The methodology of Internet-based experiments

Ulf-Dietrich Reips

Oxford Handbook of Internet Psychology
Edited by Adam N. Joinson, Katelyn Y. A. McKenna, Tom Postmes, and Ulf-Dietrich Reips

Print Publication Date: Feb 2009
Subject: Psychology, Social Psychology, Cognitive Psychology
Online Publication Date: Sep 2012 DOI: 10.1093/oxfordhb/9780199561803.013.0024

Abstract and Keywords

This article discusses methods and techniques, procedures and tools that have been found to be necessary or useful in Internet-based experimenting. While the focus is on experiments, many of the methods apply to other types of Internet-based research as well. The article is structured in a step-by-step fashion, guiding the reader through the various stages of setting up and conducting a web experiment. Apart from general issues, the relevant steps begin with planning, generating, and pre-testing an experiment. They continue with recruitment and monitoring, then analysis and archiving.

Keywords: Internet-based experimenting, experimental methods, Internet-based research

Introduction

This chapter contains information about methods and techniques, procedures and tools that have been found to be necessary or useful in Internet-based experimenting.[1] While the focus is on experiments, many of the methods apply to other types of Internet-based research as well.

The Internet-based experimental method lends itself easily to research in particular areas of psychology, more so in some than in others. For example, it is straightforward to use the method in judgement and decision-making research (Birnbaum 2001; Schulte-Mecklenbeck and Neun 2005) and other areas of cognition, while there are limitations to Internet-based research in neuropsychology (see Erlanger *et al.* [2003] for a noteworthy exemption). Interestingly, it could be shown empirically that the distribution of usage among subfields of psychology remains constant over the years. Reips and Lengler (2005) analysed data from submissions by experimenters to one of the largest web portals for Internet-based experiments, the Web Experiment List,[2] and compared the results with those from a web survey conducted by Musch and Reips (2000) in 1998 and 1999 among the early web experimenters. Table 24.1 shows their results. Note that social psychology rivals cognition in number of Internet-based studies.

It has also repeatedly been noted that the number of Internet-based experiments conducted is increasing exponentially or nearly exponentially (Reips 2002c; Birnbaum 2004; Birnbaum and Reips 2005). In a recent analysis of data from the Web experiment list this description of the trend turned out to be appropriate (see Figure 24.1) (Reips, 2006a). The number of web experiments in the list is now (July 2006) at 400.

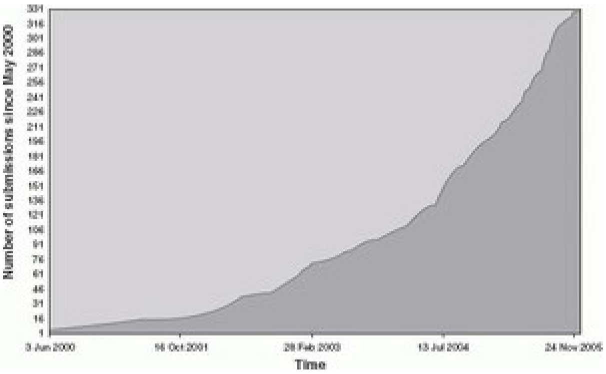

Figure 24.1 Number of submissions of web experiments to the Web Experiment List by time (adapted from Reips, 2006a).

Given the widespread and increasing use of the method it is surprising that most universities are not yet fully prepared to teach students how to conduct this type of research properly. (p. 374) One of the reasons may be that most researchers and teachers received their own methods training during a time when the Internet was little more than an obscure military network. As a result of the widespread lack of education in Internet-based research methods and the misconception that Internet-based research can be conducted by applying laboratory techniques in a 1:1 fashion, many web-based studies appear on the Internet and in manuscripts that do not meet basic methodological standards[3] (Reips, under review).

Table 24.1 Number and percentage of web experiments by field in two studies, from the end of the 1990s and 2005

Field	**Web survey in 1998/9 (Musch and Reips, 2000)**	**Web experiment list (Reips and Lengler, 2005)**
Cognitive Psychology	10 (56%)	100 (40%)
Social Psychology	4 (22%)	90 (36%)
Perception	1 (5%)	25 (10%)
Internet Science	1 (5%)	10 (4%)
Personality Psychology	1 (5%)	10 (4%)
Clinical Psychology	-	10 (4%)
Development Psychology	-	3
Neuropsychology	1 (5%)	2
Sum	18	250

