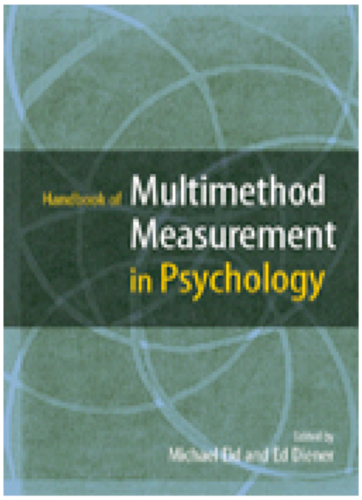
**Preprint of**

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.




Handbook of Multimethod Measurement in Psychology

Editors: Michael Eid, PhD and Ed Diener, PhD

December 2005
Hardcover
736 Pages

---

**Chapter 6:**

# Web-Based Methods

Ulf-Dietrich Reips

Department of Psychology
University of Zürich, Switzerland

To appear in:
Michael Eid & Ed Diener (Eds.),
Handbook of Multimethod Measurement in Psychology

Washington, DC: American Psychological Association

What can be gained from applying Web-based methods to psychological assessment? In the last decade it has become possible to collect data from participants who are tested via the Internet rather than in the laboratory. Although this type of assessment has inherent limitations stemming from lack of control and observation of conditions, it also has a number of advantages over laboratory research (Birnbaum, 2004; Krantz & Dalal, 2000; Reips, 1995; 1997; 2000; 2002c; Schmidt, 1997). Some of the main advantages are that (a) one can test large numbers of participants very quickly; (b) one can recruit large heterogeneous samples and people with rare characteristics; and (c) the method is more cost-effective in time, space, and labor in comparison with laboratory research.

This chapter comprises seven sections. In the first section, Web-based Methods in Psychology, I briefly look at the short history of Web-based methods in psychological research, describe their characteristics, and present a systematic overview of different types of methods. The second section, Advantages of Web-based Methods, illustrates that Web-based methods promise a great number of benefits to psychological assessment, several of which have been empirically supported or confined to specific conditions. The third section, Common Concerns Regarding Internet-based Studies, presents some typical concerns regarding Web-based research, along with findings and reasons that convincingly soften most concerns. However, the theoretical and empirical work conducted by pioneers in research on Web-based methods has also identified some basic problems and some typical errors. The fourth section, Techniques, demonstrates several techniques to avoid, solve, or alleviate these issues. The fifth section, Three Web-based Assessment Methods explains several specific methods, including log file analysis, using the randomized response technique (RRT) on the Web, and game scenarios as covers for Web experiments. The sixth section, Using Web-based Methods: An Example, gives the reader the opportunity to become active and experience Web-based methods by creating and conducting a Web-based experiment, and subsequently, a log file analysis in a step-by-step fashion. The example used is from Internet-based psychological research on framing effects. It shows how the use of Web-based tools can create a whole new type of research experience in psychology when Web-based methods of assessment are integrated with new communication and presentation modes. The concluding section looks at potential future trends and the continuing evolution of Web-based methods and their use in psychological assessment. The rapid development of Web technology and the spread of knowledge among psychologists regarding its characteristics creates the expectation that Web-based methods will inevitably impact the way psychological assessment is conducted in the future.

## WEB-BASED METHODS IN PSYCHOLOGY

Since the beginning (i.e., when the interactive Web became available with the advent of forms in HTML standard 2.0), this technology has been used in psychological research. The first psychological questionnaires appeared in 1994. 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 in September, 1995 (Web Experimental Psychology Lab: http://wexlab.eu/WebExpPsyLab.html[1]). Studies conducted via the World Wide Web (WWW) have grown exponentially since 1995, when researchers began to take advantage of the new standard for HTML, which allowed for convenient data collection (Musch & Reips, 2000).

To get an overall impression of the kinds of psychological studies currently in progress on the Web, the reader may visit studies linked at the Web Experimental Psychology Lab or at the following Web sites:

- Web experiment list (Reips & Lengler, 2005): http://wexlist.org/

[1] Because Web addresses (URLs) may change, the reader is advised to use a search engine like Google (http://www.google.com/) to access the Web pages mentioned in this chapter. In the present case, typing “Web Experimental Psychology Lab“ into the search field will return the link to the laboratory as the first listed result. The Web Experimental Psychology Lab can also be accessed using the short URL http://tinyurl.com/dwcpx

- Web survey list: http://www.wexlist.net/browse.cfm?action=browse&modus=survey
- Psychological Research on the Net by Krantz: http://psych.hanover.edu/research/exponnet.html
- International Personality Item Pool by Goldberg: http://ipip.ori.org/ipip/
- Online Social Psychology Studies by Plous: http://www.socialpsychology.org/expts.htm
- Decision Research Center by Birnbaum: http://psych.fullerton.edu/mbirnbaum/decisions/thanks.htm

## Types of Web-based Methods

Web-based studies can be categorized as *nonreactive Web-based methods, Web surveys, Web-based tests,* and *Web experiments.*

*Nonreactive Web-based methods* refer to the use and analysis of existing databases and text collections on the Internet (e.g., server log files or newsgroup contributions). The Internet provides an ocean of opportunities for nonreactive data collection. The sheer size of Internet corpora multiplies the specific strengths of this class of methods: Nonmanipulable events can be studied *in natura*, facilitating the examination of rare behavioral patterns. An early example of the use of nonreactive data is the study of communicative behavior among members of several mailing lists, conducted in 1996 and 1997 (at a time when SPAM was a rare phenomenon) by Stegbauer and Rausch (2002). These authors were interested in the so-called "lurking behavior" (i.e., passive membership in mailing lists, newsgroups, and other forums). By analyzing the number and time of postings and the interaction frequencies pertaining to e-mail headers in contributions, Stegbauer and Rausch empirically clarified several questions regarding the lurking phenomenon. For instance, 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]" (p.

267). On the other hand, 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.

The fifth section, Three Web-based Assessment Methods, describes log file analysis as an (important) example of a nonreactive Web-based method. For more examples refer to Nonreactive Methods in Psychological Research (Fritsche and Linneweber, this volume, chap. 14).

*Web surveys:* The most commonly used Web-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 surveys can be constructed, conducted, and evaluated. However, this impression is somewhat fallacious. Work by Dillman and his group (Dillman & Bowker, 2001; Dillman, Tortora, & Bowker, 1998) has shown that many Web surveys are plagued by problems of usability, display, sampling, or technology. Joinson and Reips (2004) 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 impact 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). Design factors like the decision whether a "one screen, one question" procedure is applied or not may trigger context effects that turn results upside down (Reips, 2002a). Despite these findings, converging evidence shows that Web-based survey methods result in qualitatively comparable results to traditional surveys, even in longitudinal studies (Hiskey & Troop, 2002).

*Web-based psychological testing* constitutes one specific subtype of Web surveying (unless an experimental component is part of the design, see Erdfelder & Musch, this volume, chap. 15). Buchanan and Smith (1999), Buchanan (2001), 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). Buchanan and Smith found that an Internet-based self-monitoring test not only showed similar psychometric properties to its conventional equivalent but compared favorably as a measure of self-monitoring. Their results support the notion that Web-based personality assessment is possible. Similarly, Buchanan, Johnson, and Goldberg (2005) showed that a modified International Personality Item Pool (IPIP) inventory they evaluated appears to have satisfactory psychometric properties as a brief online measure of the domain constructs of the Five-Factor Model. Across two studies using different recruiting techniques, they observed acceptable levels of internal reliability and significant correlations with relevant criterion variables. However, the issue of psychometric equivalence of paper-and-pencil versions of questionnaires with their Web-based counterparts is not a simple “all equal”. For instance, Buchanan et al. (2004) could only recover two of four factor-analytically derived subscales of the Prospective Memory Questionnaire with a sample of N = 763 tested via the Internet. The other two subscales were essentially meaningless. 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, and that Mac users scored significantly higher on Openness than PC users.

