELINES

In essence, an online-questionnaire combines questionnaire-based survey functionality with that of a webpage/site. As such, the design of an online-questionnaire should incorporate principles from both contributing fields. Hence, in order to derive a comprehensive set of guidelines for the design of online-questionnaires, we performed an environmental scan of existing guidelines for paper-based questionnaire design (e.g. [7-14]) and website design, paying particular attention to issues of accessibility and usability (e.g. [19-30]). Additionally, we reviewed the scarce existing provision of online-questionnaire design guidelines [2, 15, 16]. Principal amongst the latter is the work of Dillman [2]. Expanding on his successful *Total Design Method* for mail and telephone surveys [31], Dillman introduced, as part of his *Tailored Design Method* [2], fourteen

Figure 1. Online-Questionnaire (a) Design Process and (b) Organizational Structure (arrows show progression, a double-barred arrow indicating choice in the structure)



additional guidelines specifically aimed at directing the design of online-questionnaires. Albeit seminal, Dillman's guidelines do not incorporate much of the relevant guidance uncovered as part of our environmental scan. We therefore propose – after collating, filtering, and integrating the disparate guidelines – a comprehensive set of guidelines for online-questionnaire design that, although stemming from Dillman's guidelines, are more encompassing. Unfortunately, given space limitations imposed on this paper, it is only possible to highlight the key elements of the derived guidelines; more detail is, however, available on request.

Table 1. Guideline Organization and Topics Covered

General Organization	Formatting	Question Type & Phrasing	General Technical Issues
Welcome Page	*Text*	*General Guidance*	*Privacy & Protection*
Registration/Login Page	*Color*	Sensitive Questions	*Computer Literacy*
Introduction Page	*Graphics*	Attitude Statements	*Automation*
Screening Test Page	*Flash*	Phraseology	*Platforms & Browsers*
Questionnaire Questions	*Tables & Frames*	*Types of Question*	*Devices*
Additional Information Links	*Feedback*	Open-Ended	*Assistive Technology*
Thank You	*Miscellaneous*	Closed-Ended	
Layout	*Response Formats*	Rank-Order	
Frames	Matrix Questions	Categorical or Nominal	
Forms & Fields	Drop-Down Boxes	Magnitude Estimate	
Navigation	Radio Buttons	Ordinal Questions	
Buttons	Check Boxes	Likert Scale	
Links		Skip	
Site Maps			
Scrolling			

Table 2. Excerpt from the Online-Questionnaire Design Guidelines

There are a number of issues of importance when designing the textual content of an online-questionnaire:

a) Fonts used should be readable and familiar, and text should be presented in mixed case or standard sentence formatting; upper case (or all capitals) should only be used for emphasis;

b) Sentences should not exceed 20 words, and should be presented with no more than 75 characters per line. If elderly respondents are anticipated, then this limit should be reduced to between 50 and 65 characters per line. Paragraphs should not exceed 5 sentences in length;

c) Technical instructions (those being instructions related to the basic technical operation of the website delivering the questionnaire) should be written in such a way that non-technical people can understand them;

d) Ensure that questions are easily distinguishable, in terms of formatting, from instructions and answers;

e) For each question type, be consistent in terms of the visual appearance of all instances of that type and the associated instructions concerning how they are to be answered. In particular, keep the relative position of the question and answer consistent throughout the questionnaire. Where different types of questions are to be included in the same questionnaire, each question type should have a unique visual appearance;

f) When designing for access by users with disabilities and the elderly, employ a minimum of size 12pt font and ensure that the font color contrasts significantly with the background coloring. Text should be discernable even without the use of color. It is advisable to test font colors and size with a screen magnifier to ensure usability prior to release;

g) If targeting an elderly audience, provide a text-sizing option on each page, use bold face but avoid italics, and left-justify text. It is also advisable to increase the spacing between lines of text for ease of reading by this respondent group;

h) Make sure that text is read (by screen readers) in a logical order. Specifically, set the tab order on the pages. This is especially true for actual questions in the questionnaire – think carefully about the order in which a visually impaired user will hear the elements of a question, including the instructions and response options.

