Chapter 17

USING THE INTERNET TO COLLECT DATA

Ulf-Dietrich Reips

The Internet is not one single, monolithic medium. It consists of many services with different functions and needs for input, even more so than the medium *telephone* varies between clunky devices made from wood that we sometimes see in old movies, the iPad, smartphones, and Voice over Internet Protocol.[1]

Using the Internet can mean writing and receiving e-mails that may be purely text based or rich in media. It can mean *to surf* using a Web browser on a desktop computer, laptop, or smartphone. Driving a car often automatically means using the Internet, as information about the location of the car is sent to satellites and databases connected via the Internet. The *Internet of Things*, one of the next steps in the Internet revolution, will connect more and more of the world to the Internet. This connection can be made by the average person with the help of services like http://www.touchatag.com. According to the Gartner Group, by the end of 2012, physical sensors will generate 20% of nonvideo Internet traffic: "The extent and diversity of real-time environmental sensing is growing rapidly as our ability to act on and interpret the growing volumes of data to capture valuable information increases" (Plummer et al., 2009).

Universal addressability of things and people (sometimes called the *Semantic Web*; Berners-Lee, Hendler, & Lassila, 2001; World Wide Web Consortium [W3C], 2010) allows Internet-based data collection even about people and things that are not connected to the Internet.[2] Other agents refer to them by sending their location, images, and so on. A case vividly demonstrating this principle is *Google Street View*: Combining location information with panoramic images creates a highly informative and immersive tool to explore the world at a human scale (i.e., with the eyes of a traveler). The future combination of visual location information with the Internet of Things will create an increasingly tightly meshed representation of the world on the Internet.

This chapter shows the major steps in collecting data on the Internet. The first section, Internet-Based Research, narrates the short history of Internet-based data-collection methods in psychological research, describes their characteristics, and presents a systematic overview of the four basic types of methods. Some notions about planning Internet-based research lead to the second section, Generating a Web Experiment. The section describes an example and provides the reader with the opportunity to become active and experience Internet-based data-collection methods by creating and conducting a Web experiment in a step-by-step fashion. The example introduces the important concepts of client-side versus server-side processing and illustrates a number of important techniques. The third section, Pre-Testing, emphasizes the need to take extra care in preparing the materials and procedure and

[1]Examine some example images of phones at http://www.sparkmuseum.com/TELEPHONE.HTM and http://www.lisisoft.com/imglisi/4/Utilities/137414mla_xp_main_thumb.png

[2]From the definition favored by the W3C (2010), you may gather that the semantics of the Semantic Web are not generally agreed upon: "The term 'Semantic Web' refers to W3C's vision of the Web of linked data. Semantic Web technologies enable people to create data stores on the Web, build vocabularies, and write rules for handling data."

DOI: 10.1037/XXXXX.XXX
APA Handbook of Research Methods in Psychology: Vol. 2. Quantitative, Qualitative, Neuropsychological, and Biological, H. Cooper (Editor-in-Chief)

evaluating their usability. Useful procedures in pretesting of Internet-based data collection are introduced, and the section explains how these procedures prevent methodological problems. In the fourth section, Recruitment, the pros and cons of various ways of attracting participants to Internet-based studies are explained, concluding with the use of games as research environments on the Internet. The Data Analysis section explains a number of important issues such as raw data preservation, paradata, inclusion criteria, and technical variance. Furthermore, the section introduces several specific methods, including log file analysis. The concluding section looks at future trends and the continuing evolution of Internet-based methods and their use in behavioral and social research.

INTERNET-BASED RESEARCH

General Issues and History

The first psychological questionnaires on the World Wide Web (WWW) appeared in 1994 following the introduction of interactive form elements in the scripting language HTML that underlies the WWW (Musch & Reips, 2000). Krantz, Ballard, and Scher (1997) and Reips (1997) conducted the first Internet-based experiments in the summer of 1995, and Reips opened the first virtual laboratory for Internet-based experiments in September 1995 (Web Experimental Psychology Lab: http://www.wexlab.eu).[3] Their results had also been presented at the Society for Computers in Psychology (SCiP) conference in Chicago in 1996 (see http://psych.hanover.edu/SCiP/sciprg96.html; also see Smith & Leigh, 1997). Several studies in most areas of Internet-based data collection—surveying, Web experimentation, data mining, and social network analysis online—were presented at the German Online Research conference in Cologne in 1997 (see http://www.gor.de/gor97/abstracts.htm). The number of studies conducted via the Internet has grown exponentially since then.

To find examples of psychological studies archived or currently in progress on the Web, the reader may visit studies linked at the Web Experimental Psychology Lab or at the following websites:

- Web experiment list at http://wexlist.net/ (Reips & Lengler, 2005) (see Figure 17.1)
- Web survey list at http://www.wexlist.net/browse.cfm?action=browse&modus=survey
- Psychological Research on the Net by John Krantz at http://psych.hanover.edu/research/exponnet.html
- Online Social Psychology Studies by Plous at http://www.socialpsychology.org/expts.htm

INSERT: Scott

- Decision Research Center by Michael Birnbaum at http://psych.fullerton.edu/mbirnbaum/decisions/thanks.htm

Types of Internet-Based Research

Generally speaking, there are four types of Internet-based data collection (Reips, 2006). It can take the forms of *nonreactive Internet-based methods, Internet-based surveys and interviews, Internet-based tests,* and*Internet-based experiments.*

INSERT SPACE

Nonreactive Internet-based methods and *data mining* refer to the use and analysis of existing databases and collections of materials on the Internet (e.g., open forum contributions, server log files, or picture repositories). The Internet provides many opportunities for nonreactive data collection. The sheer size of Internet corpora multiplies the specific strengths of this type of methods: Nonmanipulable events can be studied as they happen in natural behavior on the Internet (navigating, searching, selecting, chatting, reading, timing of these behaviors, and so on), facilitating the examination of rare behavioral patterns or integrating behavioral traces in useful ways (e.g., for television program recommendations; Van Aart et al., 2009). Many of these *user behaviors* are stored in server log files or data bases. Thus, log file analysis is an important example of a nonreactive Web-based method (Reips & Stieger, 2004).

Nonreactive methods have a long tradition in psychological research (e.g., see Fritsche & Linneweber, 2006), and they were used early on in the Internet. In 1996 and 1997, Stegbauer and Rausch (2002)

[3]Because Web addresses (URLs) may change, the reader may use a search engine like Google (http://www.google.com) to access the Web pages mentioned in this chapter, if a link does not work. In the present case, typing "Web Experimental Psychology Lab" into the search field will return the link to the laboratory as the first result listed.

FIGURE 17.1. Front page of the web experiment list. Retrieved from http://wexlist.net

studied the communicative behavior among members of several mailing lists. In this early example of the use of nonreactive data on the Internet, the authors were interested in the so-called lurking behavior (i.e., passive membership in mailing lists, newsgroups, forums, and other social Internet services). They analyzed the number and time of postings and the interaction frequencies pertaining to e-mail headers in contributions (without much need for filtering: It certainly helped that *spam* was a rare phenomenon at the time). Several questions regarding the lurking phenomenon could thus be clarified empirically. For example, about 70% of subscribers to mailing lists could be classified as lurkers, and "among the majority of users, lurking is not a transitional phenomenon but a fixed behavior pattern [within the same social space]" (Stegbauer & Rausch, 2002, p. 267). However, the behavioral pattern is specific to a mailing list: The analysis of individuals' contributions to different mailing lists showed a sizeable proportion of people may lurk in one forum but are active in another. "With this result, Stegbauer and Rausch empirically supported the notion of so-called 'weak ties' as a basis for the transfer of knowledge between social spaces" (Reips, 2006, p. 74).

The most widely used services on the Internet are search engines. With billions of searches performed every day, it is obvious that these searches contain much information about many aspects of human life. A simple measure is *search engine count estimates* (SECEs) that you may have seen as an opener of a slide presentation ("Googling *X* returns 1.2 million links"). Janetzko (2008) has shown good quality (objectivity, validity, reliability) for SECEs as estimates of relative importance of searched items. Several search engine providers have recently moved beyond frequencies and subjected their data to higher order algorithms for mining, as the following examples show.

With the knowledge about new search-based prediction services still not having disseminated widely to the population, it is possible to generate relatively manipulation-free (and thus accurate) predictions. For example, Google's Eurovision site (Google Eurovision, 2010) generates a prediction from searches for performers in the yearly Eurovision song contest, which correctly predicted the winner in 2009. Figure 17.2 shows the prediction for the 2010 contest at the time of this writing, May 12, 2010. The reader may visit the Eurovision song