*Web experiments* show a certain categorical distinctiveness from experiments conducted in the laboratory or in the field (Reips, 1995, 2000). However, the underlying logical criteria are the same as those 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).

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; Erdfelder, Faul, & Buchner, 1996). 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 about 4,000 people per month (Reips, 2001). On the Internet the participants may leave at any time, and the experimental situation is usually free of the social pressure 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.

## ADVANTAGES OF WEB-BASED METHODS

One of the principal reasons why Web-based methods are so popular is the *fundamental asymmetry of accessibility*: What is programmed to be accessible from any Internet-connected place in the world, will surely also be accessible in a university laboratory, but what is programmed to work locally may most likely not be accessible anywhere else. A laboratory experiment, for instance, cannot simply be turned into a Web experiment by connecting the host computer to the Internet. But any Web experiment can also be used in the laboratory. Consequently, it is a good strategy to design a Web-based study, if possible. As demonstrated later in this chapter, however, the ease with which laboratory studies can be connected to the Web when developed with Internet software carries the danger of overlooking the specific methodological requirements of using Web-based methods. The requirements and associated techniques are outlined in the next section of this chapter, however, some primary advantages of Internet-based assessment must first be stressed.

Web-based methods offer various benefits to the researcher (for summaries, see Birnbaum, 2004; Reips, 1995, 2000, 2002c). Main advantages are that (a) one can test large numbers of participants quickly; (b) one can recruit large heterogeneous samples and people with rare characteristics (Schmidt, 1997); and (c) Web-based methods are more cost-effective in time, space, administration, and labor in comparison with laboratory research. Of course, all advantages of computerized assessment methods (see Drasgow & Chuah, this volume, chap. 7) apply to Web-based assessment methods as well. Methodological analyses and studies reveal that Web-based methods are usually valid e.g., Krantz, Ballard, & Scher, 1997; Krantz & Dalal, 2000) and sometimes even generate higher quality data than laboratory studies (Birnbaum, 2001; Buchanan & Smith, 1999; Reips, 2000) and facilitate research in previously inaccessible areas (e.g., Bordia, 1996; Coomber, 1997; Rodgers et al, 2001).

Other benefits of Web-based methods are (d) the ease of access for participants (bringing the experiment to the participant instead of the opposite); (e) the ease of access to participants from different cultures—for instance, Bohner, Danner, Siebler, and Samson (2002) conducted a study in three languages with 440 women from more than nine countries (but see the discussion about the physical and educational *digital divide* in access to Web technology); (f) truly voluntary participation (unless participants are required to visit the Web site); (g) detectability of confounding with motivational aspects of study participation; (h) the better generalizability of findings to the general population (e.g., Brenner, 2002; Horswill & Coster, 2001); (i) the generalizability of findings to more settings and situations because of high external validity—Laugwitz (2001), for instance, was able to show that a color perception effect in software ergonomics persisted despite the large variance of conditions of lighting, monitor calibration, etc. in participants' settings; (j) the avoidance of time constraints; (k) the *simultaneous* participation of very large numbers of participants is possible; (l) the reduction of experimenter effects (even in automated computer-based assessments there is often some kind of personal contact, not so in most Web-based assessments); (m) the reduction of demand characteristics (see Orne, 1962); (n) greater visibility of the research process (Web-based studies can be visited by others, and their links can be published in articles resulting from the research); (o) the access to the number of people who see the announcement link to the study, but decide not to participate; (p) the ease of cross-method comparison—comparing results with results from a sample tested in the laboratory; (q) greater external validity through greater technical variance; and (r) the heightened public control of ethical standards.

These are the reasons why 70% of those who have conducted a Web experiment intend to *certainly* use this method again (with the other 30% who are keeping this option open). “Large number of participants” and “high statistical power” were rated by surveyed researchers who had made the decision to conduct a Web experiment as the two most important benefits (Musch & Reips, 2000).

## COMMON CONCERNS REGARDING INTERNET-BASED STUDIES

Many routinely raised concerns involve the lack of proper sampling and the lack of control in Internet-based studies. There are also issues of coverage, measurement, and nonresponse (Dillman, 2001). According to D. Dillman (personal communication, April 1, 2004) the situation gets worse, partly because of the ever increasing variety of media and differences in access to and knowledge about media. Along with other researchers (e.g., Brenner, 2002; Dillman, 2000), I have continuing concerns about potential problems in both Internet-based and laboratory studies. Many unresolved issues remain in traditional studies, including contaminated student samples, experimenter effects, demand characteristics, motivational confounding, low power, and generalizability (for an extensive discussion see

Reips, 2000), and these issues can be alleviated or even resolved with Web-based methods. Experience has shown initial concerns regarding Web-based methods, like the frequency and detectability of multiple submissions, nonrepresentativeness of Internet users, dishonest or malicious behavior (false responses and "hacking"), are not as problematic as previously considered (Birnbaum, 2004; Birnbaum & Reips, 2005), and the real issues tend to be overlooked (Reips, 2002b; 2002c).

When designing a study one must find a balance between methodological advantages and disadvantages. From a multimethod perspective, the opportunity to validate findings with a new set of methods in a new setting is an exciting one: Design the study for the Web, and for comparison, run a subsample in the traditional way.

## Response Time Measurement

One of the more technical concerns about Web-based methods deals with response or even reaction time measurement. How can these times be accurate if the computer equipment is not standardized and calibrated, and if the response is transferred over a fragile net connection? The simple answer is: The noise is small enough to detect relative differences in a proper design, even with the weaker techniques of Internet-based response time measurement, like JavaScript. Reips, Morger, and Meier (2001) demonstrated in an experiment on the previously established *list context effect* with a Web and a lab condition that an effect is detectable on the Web using JavaScript time measurement. However, for the same number of participants, the power to detect effects is lower on the Web. Fortunately, as mentioned earlier, it is also much easier to recruit many participants on the Web.

One of the ways to measure response times is via JavaScript. Because JavaScript is a "client-side" language (it does not run on the server, but on the participants' computers), depending on the exact JavaScript methods used in the scripts, OS, browser type, browser version, and other software running on the client, there is a probability for variance in timing and technical problems with JavaScript. Accumulating technical interactions with JavaScript can even lead to crashes of browsers and computers (for an experiment showing that using JavaScript in a Web experiment will lead to a 13% higher overall dropout rate compared to the same Web experiment without JavaScript, see Schwarz & Reips, 2001). The likelihood for problems seems to decrease, although, with newer browsers and newer OS versions that obviously adapt well to the problems.

A second crude way of measuring response times is to calculate the time differences of when materials are accessed on the Web server. Scientific LogAnalyzer (Reips & Stieger, 2004, see this chapter: Using Web-based Methods: An Example) includes a routine to calculate these times from servers' log files.