For example, Reips (2002b) describes a number of frequent ‘configuration errors’ that may lead (p. 375) to (often undetected) methodological and/or ethical problems. *Configuration errors I and II* describe how just by design of the Web study confidential participant data may accidentally end up in the wrong hands.[4] In the case of *configuration error V* one of the legitimate responses to a question is erroneously preselected. This happens easily if the experimenter thinks in terms of usual procedures in paper-based questionnaires, where a range of categories is listed for the participant to choose from. For example, ‘below 10’ may end up as a preselected option in a drop-down menu near a question that asks the participant to select his or her age. Rather than simply transferring the question format from paper to the Internet, it is important to preselect an option that says something like ‘please choose here’ that returns a ‘missing’ code unless the participant makes a choice. If, however, a legitimate answer is preselected the experimenter will be unable to distinguish real data from those that result when participants fail to respond to the item—quite a frequent behaviour on web pages. In the example above, the sample would appear to contain a surprisingly high number of children.

How to counter the problem of frequent flaws in Internet-based studies? There are at least three solutions: one way will be to include Internet-based research methodology in curricula. This requires colleges to employ teachers who are savvy of Internet-based research and change their course programmes. A second possibility is the use of software that automatically implements techniques that maximize the advantages of Internet-based data collection and prevent the experimenter from pitfalls arising from the particularities of conducting research via the Internet—for example WEXTOR for web experiments (see Reips and Neuhaus, 2002, and http://psych-wextor.unizh.ch/wextor/en/). Finally, the knowledge about Internet-based research methods needs to be disseminated to an interested readership—an aim this section of this book will hopefully meet.

The current chapter is structured in a step-by-step fashion, guiding the reader through the various stages of setting up and conducting a web experiment. Apart from *general issues* the relevant steps begin with *planning, generating*, and *pre-testing* an experiment. They continue with *recruitment* and *monitoring*, then *analysis and archiving*.

General issues

It should be noted that conducting Internet-based experiments requires knowledge of and familiarity with methodological and technological particularities that relate to the networked structure of the Internet. Not only is it necessary to understand the basics of HTML (the code web pages are made of), the workings of the TCP/IP protocol (the Internet's address and transportation system), and the meshwork of technical components in a web experiment (e.g. Reips 2000; Schmidt, Chapter 29 this volume) but how participant behaviour changes in Internet-based studies.

Much of what we know from research in computer-mediated communication (CMC) can be applied to the participant situation in Internet-based research, because in this type of research participants use a computer by definition (even if it is a mobile phone or PDA). At least five issues are critical in CMC:

1. cues transmitted,
2. bandwidth,
3. cost constraints,
4. level and type of anonymity,
5. synchronicity and exclusivity (Joinson 2003).

In comparison to a laboratory situation and depending on the exact Internet service used, Internet-based research may thus primarily differ on these dimensions, and suffer or profit from their impact on the information transmission and on the psychology of the participant.

One of the reasons why the methods of Internet-based experimenting became widely used is the *fundamental asymmetry of accessibility* (Reips 2002c, 2006b). What is programmed to be accessible from any Internet-connected computer in the world will surely also be accessible in a university laboratory, but what is programmed to work in a local computer lab may not necessarily be accessible anywhere else. A laboratory experiment cannot (p. 376) simply be turned into a web experiment, because it may be programmed in a stand-alone programming language and lack Internet-based research methodology, but any web experiment can also be used by connecting the laboratory computer to the Internet. Consequently, it is a good strategy to design a study web-based, if possible.

Web-based methods offer several benefits to the researcher (for summaries, see Birnbaum 2004; Reips 2000, 2002c). The main advantages are that:

- ♦ it is possible to test large numbers of participants quickly;
- ♦ it is possible to recruit large heterogeneous samples and people with rare characteristics (Schmidt 1997); and
- ♦ web-based methods are more cost-effective in administration, time, space, and labor in comparison with laboratory research.

Methodological analyses and studies reveal that web-based methods are usually valid (e.g., Krantz and Dalal 2000) and sometimes even generate higher quality data than laboratory studies (Buchanan and Smith 1999; Reips 2000; Birnbaum 2001). They facilitate research in previously inaccessible areas (e.g., Bordia 1996; Coomber 1997; Rodgers *et al.* 2001).