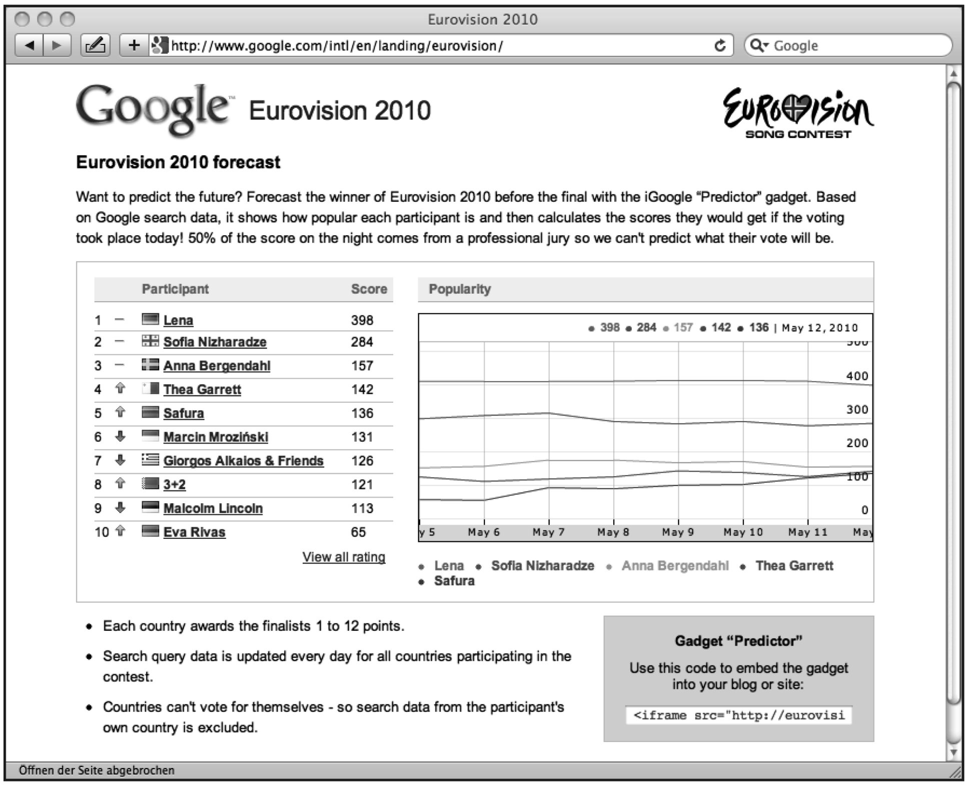


FIGURE 17.2. **Using Google Analytics to predict the results of the Eurovision Contest. As in 2009, the prediction correctly identified the winner. Retrieved from http://www.google.com/intl/en/landing/eurovision**

contest site at http://www.eurovision.tv/page/home to confirm whether the prediction was accurate. Similar approaches from Google Trends (http://www.google.com/trends) and Google Insights (http://www.google.com/insights) have been used to predict a rise in influenza (Ginsberg et al., 2009) and local unemployment rates (Askitas & Zimmermann, 2009). Psychologists working in health care, social services, or advertising may benefit directly from these new intelligent data mining tools. They can get more time for preparing the events predicted for their area, psychological researchers may use the tools to mine searches for certain terms or combinations of terms, in combination with filters like geolocation or time frame (Reips, 2009). One freely available tool that was developed for researchers to mine messages sent on the social networking platform Twitter is iScience Maps (http://tweetminer.eu). Reips and Garaizar (2011), who developed and published the tool, were able to replicate an older study on personality characteristics associated with first names and discover the frequencies of first names as a confound that explains the results of the original study. Using iScience Maps to conduct the study took only about two hours even though the replication was conducted for both the Western part of the United States and the United Kingdom plus Ireland, compared to weeks for the original study, which was conducted in Los Angeles only. Data from websites for psychologists may also be mined to discover mega trends, for example, changes in the topics studied in psychological research (Reips & Lengler, 2005).

Social websites have become valuable sources for social-behavioral research that is based on nonreactive data collection. David Crandall and colleagues from Cornell University (http://www.cs.cornell.edu/~Crandall) created detailed maps by analyzing the location information of approximately 35 million geo-tagged photos that had previously been uploaded to Flickr, a website dedicated to photo sharing. The locations of the motifs show the relative interest in places, and because the sea is always an attractive motive, the shapes of continents appeared on the maps (see Figure 17.3; Barras, 2009). This information may lead to applications in tourism, city planning, ecology, and economics. For example, city planners may trace the location maps over long periods and thus identify areas to be developed or to be made accessible via public transportation.

Internet-based surveys and interviews. The most commonly used Internet-based assessment method is the Web survey. The frequent use of surveys on the Internet can be explained by the apparent ease with which Web questionnaires can be constructed, conducted, and evaluated. Web survey methodology is a difficult matter, however, if one aims at generalizing results from a sample to a particular population. Work by Dillman and his group (Dillman & Bowker, 2001; Dillman, Smyth, & Christian, 2008; Smyth, Dillman, Christian, & Stern, 2006), among others, has shown that many Web surveys are plagued by problems of usability, display, coverage, sampling, nonresponse, or technology. Joinson and Reips (2007) have shown through experiments that the degree of personalization and the power attributable to the sender of an invitation to participate in the survey can affect survey response rates. Data quality can be influenced by degree of anonymity, and this factor as well as information about incentives also influence the frequency of dropout (Frick, Bächtiger, & Reips, 2001). The impression of anonymity on the Internet is particularly helpful in the investigation of sensitive topics. Mangan and Reips (2007) described two Web surveys on the sensitive and rare condition sexsomnia that reached more than five times as many participants from the target population than all nine previous studies from 20 years of research combined.

Design factors like the decision whether to apply a "one screen, one question" procedure may trigger context effects that turn results upside down (Reips, 2002a, 2010). Any aspect of a Web survey that may annoy participants, such as forcing a response, will likely create psychological reactance (Brehm, 1966) and subsequently lead to an increase in random answering behavior, nonresponse, and possibly even dropout (Stieger, Reips, & Voracek, 2007). Despite these findings, converging evidence shows that

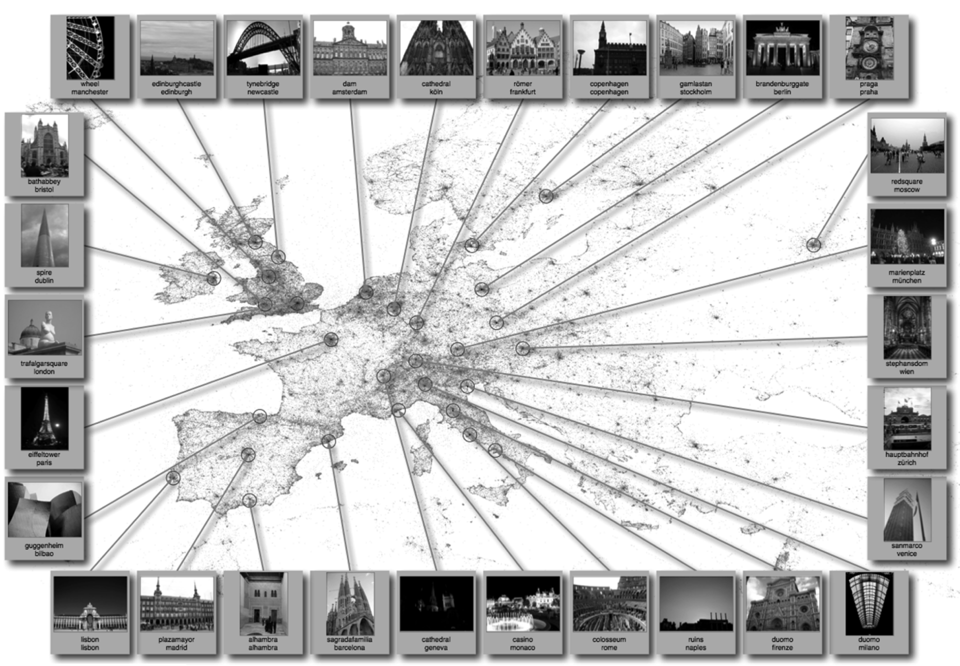


FIGURE 17.3. Map of Europe generated from information embedded in Flickr photographs. From "Gallery: Flickr Users Make Accidental Maps," by G. Barras, April 27, 2009, *New Scientist*. Image created by David Crandall. Repinted with permission. Retrieved from http://www.newscientist.com/article/dn17017-gallery-flickr-user-traces-make-accidental-maps.html

pri

Web-based survey methods result in qualitatively comparable results to traditional surveys, even in longitudinal studies (Hiskey & Troop, 2002). Recently, survey researchers have begun to explore options of mixed-mode surveys (e.g., De Leeuw, 2005; also see the meta analysis by Shih & Fan, 2007). A very good source for research on Web surveying is the Web survey methodology website at http://websm.org.

Web-based psychological testing constitutes one specific subtype of Web surveying, involving psychometric measurement. Buchanan (2001, 2007), Buchanan and Smith (1999), Preckel and Thiemann (2003), and Wilhelm and McKnight (2002), among others, have shown that Web-based testing is possible if the particularities of the Internet situation are considered (e.g., computer anxiety may keep certain people from responding to a Web-based questionnaire) and tests are used that were *validated for use on the Internet*. Buchanan and Smith found that an Internet-based self-monitoring test not only showed similar psychometric properties to its conventional equivalent but also compared favorably as a measure of self-monitoring. Similarly, Buchanan, Johnson, and Goldberg (2005) modified an International Personality Item Pool (IPIP, http://ipip.ori.org/ipip) inventory. In their evaluation it showed to have satisfactory psychometric properties as a brief online measure of the domain constructs of the Five Factor model. Across two studies using different recruitment techniques, they observed acceptable levels of internal reliability and significant correlations with relevant criterion variables. Psychometric equivalence of paper-and-pencil versions of questionnaires with their Web-based counterparts is not always the case, however. For instance, Buchanan et al. (2005) could recover only two of four factor-analytically derived subscales of the Prospective Memory Questionnaire with a sample of $N = 763$ tested via the Internet. Buchanan and Reips (2001) showed that technical aspects of how the Web-based test is implemented may interact with demography or personality and, consequently, introduce a sampling bias. In their study they showed that the average education level was higher in Web-based assessment if no JavaScript was used to code survey and website, and that Mac users scored significantly higher on the personality dimension *openness to experience* (appreciation for art, adventure, curiosity, emotion, unusual ideas, and variety of experience) than Windows users.