So, is there any way to accurately measure reaction times via the Internet? There is: Eichstaedt (2001) developed a Java-based method for very accurate response time measurements. A clever combination of applets ensures continuous synchronization and calibration of timing between server and client, which minimizes timing inaccuracies produced by the Internet.

## TECHNIQUES

Two types of techniques were developed in Internet-based research. One type guards against common errors and problems, the other one increases the usefulness of Web-based assessment methods. Also, techniques can be grouped, along the stages of the research process, according to their applications: techniques for design and procedure, techniques for recruitment, techniques for data analysis. Many of the techniques have been implemented in those Web services or software that allow the creation of Web-based assessments.

### Techniques Against Common Errors and Problems

Every coin has two sides, and so the great advantage of revealing assessment materials to a large worldwide audience via the Internet also means that the collected information may be accessible for many people. There is evidence that confidential data is often openly accessible (an estimate runs at 25%–33%, and this is a cause for concern) because of configuration errors on the part of the researcher that can be easily made in certain operating systems (Reips, 2002b). Several measures help delete this problem: (a) choosing the right (secure) combination of operating system and Web server, (b) using a pretested system to develop and run the Web-based assessment, (c) having people with good Internet knowledge test the Web-based assessment for security weaknesses.

In dealing with *multiple submissions* that may become a problem in highly motivating study scenarios (see the description of game-based Web experiments in Three Web-based Assessment Methods, this chapter), one can use techniques for *avoiding* and techniques for *controlling* the respondents' behavior (Reips, 2002c). Avoidance of multiple submissions, for instance, can be achieved by limiting participation to members of a group known to the researcher, like a class, an *online participant pool,* or *online panel* (Göritz, Reinhold, & Batinic, 2002) and working with a *password scheme* (Schmidt, 1997). A technique that helps control multiple submissions is the *sub-sampling technique* (Reips, 2000, 2002b): For a limited random sample from all data sets, every possible measure is taken to verify the participants' identity, resulting in an estimate for the total percentage of multiple submissions. This technique can help estimate the number of wrong answers by checking verifiable responses (e.g., age, sex, occupation). Applications for Web-based assessment may include routines that check for internal consistency and search for answering patterns (Gockenbach, Bosnjak, & Göritz, 2004). Overall, it has repeatedly been shown that multiple submissions are rare in Internet-based research (Reips, 1997; Voracek, Stieger, & Gindl, 2001), and that data quality may vary with a number of factors (e.g., whether personal information is requested at the beginning or end of a study, Frick et al., 2001; information about the person who issues the invitation to the study, Joinson & Reips, in press; or whether scripts are used that do not allow participants to leave any items unanswered and, therefore, cause psychological reactance, Reips, 2002c).

## Techniques to Increase the Usefulness of Web-based Assessment

One major asset available in Web-based assessment methods is the information gained from different types of nonresponse behavior (Bosnjak, 2001), particularly dropout (attrition). Dropout is always present in Web-based assessment methods because subjectively the participant is in a much more voluntary setting than in a laboratory situation. Although one may consider dropout a serious problem in any type of study, dropout can also be put to use and turned into a detection device for motivational confounding, i.e. the confounding of the motivation to continue participating in the study with any other difference caused by differing influences between conditions (Reips, 1997, 2000, 2002b; Reips, Morger, and Meier, 2001). If desired, dropout can also be reduced by implementing a number of measures, like promising immediate feedback, giving financial incentives, and by personalization (Frick et al., 2001). Or, the *warm-up technique* for dropout control can be implemented (Reips, 2000, 2002b): the actual study begins several pages deep into the material, so a high compliance is already established.

Only a selection of the available techniques can be explained in this chapter, but the reader is referred to Birnbaum (2001), Birnbaum and Reips (2005), and Reips (2000, 2002b, 2002c, 2002d) for more detailed explanations of these and other techniques of Web-based assessment.

# THREE WEB-BASED ASSESSMENT METHODS

In this section, three specific Web-based methods are presented: log file analysis as an example of a nonreactive method, using the randomized response technique in surveys conducted on the Web, and games as a cover format for Internet-based experiments.

*Log file analysis* is at the core of many nonreactive methods of behavioral research on the Web. Navigation behavior in Web sites can be captured as so-called *click streams*, both on an individual and on a group level. Scientific applications for Web log analysis can be used to extract information about behaviors from log files, calculate response times and nonresponse behavior, and find relevant differences between users' navigation behaviors. The tool STRATDYN (Berendt, 2002; Berendt & Brenstein, 2001), for instance, provides classification and visualization of movement sequences in Web navigation and tests differences between navigation patterns in hypertexts. Scientific LogAnalyzer (Reips & Stieger, 2004) is geared toward analyzing data provided on forms and was developed for the analysis of data from most types of Internet-based experimenting (the section, Using Web-based Methods: An Example, contains a description of how to use Scientific LogAnalyzer). LOGPAT (Richter, Naumann, & Noller, 2003) is useful in analyzing sequential measures, (i.e., counting the frequency of specific paths or path types in a log file). Like Scientific LogAnalyzer, LOGPAT was developed as a platform-independent, Web-based tool. In addition to these scientific applications, a large number of commercial and free log file analysis programs are available primarily focus on helping the user maintain a Web site. This type of software can help identify access errors, points of entry, and user paths through a Web site. Many of the applications are user friendly and create visually appealing graphical output. Example programs are Analog (http://www.analog.cx/), FunnelWeb (http://www.quest.com/funnel_web/analyzer/), TrafficReport (http://www.seacloak.com/), and Summary (http://www.summary.net/).

Testing large numbers of participants very quickly via the Web is particularly important for the success of research projects that depend on the availability of a large sample. Therefore, a Web-based format is always a good choice if the *randomized response technique* (RRT, Warner, 1965) is to be used. Researchers have demonstrated the feasibility of the RRT in a large number of studies (e.g., Antonak & Livneh, 1995, for an explanation of the method see Erdfelder and Musch, this volume, chap 15).

One of the better versions of the RRT, the cheater detection model by Clark and Desharnais (1998), which operates with an experimental between-subjects manipulation, has been repeatedly used on the Web (Musch, Bröder, & Klauer, 2001; Reips & Musch, 1999). Figure 6.1 shows a screen capture taken from the Web-based RRT study by Reips and Musch on the feasibility and trustworthiness of a computerized random generator. The participant is asked to click on the random wheel on the left side of the window. A click results in one of two events: If the left portion of the window turns blue then a true answer to the question is requested. If the window turns red, then the participant is asked to answer with “Yes”, independently of the true answer. This condition is compared with one in which a different “random” device independent of computers and the Internet is used: the participant’s month of birth. From various other conditions the behavior’s incidence rate and the proportion of “cheaters” (sic!) in the sample can be calculated, as well as the influence of the computerized “random wheel”. The enhanced anonymity often associated with Web-based questioning has provided additional advantages when conducting RRT surveys on the Internet.

Web experiments designed in game style are likely to attract a very large number of participants who will participate with high motivation (e.g. Ruppertsberg, Givaty, Van Veen, & Bülthoff, 2001; Reips & Mürner, 2004). Ruppertsberg et al. (2001) used games written in Java as research tools for visual perception over the Internet. They concluded that presenting games “… on the Internet resulted in large quantities of useful data, and allowed us to draw conclusions about mechanisms in face recognition in a broader, less selected participant population” (p. 157).