Guideline Organisation

At the highest level, our guidelines advise on the process that should be followed when designing a questionnaire (the sequence of steps is shown in Figure 1(a)). Providing brief assistance for the remaining steps, the guidelines focus on supporting the design and implementation of questionnaire content (step shown shaded in Figure 1(a)). To this end, we identify the general organisational structure that online-questionnaires should adopt (see Figure 1(b)), provide assistance at this level, and then progressively refine the guidance according to the issues identified in Table 1.

Since it is not possible to include the comprehensive set of guidelines, the following excerpt (Table 2) is presented as an example to provide a 'flavour' for the guidelines as a whole; the guidance relates to the formatting of text (see outlined component in Table 1) in online-questionnaires.

When reading the example, it is important to note that none of the guidelines are particularly innovative in their own right; each has been drawn from the aforementioned sources covered by the environmental scan. What *is* novel, however, is the fact that applicable guidelines from these disparate sources have been collated into a unified set which is presented methodically in order to comprehensively support online-questionnaire design.

A FRAMEWORK FOR EVALUATION OF SUPPORT

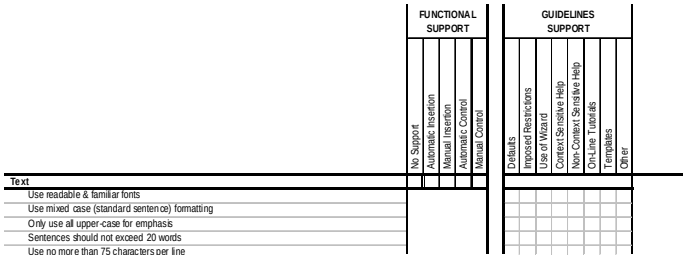
Choice of online-questionnaire development tool is complex. Developers of online-questionnaires are confronted by an increasing variety of software tools to help compose and deliver online-questionnaires. Many such tools purport to allow 'anyone' to quickly and easily develop an online-questionnaire. We wanted to assess the degree to which such tools encourage 'anyone' to develop a *good* questionnaire (where, for our purposes, 'good' is defined as following established principles for website and questionnaire design); that is, we wanted to evaluate the extent to which online-questionnaire development tools incorporate the principles of our guidelines.

Developed by Lumsden, SUIT is a means by which user interface development tools (UIDTs) can be systematically evaluated and compared [17, 18]. Centring around a framework and evaluation method, SUIT adopts a reference model-based approach to tool evaluation. Although, as published, SUIT is dedicated to UIDT evaluation, the principles of SUIT are applicable to any artefact evaluation and comparison [17]. Hence, together with the fact that a website – and therefore an online-questionnaire – is essentially a user interface, a version of the SUIT framework (modified to reflect appropriate evaluative parameters) seemed ideal for the evaluation of support for our identified guidelines within current online-questionnaire development tools.

Evaluation Framework

Figure 2 shows an excerpt from the evaluation framework that was used to assess online-questionnaire development tools. Rows in the framework represent the guidelines, each being summarized for brevity and included under a header representing the applicable online-questionnaire component (e.g. 'text' in Figure 2). Tools were evaluated according to their feature provision and specifically the means by which each feature is incorporated into an online-questionnaire. Where a tool did not support a particular feature, this was marked as 'no support'; where a feature was supported, we recorded whether it was incorporated automatically or manually and whether control over the feature (e.g. style) was manual or automatic. In essence, these measures allow us to determine the functionality and locus of control available in

Figure 2. Excerpt from the Evaluation Framework Used



online-questionnaire development tools. Tools were also evaluated according to their support for our guidelines, measured according to the manner in which the guidelines manifested; for instance, for any given guideline, if a tool restricted the use/set up of the associated feature in accordance with the principle of the guideline, support for that guideline was recorded as 'Imposed Restrictions'. Support mechanisms were not mutually exclusive; it was possible for any given guideline to be supported by more than one means (the available options are shown in Figure 2).

RESULTS AND DISCUSSION

Fifteen online-questionnaire development tools were randomly selected for inclusion in this study; seven were web-based software products (online-tools), typically also hosting the finished survey, and eight were offline software products, installed on one's own computer (offline-tools). A combination of demo software, free online accounts, and vendor tutorials was used to source the information for the study3.