Via the iScience server at http://iscience.eu the author of this chapter offers the Five Factor personality test for use on the Internet. Researchers may append the test to their own study by redirecting participants to a study-specific URL. The English and German versions of the test were previously validated for use on the Internet by Buchanan et al. (2005) and Hartig, Jude, and Rauch (2003); validation of the version in Spanish is under way.

Web experiments show several basic differences from experiments conducted in the laboratory or in the field (Reips, 2000, 2002d; Reips & Krantz, 2010). However, the underlying logic is the same as that in the other experimental methods. Hence, the definition of *experiment* used here requires manipulation of the independent variable(s), repeatability, and random assignment to conditions. Likewise, a quasi-Web experiment would involve nonrandom assignment of subjects to conditions (see Campbell & Stanley, 1963; Kirk, 1995). Birnbaum (2007) further discusses representative and systextual experiment designs.

Web experiments offer a chance to validate findings that were acquired using laboratory experiments and field experiments. The number of participants is notoriously small in many traditional studies because researchers set the Type I error probability to a conventional level (and therefore the power of these studies is low; Faul, Erdfelder, Buchner, & Lang, 2009). One of the greatest advantages in Web research is the ease with which large numbers of participants can be reached. The Web Experimental Psychology Lab, for instance, is visited by several thousand people per month (Reips, 2001, 2007). On the Internet the participants may leave a survey at any time, and the experimental situation is usually free of the subjective pressure to stay that is often inherent in experiments conducted for course credit with students. Because Web experiments are often visible on the Internet and remain there as a documentation of the research method and material, overall transparency of the research process is increased.

ADD AUTHOR NAMES TO CHAPTER REFERENCES

Planning an Internet-Based Study

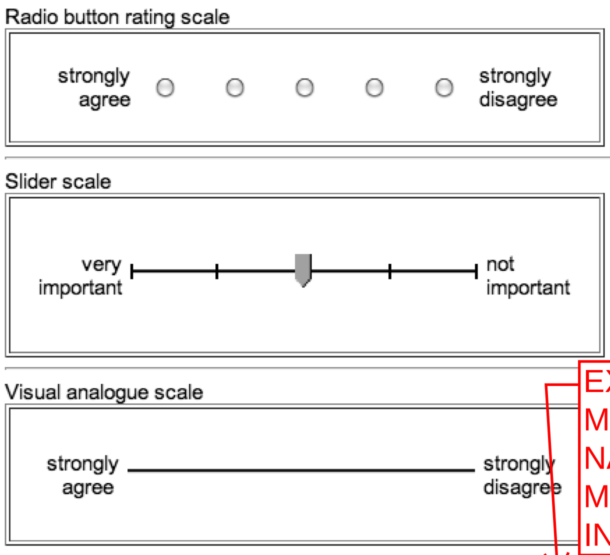
In planning an Internet-based study many of the topics covered in this handbook are of importance. For example, regarding the measurement of psychological constructs (Volume I, Chapters 12 through 31, this handbook) on the Internet we need to consider measurement scales, behavior observation (Volume 1, Chapter 12, this handbook), and use of computers (Volume 1, Chapter 16, this handbook). Regarding psychological tests we need to consider the various subtypes of tests (Volume 1, Chapters 17 through 19, this handbook) and brief instruments and short forms (Volume 1, Chapter 20, this handbook). In the next section I will focus on planning and generating a Web experiment, but many of the issues apply to all types of Internet-based data collection.

GENERATING A WEB EXPERIMENT

Important Techniques

For anyone using the Internet for research purposes it is important to know that there is a growing body of research on newly developed techniques (*tricks*) (e.g., Reips, 2002d, 2007) and the impact of design features of Web-based studies. For example, when choosing a response format for questions one has many options, some of which are not available in paper-and-pencil format (Reips, 2002a, 2010). Figure 17.4 shows a slider scale, a radio button~~s~~ scale, and a visual analogue scale. Funke, Reips, and Thomas (2011) found slider scales to lead to significantly higher break-off rates than radio button~~s~~ scales (odds ratio 6.9) and also to substantially higher response times. Problems with slider scales were especially prevalent for participants with less than average education, suggesting the slider scale format is more challenging in terms of previous knowledge needed or cognitive load. The finding resonates well with the general principle that low-tech solutions are to be preferred for the design of Web-based data collection (Buchanan & Reips, 2001; Reips 2007, 2010).

Reips (2002c) proposed 16 standards or guidelines that may help researchers and reviewers of manuscripts that are based on Internet-mediated research. The proposed standards that refer to techniques are explained in the following sections.



EXCHANGE MARKED SCALE NAMES TO MATCH ORDER IN FIGURE

FIGURE 17.4. From top to bottom: A slider scale, a radio button~~s~~ scale, and a visual analogue scale.

Standard 5: Link your study to other sites. Consider linking your Web study to several Internet sites and services (multiple-site entry technique) to determine effects of self-selection or sampling and estimate generalizability. The multiple-site entry technique is implemented by identifying each source for participants by using a slightly modified link, for example

- http://wextor.org:8080/yourname/yourstudy/index.html?so=clinic1ar (patients from clinic 1—note that two random characters were added after the "1" so as to not allow anyone to change the URL to a real existing link like "_clinic2")
- http://wextor.org:8080/yourname/yourstudy/index.html?so=clinic2gi (patients from clinic 2)
- http://wextor.org:8080/yourname/yourstudy/index.html?so=psstudents (psychology students)
- http://wextor.org:8080/yourname/yourstudy/index.html?so=frfam (friends and family)
- http://wextor.org:8080/yourname/yourstudy/index.html?so=onr (a forum for online researchers)

The resulting data table will contain a column named "so" containing the values "clinic2," "frfam," "onr," and so on, so analyses can be done by source. Following Standard 5, a win-win situation usually results: Either the results are replicated for all sources and samples, thus strengthening the argument that both do not play a role in obtaining these specific results, or some sources and samples may deviate systematically (maybe even theoretically predictable) from the rest—opening up interesting investigations into the reasons for the deviation.

Standards 7–9: Possible strategies to deal with dropout. If dropout is to be avoided, Standard 7 suggests the use of a *warm-up technique*, that is, the actual experimental manipulation only happens several pages deep into the material, after the typical initial spike in dropout, so a high compliance is already established. Any dropout or other nonresponse behavior occurring before the experimental manipulation cannot be attributed to the manipulation. Thus, the results are immune to dropout-related criticisms.

For example, in a Web experiment on list context effects by Reips, Morger, and Maier (reported in Reips, 2002c, 2003a) the target category "number" was experimentally manipulated in a categorizing task only after 31 other items had been categorized to the nonmanipulated categories "flower," "female," and "vegetable," resulting in a dropout of only about 2% (usually, dropout rates are at about 35% in Web experiments; Musch & Reips, 2000).

Standard 8 says to use dropout to determine whether there is *motivational confounding*, that is, the confounding of the motivation to continue participating in the experiment with levels of the independent variable(s). For example, in the case of one experimental condition being more boring than another, a difference in results from the two conditions could simply come from the difference in "boringness" (Reips, 2007).

Standard 9 suggests the use of the *high-hurdle technique*, *incentive information*, and *requests for personal information* to influence time and degree of dropout. The high-hurdle technique seeks to provoke an early drop in participants who likely would have dropped later during a study (Reips, 1997, 2000; but see discussion between Frauendorfer & Reips, 2009, and Göritz & Stieger, 2008). Incentive information is known to have an influence on response rates and dropout (Frick et al., 2001; Göritz, 2006; Musch & Reips, 2000)—a small amount or a chance to win a raffle in the sense of a token of appreciation seems to be the optimal strategy. Requests for personal information at the beginning of a study were shown to increase compliance with study instructions, including a reduction in dropout and other nonresponse (Frick et al., 2001). This is in line with Standard 10.

Standard 10: Ask filter questions (seriousness of participation, expert status, language skills, etc.) at the beginning of the experiment to encourage serious and complete responses. The "seriousness check" (Reips, 2000; Diedenhofen et al., 2010) has become one of the most successful techniques to increase the quality of data in Internet-based data collection. In a simple, albeit very effective strategy, participants are asked at the very beginning of the study whether they "are intending to seriously participate now" or whether they "just want to take a look." Including only the data of those who chose the first option dramatically improves the quality of data. For example, Reips (2005) repeatedly found dropout rates to differ markedly between serious (about 15%) and other (about 75%) participants. Diedenhofen, Aust, Ullrich, & Musch (2010) found that restricting analyses to serious participants allowed a more valid forecast of election results. Moreover, serious participants answered attitudinal questions in a more consistent manner than other participants.

ADD dropout
AFTER "%" IN
BOTH CASES

Standard 11: Check for obvious naming of files, conditions, and, if applicable, passwords. During the process of creating a study, many researchers tend to use names for files and folders that help them remember the meaning of conditions and files or sequences of screens. However, these names are visible to participants via the browser's location window. Consider for example an URL like http://somesite.edu/psych/survey3/controlcond/page4.html. Participants in this study would be able to jump pages by exchanging the 4 in "page4" for a different number, similarly they could investigate previous surveys by changing "survey3." From "controlcond," they could think they were sent to the control condition (Reips, 2002b). Using Web services for scientists that were constructed with these issues in mind (e.g., WEXTOR) will support the scientist in avoiding such frequent errors, for example, by mixing logical with random sequences of characters.