Reips and Mürner (2005) recently developed a Web site that allows researchers and students to develop their own Web-based Stroop experiments in an arcade game style. This Web site is available at http://www.psychologie.unizh.ch/sowi/reips/stroop/. The researcher can configure many aspects of the Stroop paradigm, like colors and names of objects, rules for events, rates for the different event types, speed, and the overall style in which the game is presented (i.e., “skins”). Access to the created Web experiment can be restricted using a login and password. The Web experiment is immediately available online, and the resulting data can be downloaded as tab-delimited text file in a format optimized for analysis in Scientific LogAnalyzer. Figure 6.2 shows the game pad page of “Stroop Invaders”.

## Using Web-based Methods: An Example

Reading about an assessment method can be useful. However, to gain insights on a deeper level and to take concrete steps in acquiring knowledge about the method, it may be even more useful to experience it. Therefore, this section provides the opportunity to create and conduct a Web experiment, in a step-by-step fashion. Along the way, several useful tools for Web-based methods are presented, that is, *WEXTOR* (Reips & Neuhaus, 2002), the *web experiment list* (Reips & Lengler, 2005), the *Web Experimental Psychology Lab* (Reips, 2001), and *Scientific LogAnalyzer* (Reips & Stieger, 2004). A portion of McKenzie and Nelson's (2003) “cup experiment” is recreated for replication on the Web. This study deals with the information implicitly conveyed by the speaker’s choice of a frame—for instance, describing a cup as being “half full” or “half empty”.

### WEXTOR

First, we use WEXTOR (Reips & Neuhaus, 2002), a Web service, to create, store, and visualize experimental designs and procedures for experiments on the Web and in the laboratory. WEXTOR dynamically creates the customized Web pages needed for the experimental procedure. It supports complete and incomplete factorial designs with between-subjects, within-subjects, and quasi-experimental (natural) factors, as well as mixed designs. It implements client-side, response time measurement and contains a content wizard for creating materials and dependent measures (button scales, graphical scales, multiple-choice items, etc.) on the experiment pages.

Several of the techniques presented earlier in this chapter are built into WEXTOR, (e.g., the warm-up and high hurdle techniques), and it automatically avoids several methodological pitfalls in Internet-based research. WEXTOR uses nonobvious file naming, automatic avoidance of page number confounding, JavaScript test redirect functionality to minimize dropout, and randomized distribution of participants to experimental conditions. It also provides for optional assignment to levels of quasi-experimental factors, optional client-side response time measurement, optional implementation of the high hurdle technique for dropout management, randomly generated continuous user IDs for enhanced multiple submission control, and it automatically implements meta tags that keep the materials hidden from search engine scripts and prevents the caching of outdated versions at proxy servers.

The English version of WEXTOR is available at http://wextor.org. WEXTOR is currently available in version 2.2. After going through a sign-up procedure, WEXTOR can be used to design and manage experiments from anywhere on the Internet using a login/password combination. For the purpose of guiding the reader through the process, I created an account in WEXTOR that already contains a complete version of the cup experiment. Readers of this chapter may log in using the login/password combination "APA/handbook". Also, a step-by-step explanation of how to create a Web-based replication of the cup experiment (Reips, 2003) is at http://www.psychologie.unizh.ch/sowi/reips/SPUDM_03/index.html. Figure 6.3 shows WEXTOR's entry page.

The process of creating an experimental design and procedure for an experiment with WEXTOR involves ten steps. The first steps are decisions that an experimenter would make whether using WEXTOR or any other device for generating the experiment, like listing the factors and levels of within- and between-subjects factors, deciding what quasi-experimental factors (if any) to use, and specifying how assignment to conditions will function. WEXTOR adapts to the user input and produces an organized, pictorial representation of the experimental design and the Web pages required to implement that design. Figure 6.4 shows the visualization of the design and procedure for the experiment by McKenzie and Nelson. It is a 2 × 2 between-subjects factorial design, resulting in four experimental conditions. Each condition is represented by one folder containing the Web pages the participants will see in that condition. Every Web page holds the dynamically created scripts that translate into the study procedure and response time measurement. After creating the experimental materials in WEXTOR they can be downloaded in one compressed archive that contains all folders (directories), scripts, and Web pages. WEXTOR contains a description of how to give these pages the “editing finish” and how to configure a Web server to post the pages on the Web (also see Birnbaum & Reips, 2005).

## Recruitment

Once the materials for a Web-based study have been assembled and are available on the Web, the recruitment phase begins. Following traditional recruitment methods, participants can be recruited offline, of course. In addition, there are now many Internet methods (e.g., recruitment via Web site, e-mail [including mailing lists], online panel, newsgroup, listings, and banner ads). Recruitment for Web-based studies can be much more effective with one or several of the techniques described by Birnbaum (2001), Birnbaum and Reips (2005), and Reips (2000, 2002b, 2002c, 2002d).

Some of the best places for recruitment are institutionalized Web sites for Internet-based assessment, like those mentioned at the beginning of this chapter. In the case of Web experiments (e.g., the cup example), the study can be announced on the web experiment list and in the Web Experimental Psychology Lab. Figure 6.5 shows the entry form that an experimenter must fill out to put a Web experiment on the web experiment list.

## Data Analysis

Because of the large numbers of possible participants recruited on the Internet within a short period of time, data analysis can often follow briefly after the recruitment process. In the case of the replication of the cup experiment, I collected 162 data sets within 8 hours (Reips, 2003). Log files contain information in a format of one line per accessed piece of material. However, for a useful statistical analysis, most often a ‘one row per participant’ format is needed. A Web-based service to do this transformation is Scientific LogAnalyzer. Several methodological features specifically needed for the analyses of data collected using Web-based assessment methods were implemented in Scientific LogAnalyzer (e.g., the detection and handling of multiple sessions, computation of response times, and a module for analyzing and visualizing dropout). Figure 6.6 shows an example of the dropout tree generated by Scientific LogAnalyzer. Each node can be expanded or collapsed, and absolute and relative frequencies of choices of paths are calculated and displayed. After a speedy analysis of even large log files (Reips & Stieger, 2004), Scientific LogAnalyzer creates output in HTML or a tab-delimited form suited for import into statistics software. A more detailed example of a log file analysis is available from Scientific LogAnalyzer’s online help.

This section presented a description of how to create, conduct, and analyze data from a Web-based study with those tools my colleagues and I developed in our group. Of course there are alternative approaches. (For the design of simple, one page Web surveys, use SurveyWiz, Birnbaum, 2000). FactorWiz, also by Birnbaum (2000), is a tool for one page within-subjects factorial experiments. Yule and Cooper (2003) recently published Express, a program for large-scale simulations also used for Internet-based experimenting. Web-based assessments can also be created with proprietary software. One example is Authorware, (McGraw, Tew, & Williams, 2000) which can be used to create functional and attractive study materials. The downside of this approach is a steep learning curve, certain timing issues

(Schmidt, 2001), and the fact that it is difficult to get participants to download and install the required plug-in.

## FUTURE DEVELOPMENTS OF WEB-BASED ASSESSMENT METHODS

The previous sections of this chapter illustrated that Web-based methods offer a number of advantages to psychological assessment. The field has evolved enough to develop techniques and applications that allow for a smooth flow of the Web-based assessment process and secure the researcher with a good quality of data. Therefore, Web-based methods are inevitably being used in psychological research with much frequency during recent years. The web experiment list, for instance, now provides more than 300 Web studies (Reips & Lengler, 2005). With continued spread of knowledge of these methods and their integration into curricula, we will see a further increase in their professional use.