Functional Support for Listed Features

On average, 74% of listed features were supported within the tools studied4; this did not differ between online- and offline-tools although the supported subset did vary slightly across the tool types. In terms of the *General Organization* related features (see Table 1), none of the tools explicitly5 supported the inclusion of screening-test pages or sitemaps, and offline-tools were not found to support development of registration/login pages. *Formatting* features (see Table 1) were better supported, with only flash missing from offline-tools and tables and frames missing from online-tools. All features related to *Question Type & Phrasing* (see Table 1) were supported irrespective of tool type. In terms of *General Technical Issues* (see Table 1), no tools supported design for assistive technology and offline-tools, surprisingly, did not support rigorous testing across platforms and browsers.

Across those that were provided, feature inclusion was achieved manually, on average, 74% of the time; this figure was slightly higher for offline-tools (80%) and slightly lower (68%) for online-tools. A similar pattern was also observed for the control of features once included (on average, 78% of control was manual). In essence, feature insertion style mirrored manipulation style. What is interesting to note here is that despite providing a distinct lack of guidance (see Section 4.2), the tools supported little automation of online-questionnaire design which could have been used in lieu of guidance to 'control' questionnaire quality to some extent.

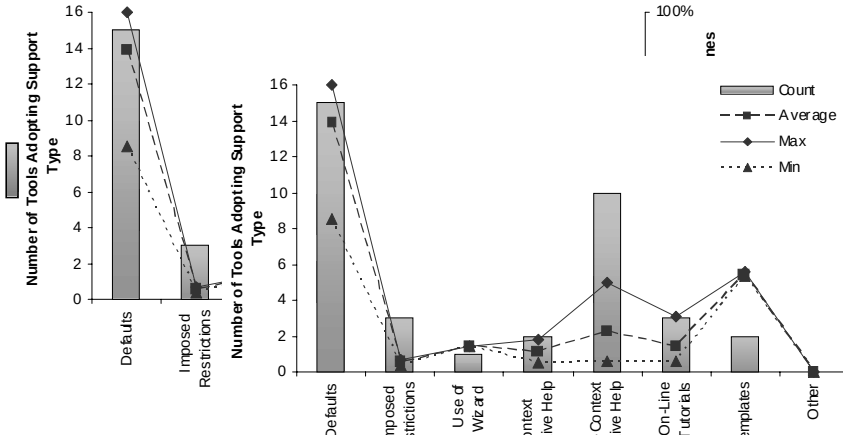
Support for Guidelines

Guideline support was only assessed relative to the features or functions that were physically present/provided within the tools. On average, only 13% of the listed guidelines (relating to supported functionality) had any form of support within the tools studied; this was true for both tool types. Of the guideline *categories* listed in Table 1, 36% had no representation at all across one or both of the tool types. Support for guidelines across 18% of categories was completely missing (where the functionality *was* available) from every tool studied; these included guidelines related to the use or design of *Additional Information Links*, *Navigation*, *Scrolling*, *Matrix Questions*, *Attitude Statements*, *Magnitude Estimate Questions*, and *Automation* of online-questionnaire components. On average, 12% of *General Organization*, 16% of *Formatting*, 6% of *Question Type & Phrasing*, and 26% of *General Technical Issues* guidelines were supported; where there was a difference (albeit, in most cases, very little) between the tool types in terms of extent of support for these high-level categories of guidelines, offline-tools generally provided more support with the exception of *General Organization*, where online-tools were more supportive.

Consider, now, the means by which the supported guidelines *were* supported. Figure 3 shows (using the primary y-axis and bar chart) the number of tools in which each support mechanism was used; using the secondary y-axis and line charts, Figure 3 shows the average, minimum, and maximum extent (as a percentage) to which the various mechanisms were used across supported guidelines. The most popular support mechanism was the use of *defaults* (used, on average, for 87% of supported guidelines). Thus, when a feature was included in an online-questionnaire, it was set up by default in adherence with the associated guideline(s); designers were, however, typically free to alter these settings without being advised of the potential disadvantages associated with their actions. Second in popularity was the use of non-context sensitive help (i.e. help which was not context-linked to actions and had to be looked up independently); it was used in 10 of the 15 tools, but its average application across supported guidelines was only 14%. The remaining support mechanisms were typically used by only one or two tools and contributed to the support of very few guidelines overall. Surprisingly, given the nature of the artifact being designed, neither wizards nor templates were much utilized; where the latter were used, they supported, on average, 34% of guidelines.