Standard 12: Consider avoiding multiple submissions by exclusively using participant pools and password techniques. Multiple submissions are a rare phenomenon—who would like to repeatedly fill in a questionnaire after all? Internet scientists observed that repeated submissions mostly happen right after the first submission (Birnbaum, 2000; Reips & Birnbaum, ~~in press~~ 2011). To ensure that each participant joins the study only one time, one can send out e-mails to a predefined list of participants. Each of the e-mails contains a unique URL to the study that works only a single time. Furthermore, by building and maintaining a participant pool (or "online panel," Göritz, 2007) for recurring participation requests to the same pool of people, one exerts a higher degree of control of participation than in open recruitment on the Internet.

Standard 13: Perform consistency checks. For example, items with slightly different wording but identical meaning should be answered in the same way, and participants claiming to have a high educational degree should be of a certain age.

Meta Tagging

Meta tags are information snippets in the headers of Web pages that inform Web browsers, caches, and proxy servers about various issues related to the page. Exhibit 17.1, for example, shows several meta tags suitable for pages following the entry page in Internet-based research studies. The *robot* tag tells search engines not to process the content of the page, so no participant will enter the study via this page. The *pragma* tag tells caches and proxy servers not to save the content (so no outdated content is served). The *expires* tag tells caches and proxy servers to consider the content expired (the date lies in the past), thus it is not stored.

Exhibit 17.1

Use of Meta Tags for Internet-Based Research Studies in Pages Following the Entry Page

```
<HTML>
<HEAD>
<meta name="author" content="Experimenter">
<meta name="robots" content="none">
<meta http-equiv="pragma" content="no-cache">
<meta http-equiv="expires" content="Thursday, 1-Jan-1991
    01:01:01 GMT">

<TITLE> </TITLE>
</HEAD>
(insert the body of the Web page here)
```

Client Versus Server

Reactive data-collection techniques on the Internet can be categorized into *server-side* and *client-side* processing (Schmidt, 2000, 2007).

Server-side processing means that all necessary computing is done at the researcher's Web server, including receiving and sending hypertext transfer protocol (HTTP) requests (communication with participant computers), recording and computing of data, communicating with a database application, writing logs, and dynamically selecting and creating materials that may depend on a user's input. Because dynamic procedures are performed on the server, server-side processing is less subject to platform-dependent issues. Sometimes, however, the server may resemble a bottleneck, causing delays.

Client-side methods distribute most tasks to the processing power of the participants' computers. Therefore, time measurements do not contain error from network traffic and problems with server delays are less likely. Server-side processing relies on the participants' computer configurations however, and thus is subject to issues of technical variance (Schmidt, 2007). Server-side and client-side processing methods can be combined, and they can be used to estimate technical error variance by comparison of measurements.

WEXTOR

WEXTOR (available at http://wextor.org) was developed by Reips and Neuhaus (2002). It is a Web-based tool that can be used to design laboratory and Web experiments in a guided step-by-step process. It dynamically creates the customized Web pages and JavaScript needed for the experimental procedure and provides experimenters with a print-ready visual display of their experimental design. WEXTOR flexibly supports complete and incomplete factorial designs with between-subjects, within-subjects, and quasi-experimental factors as well as mixed designs. It implements server- and client-side response time measurement and includes a content wizard for creating interactive materials as well as dependent measures (graphical scales, multiple-choice items, and so on) on the experiment pages.

Many of the methodological solutions discussed in this chapter were built into WEXTOR. As a Web service, WEXTOR can be used to design and manage experiments from anywhere on the Internet using a login and password combination. For support, there are tutorials, a frequently asked questions page, feedback and bug report forms, and an associated mailing list, all linked at the site. Figure 17.5 shows WEXTOR's entry page. The reader is encouraged to download the step-by-step-tutorial available from http://wextor.org/wextor_docs/WEXTOR_tutorial.pdf and re-create a Web experiment as explained.

The process of creating an experimental design and procedure as a self-contained folder of all materials and scripts for an experiment with WEXTOR involves 10 steps. An experimenter logs on to WEXTOR, clicks on the link to "Create/modify an experimental design," and then enters number and names of within- and between-subjects and quasi-experimental factors and their levels. The experimenter then specifies the number of Web pages (screens) and adds information about the type of design (e.g., complete or incomplete), the assignment to conditions, counterbalancing, and so on. Text, stimuli (for a discussion of stimulus delivery, see Krantz, 2001), and measures of various kinds can be added, including visual analogue scales that were shown to produce better measurements than radio button scales (Funke & Reips, in press; Reips & Funke, 2008). Any object can be integrated with the experimental materials, including, for example, videos or Flash content (for more information on using media in Internet-based research, see Krantz & Williams, 2010).

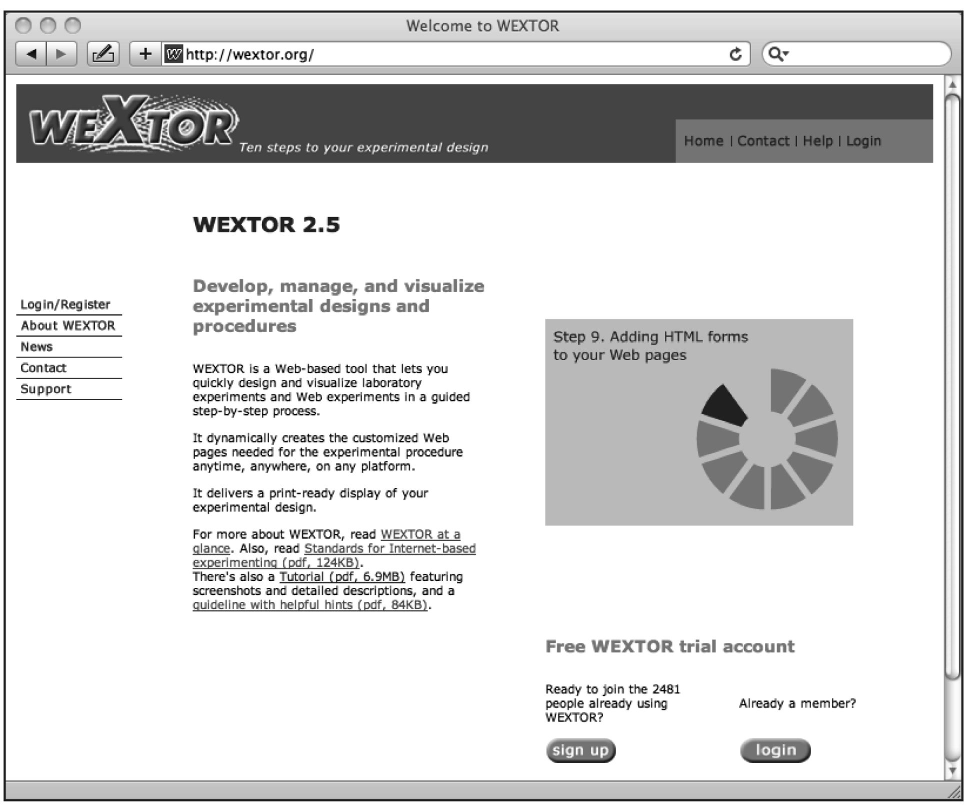


FIGURE 17.5. WEXTOR is a Web service for generating and hosting experiments that can be conducted in the lab or on the Internet. Retrieved from http://wextor.org

WEXTOR produces an organized visual and textual representation of the experimental design and the Web pages and associated JavaScript and cascading style sheet (CSS) files required to implement that design. Furthermore, a code sheet is generated. In the final step, one can then download the experimental materials in a compressed archive that contains all directories (folders), scripts, and Web pages.

After decompressing the archive, the resulting Web pages created in WEXTOR can be viewed and tested for their functionality, even when not connected to the Internet. After further editing in a hypertext markup language (HTML) editor, the whole folder with all experimental materials can be uploaded onto the WEXTOR server or to a personal Web server.

INSERT Then the experiment is ready to go.

Many techniques for Internet-based experimenting mentioned in this chapter were built into WEXTOR to automatically avoid common errors found in Internet-based data collection. Among these techniques are meta tags, the seriousness check technique, the multiple site entry technique, and the high-hurdle technique discussed in the section Important Techniques. Flexible timing out of individual pages and soft form validation can be applied to the Web pages. Skins are available to flexibly change the appearance of all Web pages in the entire experiment at once. Figure 17.6 shows options for the implementation of the high-hurdle technique (Reips, 2000, 2002c), response time measurement, session identification, and form validation in Step 9 in WEXTOR.

For experiments hosted at the WEXTOR website, data preparation can optionally be done on the server, so that data can then be downloaded and opened in spreadsheet programs like Excel or SPSS. The downloaded data file contains a column showing the path taken by each participant (e.g., to see use of the back button), and both server-side and client-side measurements of response time for each page that was accessed.