Apart from an increase in use and professionalism, a future trend of Web-based methods may be the development of more specialized Web-based methods in psychological assessment. Because many traditional methodological paradigms can somehow be transformed into a Web-based version, and the advantages are so appealing, we will likely see many more of these special applications.

The rapid development of Web technology and the spread of knowledge regarding its characteristics among psychologists assures that Web-based methods will strongly impact the way psychological assessment will be conducted in the future. The unending possibilities offered by this branch of media will perhaps be the beginning of a new era for psychological assessment and research.

# References

Antonak, R. F., & Livneh, H. (1995). Randomized response technique: A review and proposed extension to disability attitude research. *Genetic, Social, and General Psychology Monographs, 1,* 97-145.

Berendt, B. (2002). Using site semantics to analyze, visualize, and support navigation. *Data Mining and Knowledge Discovery, 6*, 37-59.

Berendt, B., & Brenstein, E. (2001). Visualizing individual differences in Web navigation: STRATDYN, a tool for analyzing navigation patterns. *Behavior Research Methods, Instruments, and Computers, 33*, 243-257.

Birnbaum, M. H. (2000). SurveyWiz and FactorWiz: JavaScript Web pages that make HTML forms for research on the Internet. *Behavior Research Methods, Instruments, and Computers, 32(2)*, 339-346.

Birnbaum, M. H. (2001). A Web-based program of research on decision making. In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet Science* (pp. 23-55). Lengerich: Pabst.

Birnbaum, M. H. (2004). Human research and data collection via the Internet. *Annual Review of Psychology, 55*, 803-832.

Birnbaum, M. H., & Reips, U.-D. (2005). Behavioral research and data collection via the Internet. In R. W. Proctor and K.-P. L. Vu (Eds.), *The handbook of human factors in Web design* (pp. 471-492). Mahwah, New Jersey: Erlbaum.

Bohner, G., Danner, U. N., Siebler, F., & Samson, G. B. (2002). Rape myth acceptance and judgments of vulnerability to sexual assault: An Internet experiment. *Experimental Psychology, 49 (4)*, 257-269.

Bordia, P. (1996). Studying verbal interaction on the Internet: The case of rumor transmission research. *Behavior Research Methods, Instruments, and Computers, 28*, 149-151.

Bosnjak, M. (2001). Participation in non-restricted Web surveys: A typology and explanatory model for item non-response. In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet Science* (pp. 193-208). Lengerich, Germany: Pabst.

Brenner, V. (2002). Generalizability issues in Internet-based survey research: Implications for the Internet addiction controversy. In B. Batinic, U.-D. Reips, & M. Bosnjak (Eds.), *Online Social Sciences* (pp. 93-114). Seattle: Hogrefe & Huber.

Buchanan, T., & Smith, J. L. (1999). Using the Internet for psychological research: Personality testing on the World Wide Web. *British Journal of Psychology, 90*, 125–144.

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

Buchanan, T., Ali, T., Heffernan, T. M., Ling, J., Parrott, A. C., Rodgers, J., & Scholey, A. B. (2005). Non-equivalence of online and paper-and-pencil psychological tests: The case of the Prospective Memory Questionnaire. *Behavior Research Methods, 37, 148-154.*

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.

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 December 27, 2001, from http://server3.uni-psych.gwdg.de/gor/contrib/buchanan-tom

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

Clark, S. J., & Desharnais, R. A. (1998). Honest answers to embarrassing questions: Detecting cheating in the randomized response model. *Psychological Methods, 3,* 160-168.

Coomber, R. (1997, June 30). Using the Internet for survey research. *Sociological Research Online, 2*. Retrieved June 16, 2002, from http://www.socresonline.org.uk/2/2/2.html

Dillman, D. A. (2000). *Mail and Internet surveys: The tailored design method* (2nd ed.). New York: John Wiley.

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.

Dillman, D. A., Tortora, R. D., & Bowker, D. (1998). Principles for constructing Web surveys. *SESRC Technical Report 98-50*. Pullman, Washington.

Drasgow, F., & Chuah, S. C. (2004). Computer-based testing. In M. Eid & E. Diener (Eds.), *Handbook of psychological assessment: A multimethod perspective* (pp. XXX-XXX). Washington, DC: American Psychological Association.

Eichstaedt, J. (2001). Reaction time measurement by JAVA-applets implementing Internet-based experiments. *Behavior Research Methods, Instruments, & Computers, 33*, 179-186.

Erdfelder, E., Faul, F., & Buchner, A. (1996). G*Power: A general power analysis program. *Behavior Research Methods, Instruments, and Computers, 28(1)*, 1-11.

Erdfelder, E., & Musch, J. (2004). Experimental methods of psychological assessment. In M. Eid & E. Diener (Eds.), *Handbook of psychological assessment: A multimethod perspective* (pp. XXX-XXX). Washington, DC: American Psychological Association.

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

Fritsche, I., & Linneweber, V. (2004). Nonreactive methods in psychological research. In M. Eid & E. Diener (Eds.), *Handbook of psychological assessment: A multimethod perspective* (pp. XXX-XXX). Washington, DC: American Psychological Association.

Joinson, A., & Reips, U.-D. (2004). Personalization, power, and online survey response rates. Manuscript submitted for publication.

Gockenbach, S., Bosnjak, M., & Göritz, A. (2003, March 30). Untersuchung stereotyper Beantwortungsmuster bei Matrixfragen in Web-Surveys [Investigation of stereotypic response patterns in matrix questions in Web surveys]. Paper presented at the German Online Research Conference 2004, Duisburg, Germany, March 30, 2004.

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

Hiskey, S., & Troop, N. A. (2002). Online longitudinal survey research: Viability and participation. *Social Science Computer Review, 20 (3)*, 250-259.

Horswill, M. S., & Coster, M. E. (2001). User-controlled photographic animations, photograph-based questions, and questionnaires: Three instruments for measuring drivers' risk-taking behavior on the Internet. *Behavior Research Methods, Instruments, and Computers, 33*, 46-58.

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

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, & Computers, 29*, 264-269.

Krantz, J. H., & Dalal, R. (2000). Validity of Web-based psychological research. In M. H. Birnbaum (Ed.), *Psychological Experiments on the Internet* (pp. 35-60). San Diego, CA: Academic Press.

Laugwitz, B. (2001). A Web experiment on color harmony principles applied to computer user interface design. In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet Science* (pp. 131-145). Lengerich: Pabst.

McGraw, K. O., Tew, M. D., & Williams, J. E. (2000). PsychExps: An on-line psychology laboratory. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet*. (pp. 219-233). San Diego, CA: Academic Press.

McKenzie, C. R. M., & Nelson, J. D. (2003). What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects. *Psychonomic Bulletin and Review, 10*, 596-602.

Musch, J., Bröder, A., & Klauer, K. C. (2001). Improving survey research on the World Wide Web using the randomized response technique. In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet Science* (pp. 179-192). Lengerich: Pabst.

Musch, J., & Reips, U.-D. (2000). A brief history of Web experimenting. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 61-88). San Diego, CA: Academic Press.

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

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17*, 776-783.

Reips, U.-D. (1995). The Web experiment method. Retrieved January 6, 2002, from http://wexlab.eu/WWWExpMethod.html

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: Hogrefe.