Overall, the study has highlighted the predominant absence of sufficient guidance when creating online-questionnaires using current development tools. Typically, most available features can be incorporated into an online-questionnaire with little or no suggestion as to best-practice; where guidelines are supported, the mechanism by which they are supported is typically implicit – there is insufficient explicit explanation provided as to how best to design an online-questionnaire.

Figure 3. Nature of Guideline Support (The primary y-axis and bar chart show the number of tools in which each support mechanism was used; the secondary y-axis and line charts show the average, minimum, and maximum extent (as a percentage) to which the various mechanisms were used across supported guidelines (note: the sum of percentages exceeds 100% since more than one support type was sometimes used per guideline).)



CONCLUSIONS AND FURTHER WORK

On the basis of our evaluation, we consider there to be a distinct need for improved support for guidelines within online-questionnaire design tools in order to facilitate online-questionnaire development that is based on *all* relevant accepted principles. Without such support, *ad hoc* online-questionnaire development will continue as is, with the anticipated result that the public will become disenchanted with such surveys, and their usefulness will therefore diminish without having been granted a fair hearing.

We are currently in the process of performing an empirical study by which we are evaluating the guidelines themselves in terms of their ability, as a comprehensive tool, to guide *better* online-questionnaire design. Ultimately, we plan to develop an online-questionnaire design tool that will *guide* a developer through the design process, highlighting contravention of advisable practice where applicable. Finally, we plan to incorporate additional new guidelines concerning the use of language in online-questionnaires. A 'structurally sound' questionnaire can be badly disadvantaged by the wording used to express questions and responses; we would like to be able to advise on the use of language, primarily via some form of natural language 'check' in our development tool.

REFERENCES

[1] N. Kwak and B. Radler, "A Comparison Between Mail and Web Surveys: Response Pattern, Respondent Profile, and Data Quality," *Journal of Official Statistics*, **18**(2), pp. 257 - 274, 2002.

[2] D. A. Dillman, *Mail and Internet Surveys: The Tailored Design Method*, Second ed. New York, NY: John Wiley & Sons, Inc., 2000.

[3] W. Bandilla, M. Bosnjak, and P. Altdorfer, "Self Administration Effects? A Comparison of Web-Based and Traditional Written Self-Administered Surveys Using the ISSP Environment Module," *Social Science Computer Review*, **21**(2), pp. 235 - 243, 2003.

[4] OECD, *Organization for Economic Co-operation and Development ITC Database*, Accessed: 2nd July 2003, http://www.oecd.org/EN/statistics/0,,EN-statistics-13-nodirectorate-no-no—13,00.html, 2002.

[5] OECD, *Bridging the "Digital Divide": Issues and Policies in OECD Countries*, Accessed: 3rd July 2003, www.oecd.org, 2001.

[6] S. D. Crawford, M. P. Couper, and M. J. Lamias, "Web Surveys: Perceptions of Burden," *Social Science Computer Review*, **19**(2), pp. 146 - 162, 2001.

[7] G. Lindgaard, *Usability Testing and System Evaluation*, 1st ed. London: Chapman & Hall Inc., 1994.

[8] A. N. Oppenheim, *Questionnaire Design, Interviewing and Attitude Measurement*. London: Pinter Publishers, 1992.

[9] A. Fink, *How to Ask Survey Questions*, vol. 2. Thousand Oaks, California: Sage Publications, 1995.

[10] American Statistical Association, *American Statistical Association Series: What is a Survey?*, Accessed: 7th July 2003, http://www.amstat.org/sections/srms/brochures/designquest.pdf, 1999.

[11] W. A. Belson, *The Design and Understanding of Survey Questions*. Great Britain: Gower Publishing Co. Ltd., 1981.