ADD BORDER TO FIGURE (AS IN FIGURE 17.5)

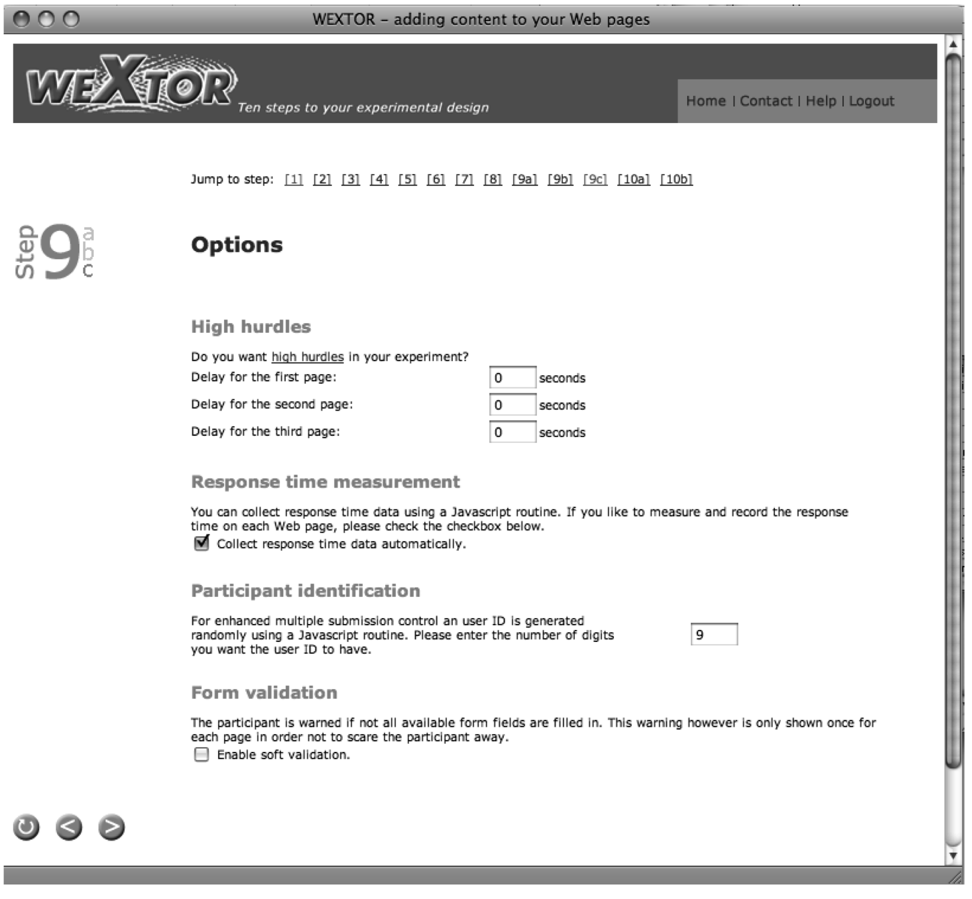


FIGURE 17.6. Step 9c in WEXTOR, showing implementation of the high hurdle, double collection of response times, session ID, and soft validation.

INSERT who are

PRETESTING

Pretesting for an Internet-based study involves gathering information about several aspects that combine to make the study materials accessible and understandable to participants, for example, the usability of the websites involved.

Steps and Procedures

It is best to follow several steps during a pretest phase: First, test your study offline on your own computer and read the ~~participant's~~ form entries in the location window. Second, test the study online and remember to mark the entries or record the time for later exclusion of test data. Third, ask a few colleagues and friends to test the study, and possibly observe some of them to see where they seem to have problems or get stuck. A good procedure is to ask participants to think aloud (speak out what they are thinking, while they are thinking), which often reveals the worst usability problems.

In studies that require a particular sample (e.g., children, participants not native English speakers), pretest a few people from the target population. Finally, check comments and data after the first two dozen people or so participants (do no recruit several hundred or more testers at once).

Preventing Methodological Problems

To prevent running into methodological problems, ask experts for advice (e.g., see the Website of the Society for Computers in Psychology, http://scip.ws). Internet-savvy colleagues may help detect technical issues, whereas those trained in Internet-based research may be able to give advice on designs and procedures.

RECRUITMENT

Portals and Lists

Portals and list sites such as the Web experiment list and others mentioned in this chapter are a good way to recruit people who are interested in taking part in research. These sites are accurately referred to by many other places on the Web as research sites that welcome participants. Over time, they have gained a stable reputation and are used and recommended by many universities around the world.

Mailing Lists, Forums, and Newsgroups

An effective way to recruit participants is to send e-mails to mailing lists, newsgroups, or forums (newsgroups were popular in the early days of the Internet, whereas forums are more of a recent development) of people who, in principle, are open to receiving invitations to participate in studies. Of course, this openness depends on a number of factors, for example, whether the study is related to the topic or membership and clientele of the mailing list, newsgroups, or forum and whether the invitation is endorsed by the moderator. Reips (2003) heard an interesting paper on the first day of a conference and decided to replicate the study overnight. Within eight hours, complete data sets from 162 participants (compared with 64 in the original study) were recorded in the Web experiment, most of which were recruited via three mailing lists of which he is a member. He included the results in his talk on the second day of the conference demonstrate how efficient Internet-based data collection can be in comparison with the methods used by the original authors.

2003b

INSERT to

Participant Pools and Online Panels

A recruitment option for institutions and for researchers who want to follow a long-term strategy in managing one's participants is the *participant pool technique*. People who sign up for this pool provide the Web experimenter with their demographic data and can be paid for participation. The technique thus is attractive for researchers who want to know much about their participants, including who participated in which Web experiments. Consequently, the technique allows for drawing stratified samples (Reips, 2000). Following a naming tradition in survey research, participant pools are in much of the literature on online research now called *online panels*—unfortunately suggesting to reviewers that the participants in the research at hand may have been a small group of experts discussing an issue in a small group meeting. Göritz has published extensively on online panels (Göritz, 2007, 2009; Göritz, Reinhold, & Batinic, 2002).

Search Engines and Banners

Search engines may be used to recruit participants to studies that remain on the Internet for lengthy periods of time, for example, when testing an item pool for the development of a new test, or for the recruitment of people who are interested in a particular phenomenon. Exhibit 17.2 shows how meta tags can be used to inform search engines, in this case, the study entrance page (and it should be only that one—as the researcher does not want to have participants enter a study later). In the example, the study will appear high in searches for "micronations" associated with terms like "ruler," "survey," "head of state," "psychology," and "research,"

The meta tag <META NAME="description" CONTENT=""> informs searches that "research psychologists invite rulers of micronations to complete a survey of experiences that we hope will contribute to understanding the needs of micronations," so this very particular audience will receive an invitation by search, but not many others.

A supplementary option to guide search engines for proper handling of Web-based studies is the Sitemaps protocol that was introduced by Google in 2005 (Sitemaps, 2010). Like the robots.txt file, it allows a webmaster to inform search engines about URLs on a website that are available for crawling. Improving on the robots.txt file, Sitemap allows for the inclusion of additional information about each URL: last update, frequency of changes, and relative importance to other URLs in the site. This information allows search engines to crawl the site more intelligently. According to Sitemaps, "Sitemaps are particularly beneficial on websites where . . . some areas of the website are not available through the browsable interface, or webmasters use rich Ajax, Silverlight, or Flash content that is not normally processed by search engines" (Sitemaps, 2010).

Exhibit 17.2
Use of Meta Tags to Recruit via Search Engine

```
<HTML>
<HEAD>
<META NAME="keywords" CONTENT="micronations,ruler,
    survey,head of state,psychology,research,">
<META NAME="description" CONTENT="research
    psychologists invite rulers of micronations to complete
    a survey of experiences that we hope will contribute to
    understanding the needs of micronations"
<TITLE>Survey of Rulers of Micronations</TITLE>
</HEAD>
<BODY>
(Further information on the study and a link to the study
    would be placed here)
</BODY>
</HTML>
```

Sitemaps containing all accessible URLs on a site can be submitted to search engines. Because the major search engines use the same protocol, having a Sitemap would let these search engines have the updated pages and links information. However, the Sitemaps protocol does not guarantee that web pages will be included in search indexes.

INSERT is

Banners ads were shown to generate only few responses despite their cost and thus this method not recommended (Tuten, Bosnjak, & Bandilla, 2000). Furthermore, any recruitment that takes the form of commercial ads may throw a negative light on the research, even research in general.

INSERT on

Offline Recruitment

One way to recruit participants for Internet-based data collection is often overlooked: traditional media and other offline routes. Simply handing out flyers or printing a study URL in documents accessed by potential participants from the desired population can be very successful. The case of the BBC Internet study (Reimers, 2007) illustrated just how successful recruitment via mostly traditional media like radio and television can be in terms of large numbers: Reimers and the BBC collected data from around 255,000 participants.

Other Sources and Techniques in Recruitment

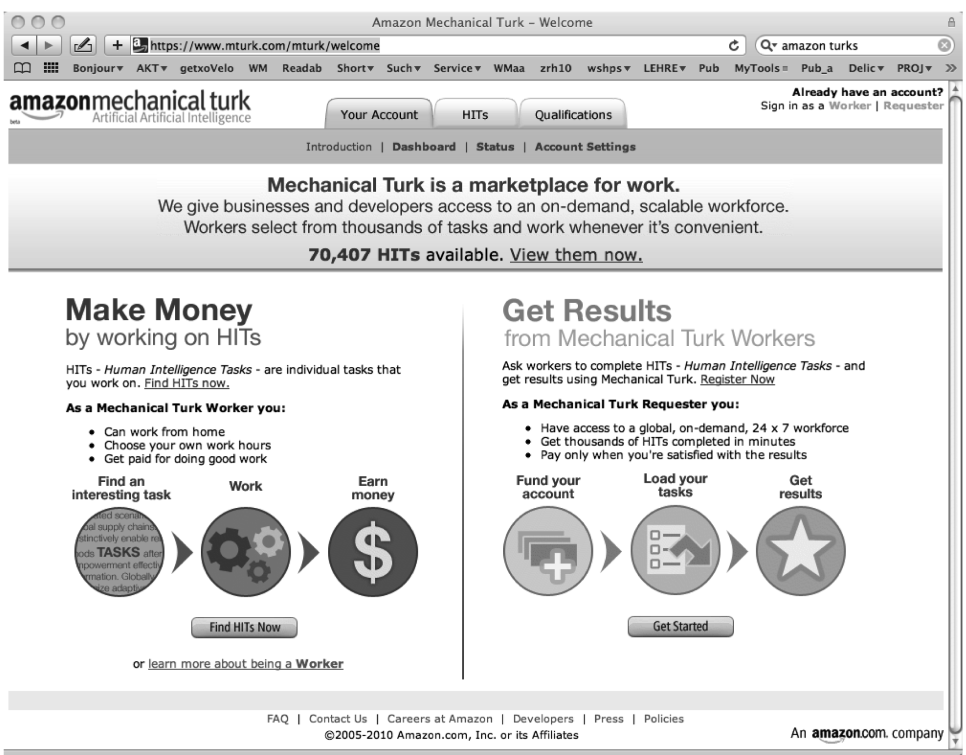
To assess Internet users' privacy concerns Paine, Reips, Stieger, Joinson, and Buchanan (2007) used a *Dynamic Interviewing Programme* (Stieger & Reips, 2008) to survey users of an instant messaging client (ICQ, "I seek you") using both closed and open question formats. Even though their final analysis was conducted only on the basis of data from 530 respondents, the Dynamic Interviewing Programme automatically contacted 79,707 ICQ users who indicated in the instant messaging service that they were open to being contacted.