Other benefits of web-based methods are

- ♦ the ease of access for participants (bringing the experiment to the participant instead of the opposite);
- ♦ the ease of access to participants from different cultures, given Internet access and the availability of the Web experiment in the respective languages—for instance, Bohner, Danner, Siebler, and Samson (2002) conducted a study in three languages with 440 women from more than nine countries;
- ♦ truly voluntary participation (unless participants are required to visit the Website);
- ♦ detectability of confounding with motivational aspects of study participation;
- ♦ the better generalizability of findings to the general population (e.g., Brenner 2002; Horswill and Coster 2001);
- ♦ the generalizability of findings to more settings and situations because of high external validity—Laugwitz (2001), for instance, was able to show that a colour perception effect in software ergonomics persisted despite the large variance of conditions of lighting, monitor calibration, etc. in participants' settings;

- ♦ the avoidance of time constraints;
- ♦ the simultaneous participation of large numbers of participants is possible;
- ♦ the reduction of experimenter effects (even in automated computer-based assessments there is often some kind of personal contact, but this is not so in most web-based assessments);
- ♦ the reduction of demand characteristics (see Orne 1962);
- ♦ 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);
- ♦ the access to the number of people who see the announcement link to the study, but decide not to participate;
- ♦ the ease of cross-mode comparison—comparing results with results from a sample tested in the laboratory;
- ♦ greater external validity through greater technical variance; and
- ♦ the heightened public control of ethical standards.

In preparation of conducting a web experiment it is important to adhere to good research ethics and Internet etiquette. Ess (Chapter 31 this volume) discusses a number of general principles and issues. Birnbaum and Reips (2005) give a set of recommendations for good practice in all stages of conducting a study via the Internet. They also stress putting dangers from this type of research into perspective, i.e. that ‘it is difficult to injure someone via the Web, except by dishonesty, so the fundamental ethical principle for web-based research is honesty’ (2005: 488). Their recommendations include:

- ♦ keeping up with promises (e.g. regarding renumeration, information);
- ♦ avoiding anything that resembles a chain letter, commercial advertising, or spamming;
- ♦ taking care of security and confidentiality;
- ♦ polite and careful communication with participants and intermediaries (e.g. administrators of mailing lists and newsgroups that are contacted for recruitment).

(p. 377) Generally, the Internet researcher should be aware of the technological basis, i.e. networking, protocols, software, hardware and the resulting technical variance (Reips 2000, 2002b, 2002c; Chapter 29 this volume). One also needs to know that technology may interact with a study's topic and/or participant demography (Buchanan and Reips 2001). Finally, there may be mode effects—participants with computer anxiety will be less likely to join web studies, tests may turn out to need re-validation (Buchanan, Chapter 28 this volume), and response behaviour with a mouse arrow on a web page differs in many ways from response behaviour with a pencil on paper (see Bosnjak 2001; Reips 2002a).

Planning a web experiment

Setting up a web experiment for the first time means an investment. Less so in terms of infrastructure: nowadays it is fairly easy to set up and maintain a web server with web pages on it and record the data in a data base or a log file for later analysis, most new computers come with the necessary software pre-installed.[5] More likely an investment for researchers is to take the time and consider, learn and implement proper methodologies for Internet-based research. Alternatively, one may use a web service that automatically implements these methodologies (see 'Generating a web experiment'). The iScience Server at http://psych-iscience.unizh.ch/ (see Figure 24.2) serves as a portal to a number of such services.

Techniques for handling of dropout

Figure 24.2 Main page of the iScience Server at the University of Zürich (http://psych-iscience.unizh.ch/), with links to web experiment generators (WEXTOR, idex), recruitment portals (Web Experimental Psychology Lab, Web Experiment List, Web Survey List), Scientific LogAnalyzer, DIP, and an online Big Five personality test.

Dropout (attrition) is built into Internet-based research. An Internet scientist can choose (p. 378) among four strategies of dealing with it: use it, control it, avoid it or suppress it. Reips (2002b) recommends using dropout as a dependent variable and describes how it may be used for detecting *motivational confounding* by looking at dropout curves by experimental condition. If there is higher dropout in one condition, then it can be concluded that the condition is more boring or otherwise less attractive to the participants. Comparisons of experimental conditions on other dependent variables would therefore be hampered by confounding. Figure 24.3 shows dropout curves for two conditions that would indicate motivational confounding, if the experiment were to be conducted in the laboratory.

