WEB SURVEY DESIGN AND IMPLEMENTATION:  
BEST PRACTICES FOR EMPIRICAL RESEARCH

Dora Luz Gonzalez-Bañales  
Durango Institute of Technology

Manuel Rodenes Adam  
Polytechnic University of Valencia

ABSTRACT
This paper presents a guideline for web survey design and implementation. The main objective is to suggest a group of recommendations which may guide a researcher through the process of planning, designing, implementing and conducting a web survey for empirical research purposes. The standards for web surveys are still developing the various implementation possibilities making hard to compare and evaluate their effects on participation in issues as behaviour, data quality and all sorts of survey errors, that is why this brief guide does not claim to be an exhaustive compilation of all noteworthy guidelines; it is just a guide of practical recommendations for empirical research.

Key words: survey online, empirical research, guidelines, web survey, E-mail contact

1. INTRODUCTION
The rapid success of Internet surveys as a new way of gathering data from large population subsets has been accompanied by an increasing number of online administered survey types, each one characterised by different objectives and features (GIOVANNINI, 2003).

Web surveys are becoming an important part of the survey industry; in the United States (US), the Esomar association (http://www.esomar.org/) estimates that more than one third of market research is now conducted in the US through online surveys (GANASSALI, 2006). Web surveys present, in many cases, a promising solution, especially in surveying populations with high use of Internet; this is due to the relatively low costs, speed of data collection and easiness of implementation (LOZAR, 2006; SAUNDERS, 2003; GANASSALI, 2006).

Web survey research, like other kind of survey, is much more than simply asking someone a set of questions. It is a multiple-step process with a clearly defined protocol at each step (MEIR, 2006). A survey is a “systematic method for gathering information from (a sample of) entities for the purposes of constructing quantitative descriptors of the attributes of the larger population of which the entities are members”. There are two main types of questionnaires: self-administered and interviewer administered (SAUNDERS, 2003), the first one is related to online and postal delivery; the second one to telephone questionnaire and structured interview. This document considers only the web survey approach (online).

In spite of the importance of web survey, the literature of this subject is more related with the marketing and social science research fields, however not all web design practices for marketing are suitable for empirical research in another fields; the web survey design for empirical research in general (independently of the research field), still seems to be an opportunity area for making contributions, this is one of the
reasons why this document presents a practical guide that proposes some best practices of web survey marketing research adapted to empirical research (mainly in the fields of system and information technology management). It is important to mention that this paper is only a brief guide of web survey design and implementation, and it does not claim to be an exhaustive compilation of all noteworthy guidelines; it is just a guide of practical recommendations for empirical research in the information technology field.

2. MAIN CHARACTERISTICS OF WEB QUESTIONNAIRES

A web questionnaire has specific characteristics, some of them are (SAUNDERS, 2003; KAPLOWITZ, 2004; WANG, 2000):

- **The population’s characteristics for which is suitable** are mainly computer-literate individuals who can be contacted mainly by E-mail, that is: for populations that regularly use Internet
- **The probability of contamination or distortion** of respondent’s answer is low
- **The response rate is variable**: 30% reasonable within organizations (intranets), 10% or lower on Internet
- **Suitable types of question**: the survey and questions must be of interest to respondent; try to avoid complex questions
- **Time taken to complete data collection**: two to six weeks from distribution (dependent on number of follow-ups)
- **In some cases implies web page design and database management**
- **Cost saving** associated with eliminating the printing and mailing of survey instruments (e.g.: printing, distributing, and data analysis preparation expenses)
- **Data are already in an electronic format**
- **Surveys in an electronic format can eliminate data entry errors and in some cases make real time data analysis**

3. STAGES OF WEB SURVEY DESIGN AND IMPLEMENTATION

Survey design and implementation can be divided into five main stages.

1. Planning the survey
2. Writing the questionnaire
3. Designing the web questionnaire
4. Sending invitations
5. Data collection

In order to be successful, a web survey needs to be properly implemented at every step. Problems at any stage can lead to incorrect data collection (MEIR, 2006; KACZMIREK, 2005).

3.1 Planning the survey

Before get started with a web survey design, it is necessary to at least determine: the research objectives, research model, indicators and resources (in terms of money and expertise). No survey project should go very far forward without a clear statement of the research purpose. If the objectives are clear, questions are easier to determine. The following are some recommendations for planning a web survey (KACZMIREK, 2005;
Outline the research: It is important to know which statistics tests will be needed for the analysis, along with what initial relationships and patterns are expected to find. Doing this will help with the data codification.

Establish a project budget: Mapping an action plan helps to determine the scope and size of the survey. Before beginning a web survey design, it is important to ask:
   a) Which method will be used: E-mail or Web-based survey?
   b) How much will the survey cost (design, software licence, web hosting…)?
   c) How long will it take to complete the survey and produce results?