Figure 17.7 shows Amazon's *mechanical turk*. Mechanical turk is a large job market that equally values and shows job seekers and job openings. Because the jobs can be very small and short in scale, the site is an ideal place to recruit participants and have the payment arranged via Amazon's well-established and trusted services. However, results from a recent study by Reips, Buffardi, and Kuhlmann (2011) show that participants recruited via mechanical turk provide lower quality data than participants from traditional sources for Internet-based data collection: for example, shorter and less variable response times per item (which one would expect to vary widely, if participants think about the items) and more responses to the middle of scales in 50 out of 64 items. At the root of these findings may be that "Mturkers" sign up as workers. Workers respond to be paid, whereas other research participants respond to help with research. A second reason why Mturkers provide lower quality data may be tied to the forums they have established where jobs are discussed, including online studies. It may well be that rumors and experiences shared in these forums lead to a decrease in data quality.

Many potential participants of research studies are more easily recruited if they believe that participating will be fun. Consequently, a way to create Internet-based research is to use game scenarios as covers or design feature. The iScience server at http://iscience.eu includes one link to a Web service named *idex*—on it, researchers and students can create their own stroop-like Web experiments, then invite participants, and later download the data from the site (Figure 17.8).

DATA ANALYSIS

Scientific investigation relies on the principle of preserving raw data. Raw data need to be saved for investigations by other researchers from the community (American Psychological Association, 2010), for example, to aid in deeper understanding of the research, reanalysis, or meta-analysis. This principle applies to Internet-based data collection as well. Raw



ADD BORDER TO FIGURE (AS IN FIGURE 17.5)

FIGURE 17.7. Amazon's Mechanical Turks, an online job market for small- and large-scale jobs that can be used in recruiting and paying participants. Retrieved from https://www.mturk.com/mturk

ADD BORDER TO FIGURE (AS IN FIGURE 17.5)

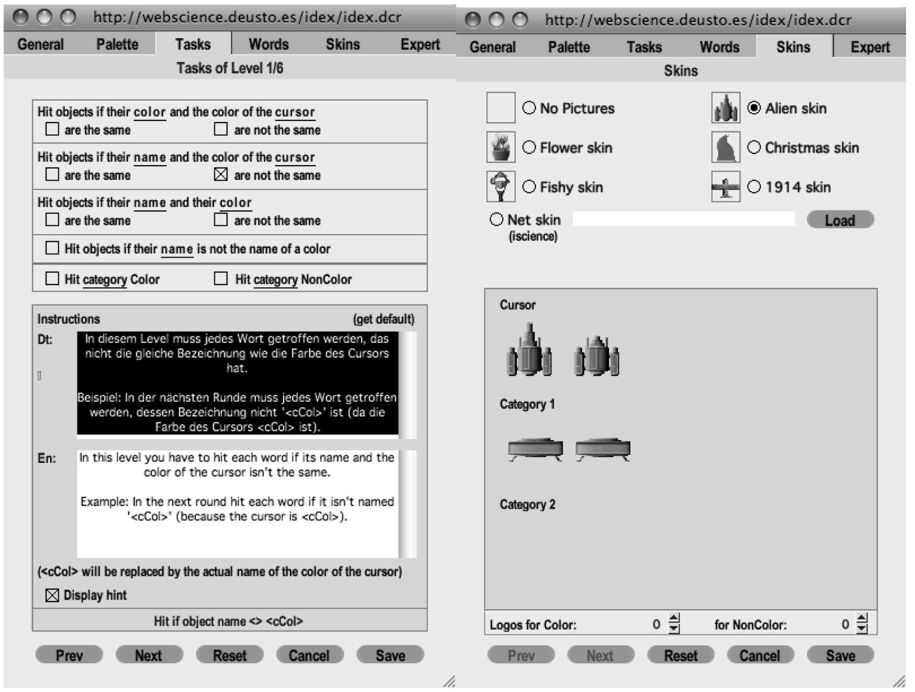


FIGURE 17.8. Two of the panels in *idex* needed to configure stroop-like Web experiments as arcade-style games. Retrieved from http://webscience.deusto.es/idex

data in Internet-based data collection can include results from searches, server log files, or whole databases (Reips, 2007; Reips & Stieger, 2004).

It may be difficult to retain raw data if data collection or parts thereof are delegated to Web services, especially those that may not allow easy download of all data. Figure 17.9 shows Google Analytics, a useful free service for collecting, analyzing, and visualizing statistics about visitors to websites. Collecting data with Google Analytics may be difficult, however, because Google follows a strategy that requires the owner of a Google Analytics account to login frequently to view and save the data.

Paradata

Stieger and Reips (2010) have developed a tool, the *UserActionTracer*, that allows the recording and later reconstruction of paradata, such as timing, clicks, scrolling, and so on, as a reconstructed video sequence. They applied *UserActionTracer* to a large web sample (N = 1,046) in an online questionnaire of average length and found that 10.5% of participants showed more than five single behaviors (out of 132 theoretically possible) that have a clearly negative impact on data quality.

Inclusion Criteria

Data collected on the Internet in their raw form usually contain a number of entries that resemble *noise* and are not to be included in data analysis, for example, accesses from automatic scripts that are used by search engines to assess the content of Web pages. Also, because participants on the Web exert much more freedom to withdraw their participation, criteria should be preset in each Internet-based data-collection effort as to which entries are to be included and which not. For example, a preset inclusion criterion may be that at least half of all items in a questionnaire have to be answered. The following is a typical series of inclusion–exclusion steps in an Internet-based study.

The number of hits is taken of the very first page. Then, all cases are excluded that ended here, that is, no further pages were accessed.[4] This exclusion step may also be defined with a different threshold, for example, the first major point in the study was not

[4]If the meta tags for the exclusion of search engine spiders are used as advised elsewhere in this chapter, then no one will enter the study materials after the first page.

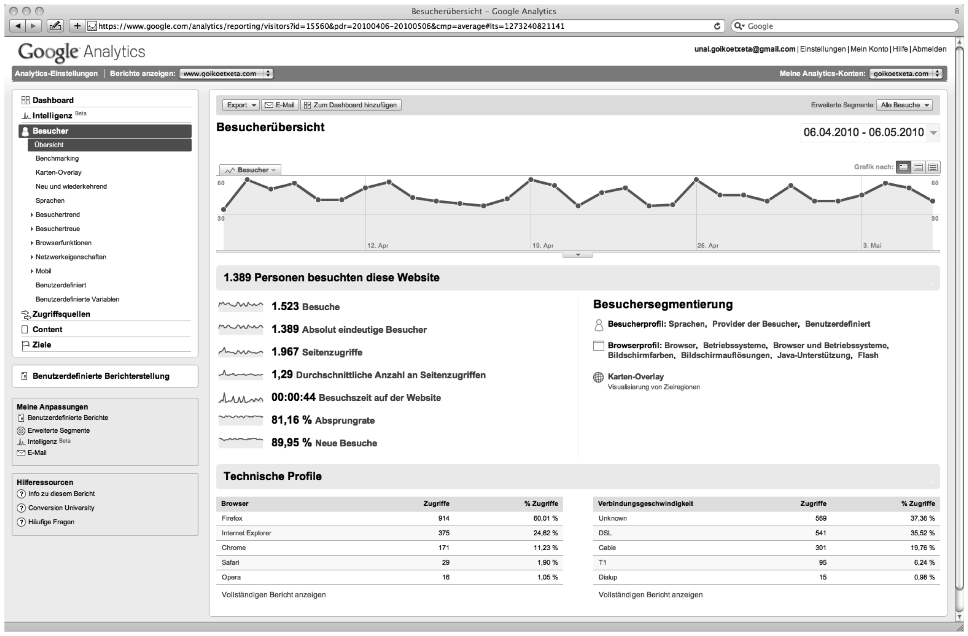


FIGURE 17.9. Google Analytics freely collects, analyzes, and visualizes statistics about visitors to websites.

reached. Such major points can be agreeing to having read an informed consent page, providing demographic information (e.g., so a minimum age can be assumed), distributing to experimental conditions, displaying the instructions, and so on. Next, only those participants are included who answered positively to the seriousness check (possibly also excluding those who did not respond at all to the check). Exclusions to avoid data from multiple submissions can conservatively be determined on the basis of Internet protocol (IP) numbers, even though two or more different participants may be assigned the same number by a large provider. Further criteria for inclusion or exclusion may include the following: whether the browser back button was used, whether a certain proportion of items was answered, whether particularly important items were answered, whether a certain minimum or maximum time was reached, or whether a longer break was taken.