The *high entrance barrier* or *high hurdle technique* (Reips 2000, 2002d) is a control procedure that can be applied to provoke dropout to happen early and ensure continued participation after someone makes the decision to stay. This means bundling of demotivating factors (i.e. long loading times from large images, long texts) at the very beginning of a web experiment, so visitors with a low motivation for continued participation will likely drop out early. Motivating factors should be implemented increasingly thereafter, enticing participants to continue with the experiment. Reips (2000: 110–111) recommends several measures that are likely to support this strategy:

- Tell participants participation is serious, and that science needs good data.
- Personalize—ask for an email address and/or phone number.
- Tell them you can trace them (via their computer's IP address).
- Be credible: tell them who you are, and your institutional affiliation.
- Tell them how long the Web experiment will take.
- Prepare them for any sensitive aspects of your experiment (e.g. ‘you will be asked about your financial situation’).
- Introduce your experimental manipulation after a warm-up phase.
- Tell them what software they will need (and provide them with hyperlinks to get it).
- Perform Java, Javascript, and plug in tests (but avoid Active-X).
- Make compliance a prerequisite for winning the reward.

Participant dropout during a web experiment can be avoided via the *warm-up technique*. It is based on the observation that most dropout will take place at the beginning of an on-line study, forming a ‘natural dropout curve’ (Reips 2002d). A main reason for the initial dropout is the short orientation period many participants show before making a final decision on their participation.

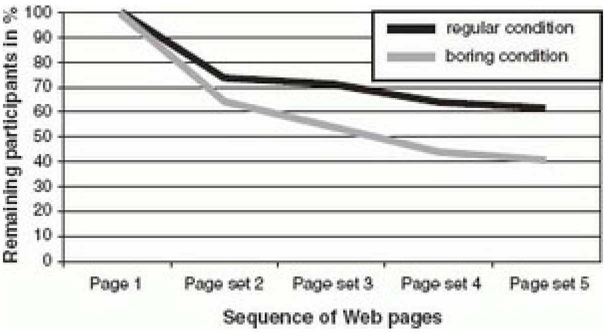

Figure 24.3 Dropout curves for two conditions that differ in attractiveness and may therefore cause motivational confounding if conducted as a laboratory experiment.

Finally, the most simple and elegant technique of avoiding a large portion of the dropout, namely one that results from mere curiosity, is the *seriousness technique*. Participants are merely asked ‘How do you intend to browse the Web pages of this study?’ and are given the two options ‘I would like to seriously participate now’ and ‘I would like (p. 379) to look at the pages only.’ By far the largest proportion of those who later drop out will indicate so by choosing the latter option.

Meta-tagging techniques

Meta tags are snippets of information beginning with ‘<meta’ that are located in the ‘head’ section of the source code of web pages.[6] Meta tags can be used for a variety of tasks, here the focus is on guiding participants' Web browsers, so-called ‘proxy caches’, and search engines.

The first meta tag displayed in Table 24.2 (<meta name=‘ROBOTS’ content=‘NONE’>) will keep search engines from indexing pages subsequent to the designated entry page, preventing scores of participants entering on pages where they will be asked for their sex and other delicate questions. The meta tag will be obeyed by most of the indexing requests (called ‘robots’, ‘bots’, ‘spiders’, or ‘crawlers’) sent by search engines, but unfortunately there are exemptions.

The second meta tag displayed in Table 24.2 keeps caches and proxy servers from serving oldversions of research materials. As will be seen in ‘Pre-testing of web experiments’ below, it rarely happens that a newly designed web experiment is without errors. During pre-testing, a study needs to be updated frequently, as new versions are developed and tested again. The pragma no-cache meta tag tells servers that function as intermediary storages spaces (caches) for web content flowing in and out of institutions to skip the content of the current web page. If the content were stored (cached) old versions of the research materials would go to all participants using computers in the respective institution until the cache is updated. Here the third meta tag is asked to play its role. It suggests that the content of the current web page has long expired, triggering servers to delete the content from their caches.

Table 24.2 Use of meta tags in pages after the entry page of a web study

<HTML>
<HEAD>
<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>
(followed by the body of the web page)

Note. This method can be used to keep search engines from indexing pages subsequent to the first page of a web experiment, and to prevent caches and proxy servers from serving old versions of research materials. (These meta tags are automatically added to pages in web experiments created with WEXTOR.)

Other techniques to consider in the planning stage

Password techniques (e.g. Schmidt 1997) provide web experimenters with ways of limiting access to the study that may be important, for example if a clearly defined sample is required. If passwords are individualized there is a very good chance that indeed the individual using the password is the person it was given to. On an even higher level of security, passwords expire after they have been used once. Some ways of conducting Internet-based experiments, for example in online panels, automatically include authentication (see Göritz, Chapter 30 this volume).