Plan a schedule: It helps to be more efficient with the time used in the web survey design and implementation process. In this point, it is practical to review and/or use existing questionnaires of another empirical research works; it helps to reduce the time used in the web survey design stage.

Define the population: before developing questions and format of a web survey, it is important to identify the population’s characteristics to which the study is focused to. Some aspects to consider are the population’s familiarity with the use of Internet, E-mail and web surveys. Be aware that some respondents may be unable or unwilling to devote much time to a web survey.

Estimate the required sample size: If the sample size is too small, there is a risk of finding no significant results and it will have to start over. On the other hand, if the sample size is too big, the researcher will waste time and resources because a smaller sample size would be sufficient. That is why it is important to determine a reliable sample size.

Select a method of data collection: the researcher needs to determine which method should use to collect data from respondents: E-mail or web survey. Sometimes, it takes a combination of both methods to get the best results. The method you use depends on factors such as:
   a) type of population to be surveyed
   b) method that the researcher uses to get a representative sample
   c) type of questions to be asked (open, multiple choice…)
   d) funds needed and personnel available

Choose a software solution that meets the research needs. Web surveys software helps to avoid the most common pitfalls in data integrity and saves many hours of work, and delivers ready to use datasets (see LOZAR (2006) and websm.org for a software overview, there are also open software solutions, E.g. www.phpsurveyor.org).

Plan how to use the results: one part of the planning stage is thinking on how the researcher will use the data collected. Some aspects to consider are:
   a) type of analysis that researcher is planning to do
   b) what reports will be written
   c) how information will be used
• **Organize the sequence of the survey items** in accordance with the research model. A recommendation is to start and finish with the easiest questions, and placing the most important questions in the middle of the survey. The survey design must be interesting for the respondent (web design, sequence of the questions, and writing), the respondents must feel that they not only are answering a survey, but also they are learning something. If the researcher is going to write questions on his/her own, he/she might consider taking a training course to learn proven methods for question writing.

• **Quality of survey questions**: it is important to remember that the survey primary goal in empirical research is to obtain data, due to it, the importance of the quality of the survey questions is vital, remember: GIGO (garbage in – garbage out) (LICEN, 2006).

• **Length of the survey.** Some contributions recommend an optimal length ranging between 15 and 30 questions for self-administrated questionnaires. Apart from the number of questions, **the number of screens** is an important issue (GANASSALI, 2006).

• **Make the survey as short as possible.** To maximize the response rates and the number of completions: keep the survey short, maximum 20-25 minutes. For a survey to be below 8 minutes, 15 questions are sensible. The objective when creating survey questions is to keep the survey as brief as possible while still gathering all pertinent information.

**Writing the questionnaire**

For web surveys, the questionnaire design is crucial (GANASSALI, 2006). 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. There are two key goals to keep in mind when writing a questionnaire: minimizing measurement error and reducing non-response; the questionnaire should be constructed so that (MEIR, 2006):

a) respondents are motivated to complete it  
b) the questions are all read correctly and thoroughly  
c) respondents understand how to respond to each question or how to skip with clear instructions throughout the document  
d) returning the questionnaire must be an easy and straightforward task

The following are some practices for writing a questionnaire for the web (KACZMIREK, 2005; LICEN, 2006; SAUNDERS, 2003; MEIR, 2006):

• **Search for already developed questionnaires and scales.** Three widely established and accepted quality criteria for questionnaires and tests are: objectivity, reliability and validity. To assess and improve the quality a lot of work has been done in the creation of scales and questionnaires, therefore, it is a good idea to start by searching for questionnaires or scales, which concern the information needs; but the researcher must be careful, not all questions are usable in every situation, so it is important to examine the questions and the situation.

• **Use as few answer types as possible and be consistent.** Each new answer type or new list of categories cause additional burden to the respondent. The possibility for misunderstandings increases when more answer types are
used. The possibility that respondents are not always reading the complete instruction might be problematic, so some will answer according to the scaling of the previous questions. Thus, the researcher must be consistent with the scales. Avoid switching between different scales.

3.2 Designing the web questionnaire

The following are some useful recommendations for web survey design (SAUNDERS, 2003; KACZMIREK, 2005; GIOVANNINI, 2003):

- **Use compatible data entry**: it is recommendable to use web survey software that is compatible with the statistics software that is being used by the researcher (e.g.: Excel, SPSS, MiniTab…). If the programs are not compatible, the researcher might have to expend more time on recapturing data, which increases the chances for data-entry errors.

- **Do not introduce problems in the online questionnaire that would not happen in a paper and pencil questionnaire**. Respondents are used to the possibility of going back to previous questions and change their answers, pause the survey and resume it later without losing all answers (“Save and continue” button).