Checking for Technical Variance

In a typical Internet-based study, the participants will access the materials using hundreds of combinations of computers, operating systems, version and type of Web browsers, and network speeds. Information about these can be gathered, many are part of the HTTP (operating systems, version, and type of Web browsers) and others can be computed (network speeds) or gathered via JavaScript (e.g., size and resolution of screen used).

The impact of technical variance can then be checked by analyzing the data via analysis of variance, using the technology parameters as factors. Often, one will find well-established differences, such as different results on personality tests for Mac versus PC users and JavaScript turned on or off depending on level of education (Buchanan & Reips, 2001). It will depend on the aim of the study whether any other differences that may appear are a challenge or an asset to the hypothesis at hand.

CONCLUSION

Even though Internet-based data collection has been around for more than 15 years and is widely accepted, there are still ongoing debates in some fields about whether this new method should be accepted (e.g., Kendall, 2008 versus Honing & Reips, 2008, for the field of musicology). In the meantime, numerous articles appear that report data from Internet-based data collection, even in areas in which the technique is still debated (such as Egermann, Nagel, Altenmüller, & Kopiez, 2009, in the case of musicology research).

The Internet of Things will gain more room and become more of a topic for the general public as well as for researchers over the next years and decades. It will be a question whether and when issues of privacy with various Internet services that recently were discussed more broadly will lead to a compromise between Internet penetration of daily life and the desire and legal requirement of anonymity and privacy.

With this chapter, I hope to have contributed to the spreading of helpful information that counters two popular misconceptions (Reips, 2002c): Internet-based data collection is neither just like offline data collection nor is it completely different from offline data collection.

References

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

Askitas, N., & Zimmermann, K. F. (2009). Google econometrics and unemployment forecasting. *Applied Economics Quarterly*, *55*, 107–120. doi:10.3790/aeq.55.2.107

Barras, G. (2009). Gallery: Flickr users make accidental maps. *New Scientist*. Retrieved from http://www.newscientist.com/article/dn17017-gallery-flickr-user-traces-make-accidental-maps.html

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, *284(6)*, 34–35. doi:10.1038/scientificamerican0501-34

Birnbaum, M. H. (2000). Decision making in the lab and on the Web. In M. H. Birnbaum (Ed.), *Psychological experiments on the internet* (pp. 3–34). San Diego, CA: Academic Press. doi:10.1016/B978-012099980-4/50002-2

Birnbaum, M. H. (2007). Designing online experiments. In A. N. Joinson, K. Y. McKenna, T. Postmes, & U.-D. Reips (Eds.), *The Oxford handbook of Internet psychology* (pp. 391–403). Oxford, England: Oxford University Press.

Brehm, J. W. (1966). *A theory of psychological reactance*. New York, NY: Academic Press.

Buchanan, T. (2001). Online personality assessment. In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet science* (pp. 57–74). Lengerich, Germany: Pabst Science.

Buchanan, T. (2007). Personality testing on the Internet: What we know, and what we do not. In A. N. Joinson, K. Y. McKenna, T. Postmes, & U.-D. Reips (Eds.), *The Oxford handbook of Internet psychology* (pp. 447–459). Oxford, England: Oxford University Press.

Buchanan, T., Johnson, J. A., & Goldberg, L. R. (2005). Implementing a five-factor personality inventory for use on the Internet. *European Journal of Psychological Assessment*, *21*, 115–127. doi:10.1027/1015-5759.21.2.115

Buchanan, T., & Reips, U.-D. (2001, October 10). Platform-dependent biases in Online Research: Do Mac users really think different? In K. J. Jonas, P. Breuer, B. Schauenburg, & M. Boos (Eds.), *Perspectives on Internet research: Concepts and methods*. Retrieved from http://server3.uni-psych.gwdg.de/gor/contrib/buchanan-tom

Buchanan, T., & Smith, J. L. (1999). Research on the Internet: Validation of a World-Wide Web mediated personality scale. *Behavior Research Methods, Instruments, and Computers*, *31*, 565–571. doi:10.3758/BF03200736

Campbell, D., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton, Mifflin.

De Leeuw, E. D. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, *21*, 233–255.

Diedenhofen, B., Aust, F., Ullrich, S., & Musch, J. (2010, May). Seriousness checks are useful to improve data validity in online research. Poster presented at 12th General Online Research Conference, Pforzheim, Germany.

Dillman, D. A., & Bowker, D. K. (2001). The Web questionnaire challenge to survey methodologists. In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet science* (pp. 159–178). Lengerich, Germany: Pabst Science.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2008). *Internet, mail, and mixed-mode surveys: The tailored design method*. New York, NY: Wiley.

Egermann, H., Nagel, F., Altenmüller, E., & Kopiez, R. (2009). Continuous measurement of musically-induced emotion: A Web experiment. *International Journal of Internet Science*, *4*, 4–20.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. doi:10.3758/BRM.41.4.1149

Frauendorfer, D., & Reips, U.-D. (2009, April 6–8). *Investigating the high hurdle technique*. Paper presented at 11th General Online Research conference, Vienna.

Frick, A., Bächtiger, M. T., & Reips, U.-D. (2001). Financial incentives, personal information, and drop-out in online studies. In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet science* (pp. 209–219). Lengerich, Germany: Pabst Science.

Fritsche, I., & Linneweber, V. (2006). Nonreactive methods in psychological research. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 189–203). Washington, DC: American Psychological Association. doi:10.1037/11383-014

Funke, F., & Reips, U.-D. (in press). Why semantic differentials in Web-based research should be made from visual analogue scales and not from 5-point scales. *Field Methods*.

Funke, F., Reips, U.-D., & Thomas, R. K. (2011). Sliders for the smart: Type of rating scale on the Web interacts with educational level. *Social Science Computer Review, 29*, 221–231.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*, 1012–1014. doi:10.1038/nature07634

Google Eurovision. (2010). *Eurovision 2010 forecast*. Retrieved from http://www.google.com/intl/en/landing/eurovision

Göritz, A. S. (2006). Incentives in web studies: Methodological issues and a review. *International Journal of Internet Science, 1*, 58–70.

Göritz, A. S. (2007). Using online panels in psychological research. In A. N. Joinson, K. Y. McKenna, T. Postmes, & U.-D. Reips (Eds.), *The Oxford handbook of Internet psychology* (pp. 473–485). Oxford, England: Oxford University Press.

Göritz, A. S. (2009). Building and managing an online panel with phpPanelAdmin. *Behavior Research Methods*, *41*, 1177–1182. doi:10.3758/BRM.41.4.1177

Göritz, A. S., Reinhold, N., & Batinic, B. (2002). Online panels. In B. Batinic, U.-D. Reips, & M. Bosnjak (Eds.), *Online social sciences* (pp. 27–47). Göttingen, Germany: Hogrefe & Huber.

Göritz, A. S., & Stieger, S. (2008). The high-hurdle technique put to the test: Failure to find evidence that increasing loading times enhances data quality in Web-based studies. *Behavior Research Methods*, *40*, 322–327. doi:10.3758/BRM.40.1.322

Hartig, J., Jude, N., & Rauch, W. (2003). *Entwicklung und Erprobung eines deutschen Big-Five-Fragebogens auf Basis des International Personality Item Pools (IPIP40)* [Development and testing of a German Big Five questionnaire that is based on the International Personality Item Pool (IPIP40)]. Frankfurt, Germany: Arbeiten aus dem Institut für Psychologie der Johann Wolfgang Goethe-Universität.

Hiskey, S., & Troop, N. A. (2002). Online longitudinal survey research: Viability and participation. *Social Science Computer Review*, *20*, 250–259.

Honing, H., & Reips, U.-D. (2008). Web-based versus lab-based studies: A response to Kendall (2008). *Empirical Musicology Review*, *3(2)*, 73–77.

Janetzko, D. (2008). Objectivity, reliability, and validity of search engine count estimates. *International Journal of Internet Science*, *3*, 7–33.

Joinson, A. N., & Reips, U.-D. (2007). Personalized salutation, power of sender, and response rates to Web-based surveys. *Computers in Human Behavior, 23*, 1372–1383.

Kendall, R. (2008). Commentary on "The potential of the Internet for music perception research: A comment on lab-based versus Web-based studies" by Honing & Ladinig. *Empirical Musicology Review*, *3(1)*, 8–10.

Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Krantz, J. H. (2001). Stimulus delivery on the Web: What can be presented when calibration isn't possible? In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet science* (pp. 113–130). Lengerich, Germany: Pabst Science.