Reips, U.-D. (2000). The Web experiment method: Advantages, disadvantages, and solutions. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 89-114). San Diego, CA: Academic Press.

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.

Reips, U.-D. (2002a). Context effects in Web surveys. In B. Batinic, U.-D. Reips, & M. Bosnjak (Eds.), *Online Social Sciences* (pp. 69-79). Seattle: 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.

Reips, U.-D. (2002d). Theory and techniques of Web experimenting. In B. Batinic, U.-D. Reips, & M. Bosnjak (Eds.), *Online Social Sciences*. Seattle: Hogrefe & Huber.

Reips, U.-D. (2003). 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, August 25–27, 2003.

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.

Reips, U.-D., Morger, V., & Meier B. (2001). "Fünfe gerade sein lassen": Listenkontexteffekte beim Kategorisieren ["Letting five be equal": List context effects in categorization]. Unpublished manuscript, available at http://personalwebpages.deusto.es/reips/pubs/papers/re_mo_me2001.pdf

Reips, U.-D., & Mürner, B. (2005). Stroop Invaders: A Web tool for creating Internet-based Stroop experiments in game format. Manuscript submitted for publication.

Reips, U.-D., & Musch, J. (1999, November 18). Using the randomized response technique on the WWW. Paper presented at the 29th Society for Computers in Psychology (SCiP) Conference, Los Angeles, USA.

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.

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, & Computers, 36, 304-311.*

Richter, T., Naumann, J., & Noller, S. (2003). LOGPAT: A semi-automatic way to analyze hypertext navigation behavior. *Swiss Journal of Psychology,* 62, 113-120.

Rodgers, J., Buchanan, T., Scholey, A. B., Heffernan, T. M., Ling, J., & Parrott, A. (2001). Differential effects of Ecstasy and cannabis on self-reports of memory ability: A web-based study. *Human Psychopharmacology: Clinical and Experimental, 16*, 619-625.

Ruppertsberg, A. I., Givaty, G., Van Veen, H. A. H. C., & Bülthoff, H. (2001). Games as research tools for visual perception over the Internet. In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet Science* (pp. 147-158). Lengerich: Pabst.

Schmidt, W. C. (1997). World Wide Web survey research: Benefits, potential problems, and solutions. *Behavior Research Methods, Instruments, and Computers, 29*, 274-279.

Schmidt, W. C. (2001). Presentation accuracy of Web animation methods. *Behavior Research Methods, Instruments and Computers,* 33, 187-200.

Schwarz, S., & Reips, U.-D. (2001). CGI versus JavaScript: A Web experiment on the reversed hindsight bias. In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet Science* (pp. 75-90). Lengerich: Pabst.

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: Hogrefe & Huber.

Voracek, M., Stieger, S., & Gindl, A. (2001). Online replication of evolutionary psychology evidence: Sex differences in sexual jealousy in imagined scenarios of mate's sexual versus emotional infidelity. In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet Science* (pp.91-112). Lengerich: Pabst.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association, 60,* 63-69.

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: Hogrefe & Huber.

Yule, P., & Cooper, R. P. (2003). Express: A Web-based technology to support human and computational experimentation. *Behavior Research Methods, Instruments, and Computers, 35 (4)*, 605-613.

Figure 6.1. A Web-based survey using the randomized response technique (RRT) in a study on the trustworthiness of computerized random generators.

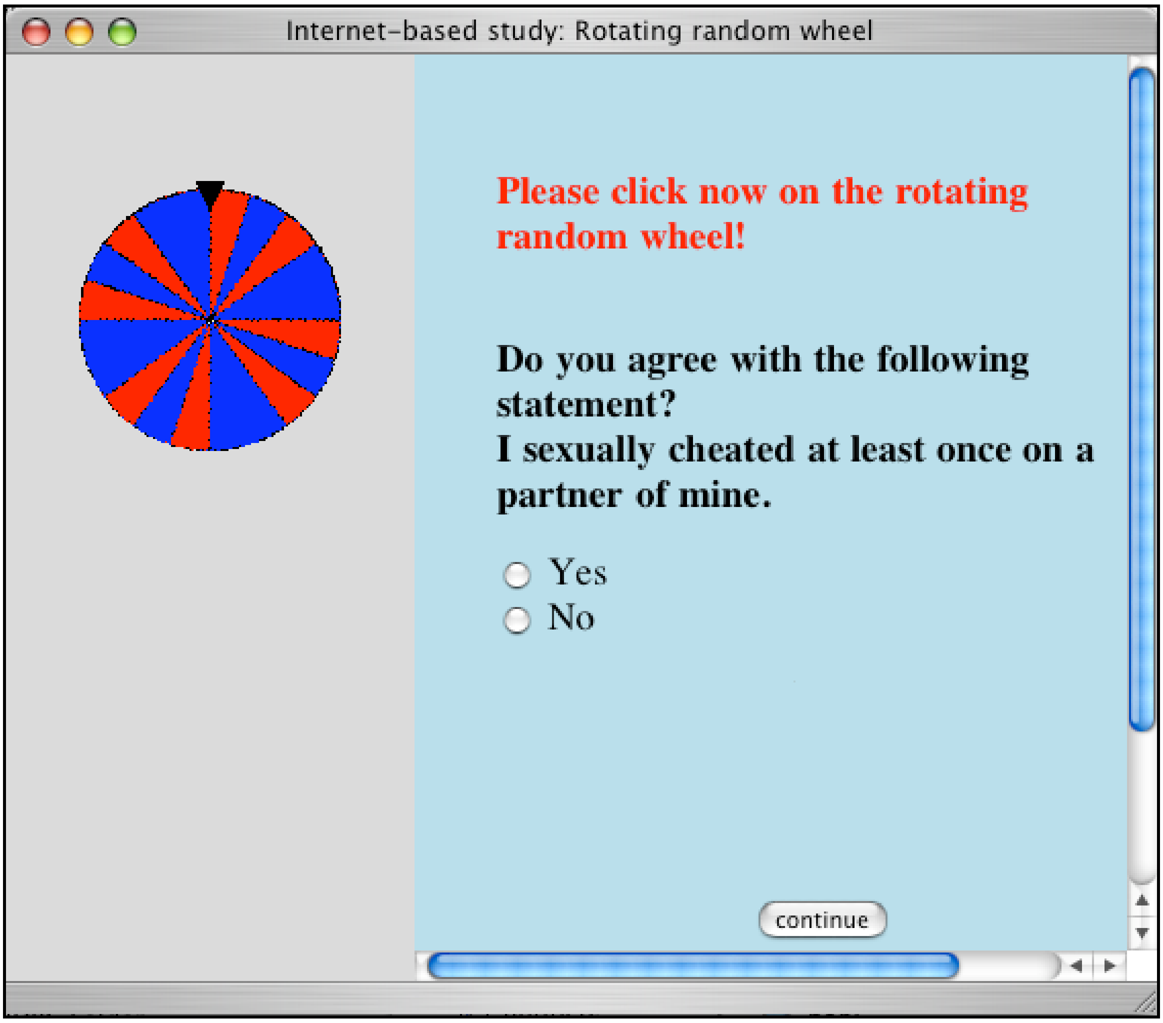


Figure 6.2. “Stroop Invaders”: A Web site that allows researchers, teachers, and students to design and conduct Web-based Stroop experiments.