- **Be user-friendly**: User-friendliness is an important aspect to achieve valid data. The implementation of a web survey should proceed in a way using technical advances without imposing technical complexity on the participants, some recommendations are (KACZMIREK, 2005):
  - Provide an hyperlink in the E-mail invitation, which leads directly to the questionnaire without an extra login procedure
  - Avoid the need to install plug-ins (a program that allows a web browser to display a wider range of content than originally intended. For example: the Flash plug-in allows web browsers to display Flash content)
  - Give participants as much control over the survey process as with an equivalent paper and pencil questionnaire or more
  - Provide answer controls where a response is necessary to proceed (for example: if the answer is no, go to the next question, if the answer is yes display another question(s))

- **Be trustworthy**: If the researcher build up trust, it becomes more likely for respondents to participate in the survey. Use established codes of conduct (e.g.: International Code CCI/ESOMAR), choose an appropriate implementation for the web survey so that participants are in full control of their answers during their participation, also (KACZMIREK, 2005):
  - Empower the participants, so they can change their previous answers and provide comments
  - Disclose the identity research and contact information and give information about the goals of the survey
  - Ensure data integrity and virtual anonymity as much as possible

- **Be explicit**: Use suitable and intelligible language. Being explicit makes the report and data quality more convincing. Thus, disclose the methodology, some recommendations are:
**Provide information about the sample, and the context in which the survey is conducted (topic and institution)**

**Share screen views of the first page (usually the invitation and explanatory page) and/or one example page of the survey. This is the easiest way to demonstrate the look and feel of the survey**

**Inform of the average time it takes to answer the survey**

**Give information about the objective of the survey**

**Use soft controls.** An implementation of “soft” control would allow proceeding with the next question without forcing a valid response. With hard validation controls, e.g. force valid responses, some respondents will quit. Using data collection software can help to save time by streamlining the process and establish soft controls.

**Be careful with drop-down-menus.** Only use drop-down menus and listboxes if respondents know the answer without having to look at all entries (a listbox allows the user to select one or more items from a list contained within a static, multiple line text box).

**Avoid scrolling.** Display only so many questions on one page as fit within the window. Distribute the questions on several pages. This allows the save and continue feature to work, gives the data response times per page and allows soft controls.

**Answer Boxes (□ or .).** Answer boxes positioned on the right of the screen facilitate the response process in web questionnaires (GIOVANNINI, 2003).

**Implement filters where appropriate.** Filters, jumps, skips, conditioning are means to adapt the questionnaire to the response behaviour of each participant. They allow skipping non-appropriate questions depending on the previous answers given and thus reduce the length of a survey to the individual minimum.

**Use a corporate design.** Use header and background (colours and logos) to design the survey in your corporate, university or research group design.

**Run pre-tests with the web survey.** A pre-test is a formal review of a questionnaire and the associated data collection methodology. It is used to discover potential problems and make improvements, so the researcher gets the best possible results. The pre-test population should be similar in characteristics to the population of the real survey. By conducting a pre-test, it is possible to:

a) find problematic questions and get the opportunity to rewrite them
b) estimate the cost of data collection
c) estimate the response rate – and thus the necessary sample size
d) estimate the distribution of key questions (variables)
e) check the matching of the question wording and their answer categories, the spelling and the flow of the questionnaire
f) check for consistent layout
g) test the survey with subjects similar to the final sample
h) check the survey with different browsers, different settings (e.g. font sizes and window sizes), different operating...
different Internet access speed (modem, ADSL, WiFi...)

i) identify questions where respondents might get stuck

- **E-mail list** (database): The E-mail validation helps to identify survey respondents, to eliminate unwanted responses, to eliminate invalid E-mail addresses, and to delete duplicate E-mails. It is important to have a depurated E-mail list before sending invitations, it can avoid that a same survey respondent receives the invitation twice.
- The **questionnaire should start** with a question that applies to everyone, in an easy and interesting way (GIOVANNINI, 2003)

Final recommendation: **redesign the web survey as many times as it is necessary, before placing it online.**

**Sending invitations**

It is important to remember that the respondents’ participation is voluntary; they have the right to decline to answer a question, a set of questions or the entire survey; some useful suggestions for increasing the response rate for an invitation letter are (KACZMIREK, 2005; SAUNDERS, 2003; ESOMAR, 1999; KAPLOWITZ, 2004; PORTER, 2005):