[12] W. Jackson, *Research Methods: Rules for Survey Design and Analysis*. Scarborough: Prentice-Hall, 1988.

[13] E. Taylor-Powell, "Questionnaire Design: Asking Questions with a Purpose," University of Wisconsin - Extension, Wisconsin G3658-2, May 1998.

[14] CASRO, *Council of American Survey Research Organization Guidelines for Survey Research Quality*, Accessed: 7th July 2003, http://www.casro.org/guidelines.cfm, 1998.

[15] K. L. Norman, S. Lee, P. Moore, G. C. Murray, W. Rivadeneira, B. K. Smith, and P. Verdines, *Online Survey Design Guide*, Accessed: 17th June 2003, http://lap.umd.edu/survey_design/tools.html, 2003.

[16] M. Schonlau, R. D. Fricker, and M. N. Elliott, *Conducting Research via E-mail and the Web*, Accessed: 16th June 2003, http://www.rand.org/publications/MR/MR1480/, 2001.

[17] J. Lumsden, "Selecting the 'Invisible' User Interface Development Tool," in People and Computers XVI - Memorable Yet Invisible, Human Computer Interaction 2002 (HCI'2002), London, UK, 2002.

[18] J. Lumsden, "Evaluating the Evaluator: Modelling Systematic Data Analysis Strategies for Software Selection," in 2004 Information Resources Management Association Conference - IRMA'2004, New Orleans, USA, 2004.

[19] A. N. Badre, *Shaping Web Usability: Interaction Design in Context*. Boston, MA: Pearson Education Inc., 2002.

[20] K. P. Coyne and J. Nielsen, "Beyond ALT Text: Making the Web Easy to Use for Users with Disabilities," Nielsen Norman Group, Fremont, CA October 2001.

[21] P. J. Lynch and S. Horton, *Web Style Guide*. Italy: Yale Center for Advanced Instructional Media, 1997.

[22] National Cancer Institute, *National Cancer Institute's Research Based Web Design & Usability Guidelines*, Accessed: 10th June 2003, http://usability.gov/guidelines/index.html, 2002.

[23] S. M. Hinton, "From home page to home site: effective Web resource discovery at the ANU," in Seventh International World Wide Web Conference (WWW7), Brisbane, Australia, 1998.

[24] K. P. Coyne and J. Nielsen, "Web Usability for Senior Citizens," Nielsen Norman Group, Fremont, CA April 2002.

[25] R. Kothari and J. Basak, "Perceptually Automated Evaluation of Web Page Layouts," in Eleventh International World Wide Web Conference, Honolulu, Hawaii, 2002.

[26] W3C, *Web Content Accessibility Guidelines 1.0*, Accessed: 8th June 2003, http://www.w3.org/TR/1999/WAI-WEBCONTENT-19990505, 1999.

[27] A. Stover, K. P. Coyne, and J. Nielsen, "Designing Usable Site Maps for Websites," Nielsen Norman Group, Fremont, CA January 2002.

[28] J. Brewer, "How People with Disabilities Use the Web," W3C Working Draft 4th January 2001.

[29] National Institute on Aging and the National Library of Medicine, *Making Your Web Site Senior Friendly*, Accessed: 19th June 2003, http://www.nlm.nih.gov/pubs/checklist.pdf, 2001.

[30] A. Stover and J. Nielsen, "Accessibility and Usability of Flash for Users with Disabilities," Nielsen Norman Group, Fremont, CA October 2002.

[31] D. A. Dillman, *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley-Interscience, 1978.

ENDNOTES

1 In the absence of appropriate measures to address this.

2 Note, this research is not concerned with coverage errors which are orthogonal to good questionnaire design; mixed-mode delivery is suggested as a means to combat such errors.

3 The authors recognize that, as a result of limited access to some tools, some available functionality may have been missed. All results should be considered in light of this caveat.

4 None of the tools supported any of the design process steps (see Figure 1(a)) – other than the 'design and implement content' step – in any way; as such, these functions are not included in any of the following discussion.

5 For the purpose of fairness and simplicity, a tool was only assessed as supporting functionality if the support was explicit; it is recognized that some tools will provide 'hidden' means by which to achieve functional goals but these are not included here.