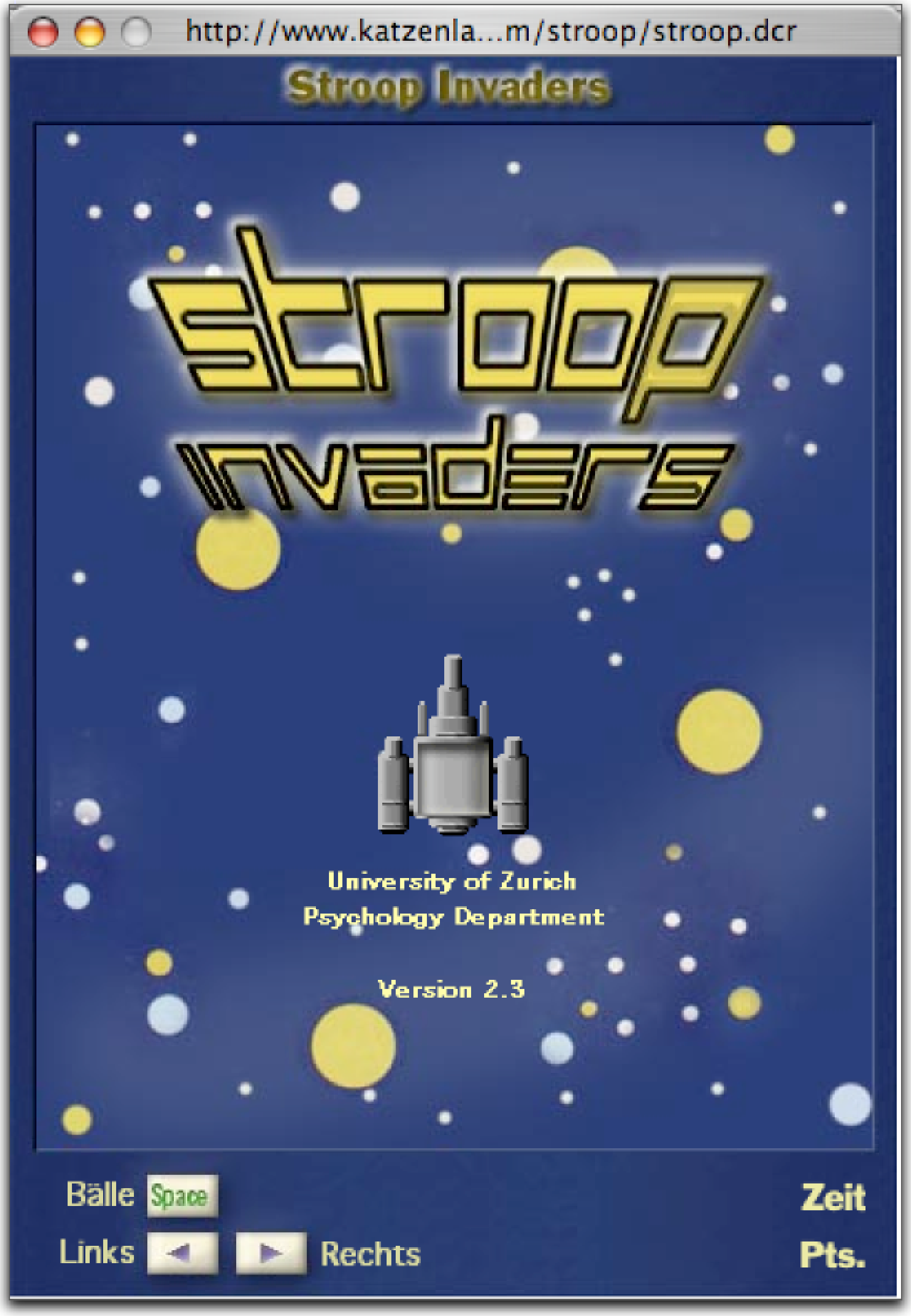


Figure 6.3. WEXTOR's entry page.

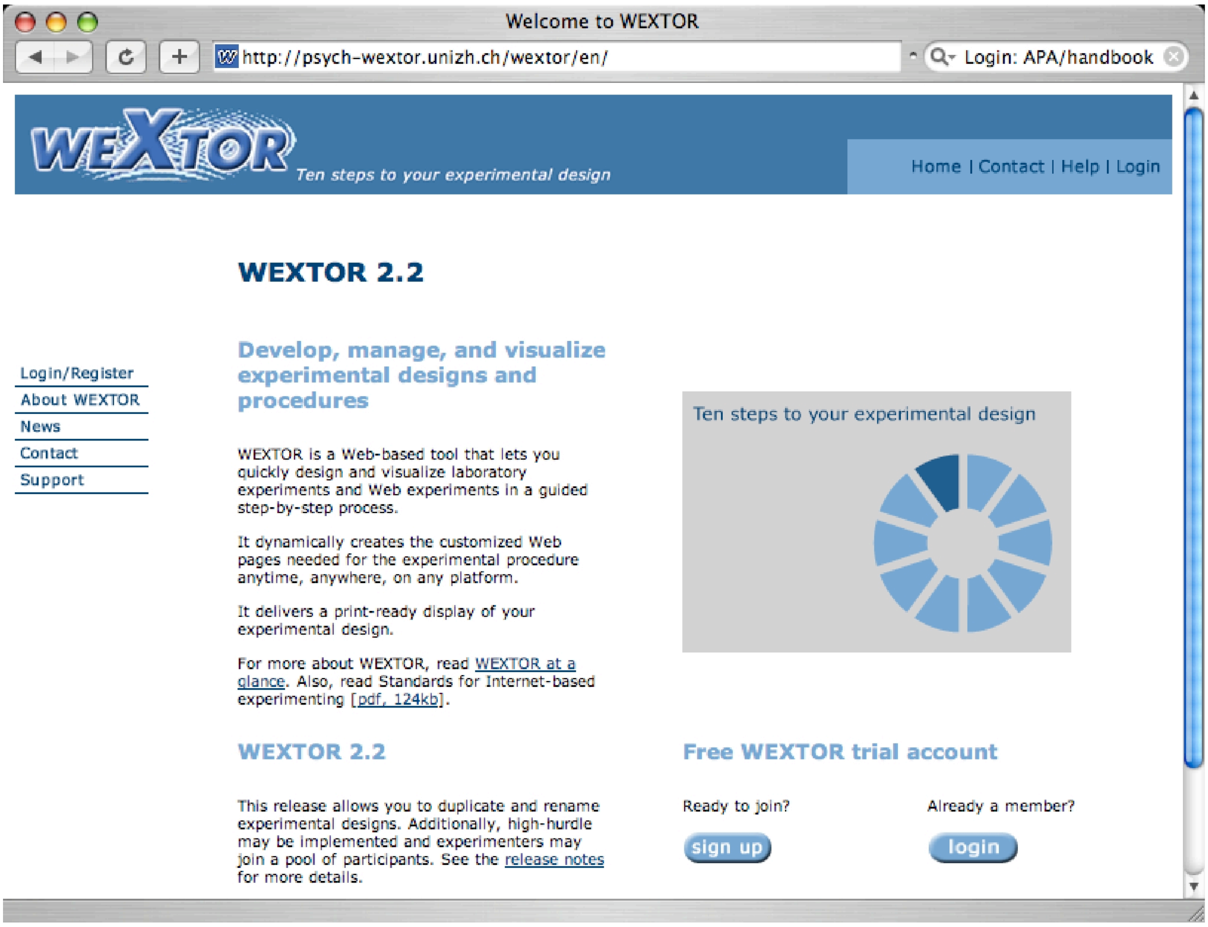


Figure 6.4. A visual display of the design and materials in the cup Web experiment, showing the four experimental conditions to which participants are randomly distributed as well as the folders and Web pages used in the Web experiment. The display is created in Step 9 in WEXTOR.