- Be sure that the person who will receive the E-mail is the right one for answering the survey (importance of reaching a specific person as respondent).
- Get an E-mail list-based management (mailing list).
- Combinations of subject lines are useful: include briefly the reason for the E-mail contact (survey), the name of E-mail sender and institution (e.g.: Dora Glez –UPolitecnica Valencia-). Try to use an interesting subject line; the purpose is that the E-mail does not be deleted before it is read.
- Sending the invitation in plain text is a good idea; in general the recommendation is to avoid attachments and HTML, unless the researcher is using a reliable software application for sending E-mail list with HTML format, if it is decided to send the invitation letter in a HTML format it is a good idea to include also the text version (due to not all E-mail clients support or allow E-mails with HTML design).
- Provide an URL (link) which leads directly to the survey.
- Test the URL in different E-mail clients, some E-mail clients could cut the entire URL.
- Use a software application for managing E-mail lists (e.g. PHP List, ConstantContact…), that will allow reducing the risk that the E-mail invitation could be consider as a “spam” message or “junk mail”.
- Use a meaningful subject line and adjust it lightly with every new E-mail to the same respondent. Use personalized E-mails if possible.
- Keep the message simple and concise: who the researcher is, where he/she works (department, organization, university…), which is his/her job position, survey objective, how to participate, the average time that will take to answer the survey, benefits for the participant, link to a scanned image that shows an official document that proves the researcher identity,
participant anonymity and data confidentiality; thank in advance to the participant his/her participation.
- Check the E-mails that were not delivered and analyze why them were returned (bounces: non-existent address, undeliverable, mailbox full, vacation/auto reply, blocked).
- Allow the respondents to unsubscribe the E-mail list survey. The researcher can include at the end of the message a link where the respondents can notify they do not want to participate or receive more invitations.
- Be sure that all the links included on the E-mail message work properly.
- For a higher response rate, it is recommendable to send the survey invitation on Tuesday, Wednesday or Thursday, and in the second or third week of the month

**Data collection**
Research depends upon public confidence that it is being carried out honestly, objectively, without unwelcome intrusion or disadvantage to respondents, and that it is based upon the respondent’s willing cooperation, so data collection is a crucial stage (SAUNDERS, 2003; MEIR, 2006). At this stage the researcher must be aware that web survey systems have a potential problem as with mailed questionnaires: researchers usually have no control over who completes the survey questionnaire (WANG, 2000). In this point, the researcher needs to collect clean, unbiased and up-to-date data in an efficient manner, it means that self-administered web surveys require clear, direct, logical, and validated questions (MEIR, 2006; WANG, 2000).

In the next list there are some recommendations for web survey data collection:

- **Methods of data collection:** There is no best method of web data collection. The most common are: E-mail and web surveys. Surveys by E-mail may have even lower cost than web surveys. Electronic methods, in general, can collect data more rapidly, but can not be used for general population, because the researcher can not assume that everyone has access to E-mail or Internet. Be sure that the population/sample is Internet/E-mail user.
- **Setting up a “codebook”:** During the data preparation and management step, it is useful to set up “codebook” information, which is any variable definition information. This includes variable names, variable formats and descriptive variable labels and value labels (numbers assigned to data, such as “1” is for male, “2” is for female). This information could be stored in a data dictionary.
- **Failure from non-response:** Many web surveys fail to collect data from an appreciable fraction of those contacted. This problem is called “non-response”. Non-response, by itself, should not be a cause for concern, but when non-respondents differ from respondents in relevant ways, the researcher has to account for the bias. Software solutions can help to analyze missing data patterns and account for these non-response questions (see (LOZAR, 2006) and websm.org for a software overview). If the researcher finds that respondents are systematically different than the non-respondents, the researcher can weight his/her results so that the sample matches known population values. Then, the researcher can draw conclusions more carefully due to missing responses.
• **Failure from item non-response:** Another problem the researcher might encounter in the data collection stage is “item non-response”. This can be avoided completely through electronic surveys. The researcher can use software that creates rules that will not allow the respondents to continue if they do not answer a question, but remember this must be “soft” controls.

• **Control the survey:** After two or three weeks of the first invitation to participate, sending a second notification for the people who do not answer the survey is recommendable. In this second invitation the researcher can improve or correct the way the survey was designed and sent the first time, this is the reason why is important to control the survey by numbers:

  o Number of total invitations sent
  o Number of total E-mail bounces (E-mail message that is returned undelivered).
  o Number of visits to the first webpage
  o Number of responses to the survey
  o Average time it takes to answer the survey


4. CONCLUSION
In general, in order to get reliable results from a survey research, the researcher must be able to plan the survey research project, collect data, access and manage the data easily, as well as report relevant results. One of the main keys of successful web survey design is to ensure that the questionnaire is concise, easy-to-understand, short, and helps to enrich the research objectives; in addition to be sure the population/sample has some expertise level with the use of Internet and E-mail it is a success key. In this way it is possible to get valid, reliable and relevant information. Pre-testing questions are the best method to determine whether a question is correct for a specific web survey. Use software solutions, and well-written, tested pre-existing questions as much as possible, especially from surveys done in a specific industry or topic area, helps to design a better web survey. However, no question is usable in every situation, so it is important to examine the questions for every particular survey research. Web survey is not necessarily the best one; it has advantages and disadvantages like other methods.

**References**


---