Krantz, J. H., Ballard, J., & Scher, J. (1997). Comparing the results of laboratory and World-Wide Web samples on the determinants of female attractiveness. *Behavior Research Methods, Instruments, and Computers*, *29*, 264–269. doi:10.3758/BF03204824

Krantz, J. H., & Williams, J. E. (2010). Using graphics, photographs, and dynamic media. In S. Gosling & J. A. Johnson (Eds.), *Advanced methods for behavioral research on the Internet* (pp. 45–61). Washington, DC: American Psychological Association. doi:10.1037/12076-004

Mangan, M. A., & Reips, U.-D. (2007). Sleep, sex, and the Web: Surveying the difficult-to-reach clinical population suffering from sexsomnia. *Behavior Research Methods*, *39*, 233–236. doi:10.3758/BF03193152

Musch, J., & Reips, U.-D. (2000). A brief history of Web experimenting. In M. H. Birnbaum (Ed.), *Psychological Experiments on the Internet* (pp. 61–87). San Diego, CA: Academic Press. doi:10.1016/B978-012099980-4/50004-6

Paine, C., Reips, U.-D., Stieger, S., Joinson, A., & Buchanan, T. (2007). Internet users' perceptions of "privacy concerns" and "privacy actions." *International Journal of Human-Computer Studies*, *65*, 526–536. doi:10.1016/j.ijhcs.2006.12.001

Plummer, D. C., Hafner, B., Hill, J. B., Redman, P., Brown, R. H., Dulaney, K., . . . Rosser, B. (2009). *Gartner's top predictions for IT organizations and users, 2009 and beyond: Where is the money?* Stamford, CT: Gartner Group. Retrieved from http://www.gartner.com/DisplayDocument?id=874312

Preckel, F., & Thiemann, H. (2003). Online versus paper-pencil version of a high potential intelligence test. *Swiss Journal of Psychology*, *62*, 131–138. doi:10.1024//1421-0185.62.2.131

Reimers, S. (2007). The BBC Internet study: General methodology. *Archives of Sexual Behavior*, *36*, 147–161. doi:10.1007/s10508-006-9143-2

Reips, U.-D. (1997). Das psychologische Experimentieren im Internet. [Psychological experimenting on the Internet]. In B. Batinic (Ed.), *Internet für Psychologen* (pp. 245–265). Göttingen, Germany: Hogrefe.

Reips, U.-D. (2000). The Web experiment method: Advantages, disadvantages, and solutions. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 89–117). San Diego, CA: Academic Press. doi:10.1016/B978-012099980-4/50005-8

Reips, U.-D. (2001). The Web Experimental Psychology Lab: Five years of data collection on the Internet. *Behavior Research Methods, Instruments, and Computers*, *33*, 201–211. doi:10.3758/BF03195366

Reips, U.-D. (2002a). Context effects in Web surveys. In B. Batinic, U.-D. Reips, & M. Bosnjak (Eds.), *Online social sciences* (pp. 69–79). Göttingen, Germany: Hogrefe & Huber.

Reips, U.-D. (2002b). Internet-based psychological experimenting: Five dos and five don'ts. *Social Science Computer Review*, *20*, 241–249.

Reips, U.-D. (2002c). Standards for Internet-based experimenting. *Experimental Psychology*, *49*, 243–256. doi:10.1026//1618-3169.49.4.243

Reips, U.-D. (2002d). Theory and techniques of conducting Web experiments. In B. Batinic, U.-D. Reips, & M. Bosnjak (Eds.), *Online social sciences* (pp. 229–250). Seattle, WA: Hogrefe & Huber.

2003b

Reips, U.-D. (2003, August 25–27). *Seamless from concepts to results: Experimental Internet science*. Paper presented at the symposium "Decision Making and the Web," 19th biannual conference on Subjective Probability, Utility, and Decision Making (SPUDM), Swiss Federal Institute of Technology (ETH), Zurich, Switzerland. Retrieved from http://iscience.deusto.es/archive/reips/SPUDM_03/index.html

Reips, U.-D. (2005, November). *Collecting data in surfer's paradise: Internet-based research yesterday, now, and tomorrow*. Paper presented at the Society for Computers in Psychology (SCiP) conference, Toronto, Ontario, Canada.

Reips, U.-D. (2006). Web-based methods. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 73–85). Washington, DC: American Psychological Association. doi:10.1037/11383-006

Reips, U.-D. (2007). The methodology of Internet-based experiments. In A. N. Joinson, K. Y. McKenna, T. Postmes, & U.-D. Reips (Eds.), *The Oxford handbook of Internet psychology* (pp. 373–390). Oxford, England: Oxford University Press.

Reips, U.-D. (2009). Schöne neue Forschungswelt: Zukunftstrends [Beautiful new world of research: Future trends]. *Nicht-reaktive Erhebungsverfahren* (pp. 129–138). Bonn, Germany: GESIS Schriftenreihe, Band 1.

Reips, U.-D. (2010). Design and formatting in Internet-based research. In S. D. Gosling & J. A. Johnson (Eds.), *Advanced methods for conducting online behavioral research* (pp. 29–43). Washington, DC: American Psychological Association.

Reips, U.-D., & Birnbaum, M. H. ~~(in press~~). Behavioral research and data collection via the Internet. In R. W. Proctor & K.-P. L. Vu (Eds.), *The handbook of human factors in Web design* (2nd ed., pp. 563–585). Mahwah, NJ: Erlbaum. (2011

Reips, U.-D., Buffardi, L., & Kuhlmann, T. (2011, March). *Using Amazon's mechanical turk for the recruitment of participants in Internet-based research*. Paper presented at the 13th General Online Research Meeting, University of Düsseldorf, Germany.

Reips, U.-D., & Funke, F. (2008). Interval level measurement with visual analogue scales in Internet-based research: VAS Generator. *Behavior Research Methods*, *40*, 699–704. doi:10.3758/BRM.40.3.699

Reips, U.-D., & Garaizar, P. (2011). Mining Twitter: Microblogging as a source for psychological wisdom of the crowds. *Behavior Research Methods, 43*, 635–642. doi:10.3758/s13428-011-0116-6

Reips, U.-D., & Krantz, J. H. (2010). Conducting true experiments on the Web. In S. D. Gosling & J. A. Johnson (Eds.), *Advanced methods for conducting online behavioral research* (pp. 193–216). Washington, DC: American Psychological Association.

Reips, U.-D., & Lengler, R. (2005). The Web Experiment List: A Web service for the recruitment of participants and archiving of Internet-based experiments. *Behavior Research Methods*, *37*, 287–292. doi:10.3758/BF03192696

Reips, U.-D., & Neuhaus, C. (2002). WEXTOR: A Web-based tool for generating and visualizing experimental designs and procedures. *Behavior Research Methods, Instruments, and Computers*, *34*, 234–240. doi:10.3758/BF03195449

Reips, U.-D., & Stieger, S. (2004). Scientific LogAnalyzer: A Web-based tool for analyses of server log files in psychological research. *Behavior Research Methods, Instruments, and Computers*, *36*, 304–311. doi:10.3758/BF03195576

Schmidt, W. C. (2000). The server-side of psychology Web experiments. In M. H. Birnbaum (Ed.), *Psychological Experiments on the Internet* (pp. 285–310). San Diego, CA: Academic Press. doi:10.1016/B978-012099980-4/50013-7 e

Schmidt, W. C. (2007). Technical considerations when implementing online research. In A. N. Joinson,

INSERT Reips, U.-D. (2003a). Web-Experimente: Eckpfeiler der Online-Forschung (2. überarbeitete Auflage) [Web experiments: Cornerstones in Online Research (Rev. ed.)]. In A. Theobald, M. Dreyer & T. Starsetzki (Eds.). *Online-Marktforschung: Beiträge aus Wissenschaft und Praxis*. Wiesbaden, Germany: Gabler.

K. Y. McKenna, T. Postmes, & U.-D. Reips (Eds.), *The Oxford handbook of Internet psychology* (pp. 461–472). Oxford, England: Oxford University Press.

Shih, T.-H., & Fan, X. (2007). Response rates and mode preferences in Web-mail mixed-mode surveys: A meta-analysis. *International Journal of Internet Science*, *2*, 59–82.

Sitemaps. (2010). Retrieved May 5, 2010, from http://en.wikipedia.org/w/index.php?title=Sitemaps&oldid=360892548

Smith, M. A., & Leigh, B. (1997). Virtual subjects: Using the Internet as an alternative source of subjects and research environment. *Behavior Research Methods, Instruments, and Computers*, *29*, 496–505. doi:10.3758/BF03210601

Smyth, J. D., Dillman, D. A., Christian, L. M., & Stern, M. J. (2006). Effects of using visual design principles to group response options in Web surveys. *International Journal of Internet Science*, *1*, 6–16.

Stegbauer, C., & Rausch, A. (2002). Lurkers in mailing lists. In B. Batinic, U.-D. Reips, & M. Bosnjak (Eds.), *Online social sciences* (pp. 263–274). Seattle, WA: Hogrefe & Huber.

Stieger, S., & Reips, U.-D. (2008). Dynamic Interviewing Program (DIP): Automatic online interviews via the instant messenger ICQ. *CyberPsychology and Behavior*, *11*, 201–207. doi:10.1089/cpb.2007.0030

Stieger, S., & Reips, U.-D. (2010). What are participants doing while filling in an online questionnaire: A paradata collection tool and an empirical study. *Computers in Human Behavior*, *26*, 1488–1495. doi:10.1016/j.chb.2010.05.013

Stieger, S., Reips, U.-D., & Voracek, M. (2007). Forced-response in online surveys: Bias from reactance and an increase in sex-specific dropout. *Journal of the American Society for Information Science and Technology*, *58*, 1653–1660. doi:10.1002/asi.20651

Tuten, T. L., Bosnjak, M., & Bandilla, W. (2000). Banner-advertised Web surveys. *Marketing Research*, *11(4)*, 17–21.

Van Aart, C., Siebes, R., Buser, V., Aroyo, L., Raimond, Y., Brickley, D., . . . Mostarda, M. (2009, October). *The NoTube beancounter: Aggregating user data for television programme recommendation*. Paper presented at 8th International Semantic Web Conference (ISWC2009), Washington, DC. Retrieved from http://ceur-ws.org/Vol-520/paper01.pdf

Wilhelm, O., & McKnight, P. E. (2002). Ability and achievement testing on the World Wide Web. In B. Batinic, U.-D. Reips, & M. Bosnjak (Eds.), *Online social sciences* (pp. 151–180). Seattle, WA: Hogrefe & Huber.

World Wide Web Consortium. (2010). *Semantic Web*. Retrieved from http://www.w3.org/standards/semanticweb