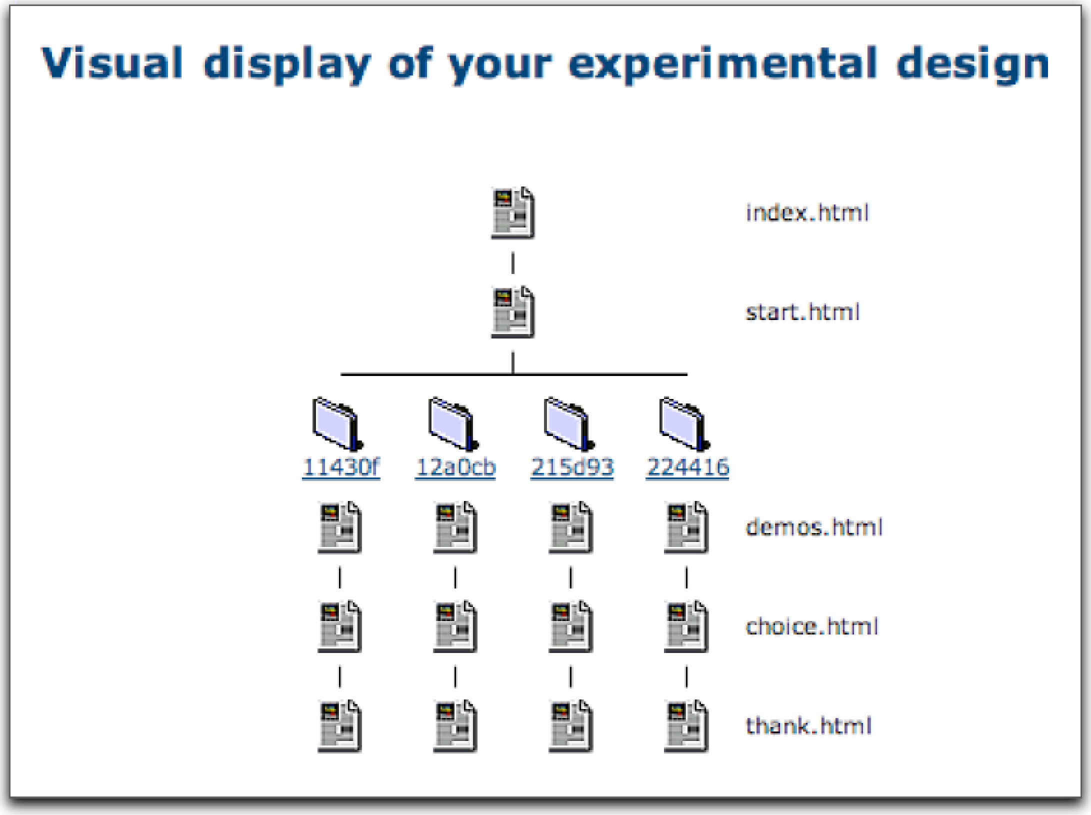


Figure 6.5. The form to be used to submit a Web experiment to be linked on the web experiment list.

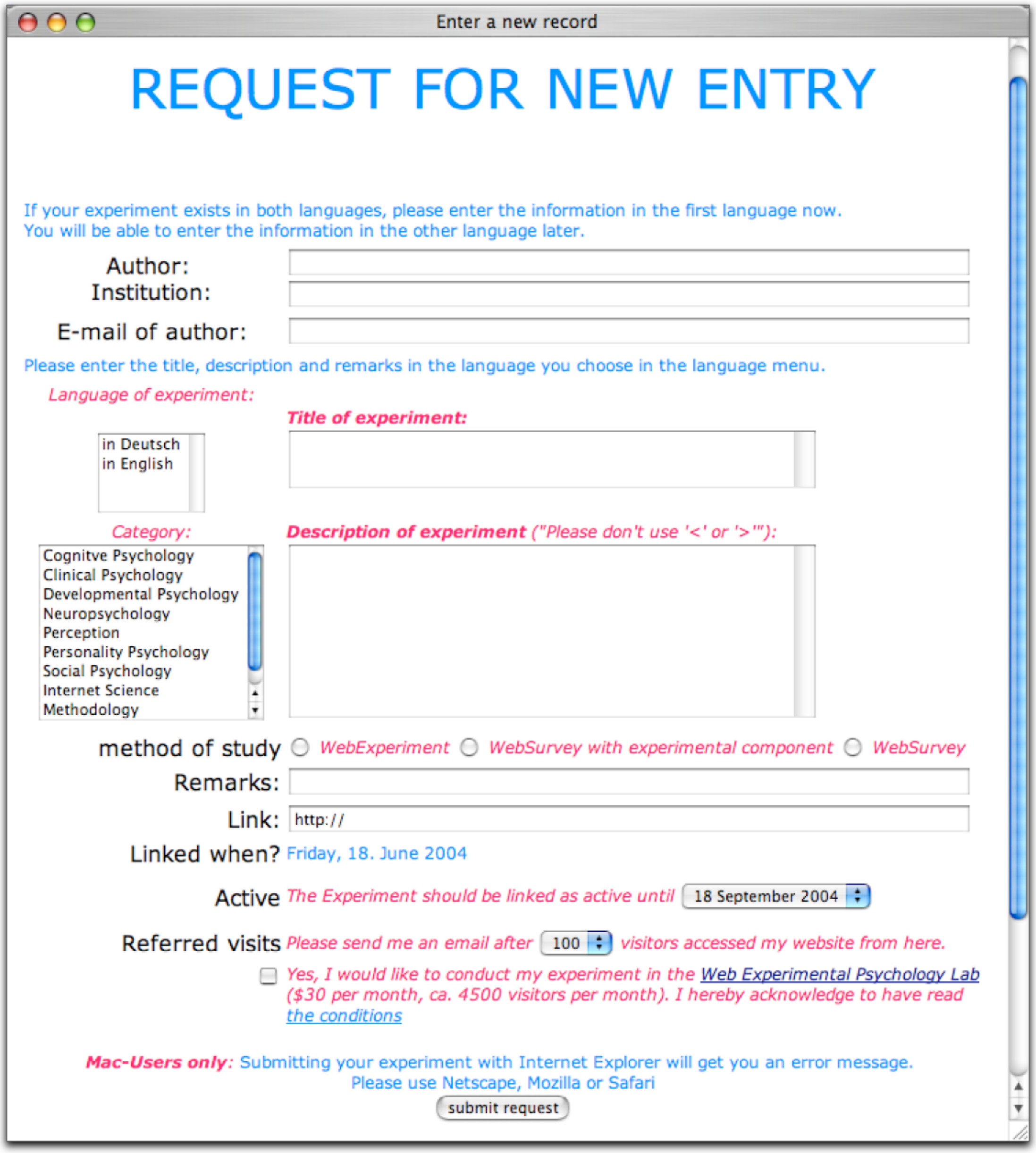


Figure 6.6. An example for the dropout tree that can be generated with Scientific LogAnalyzer. Each node can be expanded or collapsed and absolute and relative frequencies of choices of paths are calculated and displayed.