Birnbaum and Reips (2005) recommend building some redundancy and cross-examination into questionnaires on the Internet. A number of techniques can be used to re-examine answers before sending people to different questionnaires. For example, Javascript can provide cross-examination when a person clicks a link to identify his or her gender. If the person clicks ‘male’, the prompt opens a new box with the question, ‘is it correct: you are male?' requiring a yes/no answer before the person can continue. A similar check cross-examines those who click 'female'. A similar technique is to provide HTML links that provide a 'second chance' to link to the correct gender. The person who clicks (p. 380) 'male' then receives a page with a link to click 'if you are female', which would send the person to the female questionnaire, where there is also a chance to revert.

One of the most helpful measures in planning an experiment is to visualize the design. WEXTOR, the experiment generator for web and lab described further below, includes visualization of experimental designs and procedures. In addition, it automatically implements most of the techniques discussed above.

Generating a web experiment

The simplest way to create a web experiment is by creating a series of web pages or parallel strings of web pages that contain experimental materials (instructions, pictures, sound files etc.) and measures (scales made from radio buttons etc.) and implement random distribution to conditions. There are a number of options for the latter:

1. Client-side scripting, meaning that a snippet of script code is included with a web page and tells the web browser to run certain processes on the participant's computer as soon as the web page is opened and the script code is started;
2. Server-side scripting, by which all of the computing is done on the server and only the result is sent to the client;
3. HTML-based pseudo-randomization, for example the birthday technique, where participants select their experimental conditions by mouse-clicking on their birthday or -month (e.g., Birnbaum and Reips 2005).

Client-side versus server-side

Client-side scripting can be an advantage, in that it frees the server from making calculations and having a lot of traffic delays sending information back and forth. It can also be a disadvantage, if the participant computer or web browser does not understand the scripting language or has it turned off. At a time when Internet Explorer used an error-prone version of a Javascript compiler and some people had turned it off (fearing to allow other people's programs to run on one's own machine), Schwarz and Reips (2001) found that Javascript caused a higher rate of dropout in web studies compared with methods that did not require client-side scripting. However, in recent years, Javascript is so prevalent in websites that few people have it turned off.

By doing the computing on the server side, one guarantees that any participant at a web browser can complete the study (Schmidt 2000 see Chapter 29, this volume). On the other hand, server-side programs may introduce delays as the participant waits for a response from the server. When there are delays, some participants may become frustrated or think the program has stopped, and may terminate their participation. There are, however, certain tasks that can and should only be done by the server, such as saving data or handling issues of security (e.g., passwords, using an exam key to score a test, etc.). Perl and PHP are the two most popular programming languages that can be used to write server-side programs. Of course, to program a server, one needs to have access to the server and the knowledge to program it.

An example for client-side scripting is Javascript, it is used by WEXTOR, the (web) experiment generator that will now be described.

WEXTOR

There are not many software systems for Internet-based experimenting yet, and even some of these few are not freely available. In the following a system is presented that integrates many of the Internet-based research methods discussed in this chapter. WEXTOR is a free web service that serves as an experiment generator and teaching tool on the WWW that can be used to design laboratory experiments and web experiments in a guided step-by-step process. Web services are WWW-based 'software', a new type of interactive application that is accessible wherever one has access to the Web. WEXTOR dynamically creates the custom-tailored web pages and Javascripts needed for the experimental procedure, and it provides the experimenter with a print-ready visual display of the experimental design. WEXTOR flexibly supports complete and incomplete factorial designs with between-subjects, within-subject, and quasi-experimental factors as well as mixed designs. (p. 381) The software implements both client-side and server-side time measurement and contains a content wizard for the creation of interactive materials as well as dependent measures (visual analogue and radio button scales, multiple choice items, etc.) on the experiment pages, but its aim is not to replace a full-fledged HTML editor. Several of the methodological features described in this chapter that are specifically necessary in web experimental design have been implemented in the Web-based tool, for example the seriousness check technique and the high hurdle technique. WEXTOR is platform-independent, so it runs on any type of operating system in any type of web browser, as long as Javascript is supported. The created web pages can be uploaded to any type of web server, where data may be recorded in log files or via database, or they can simply be served from the WEXTOR website, where hosting of the experiment is offered as an additional service.

Many human factors considerations are built into WEXTOR, and it automatically prevents several methodological pitfalls in Internet-based research. This web service uses non-obvious file naming, automatic avoidance of page number confounding, Javascript test and redirect functionality to minimize drop-out and randomized distribution of participants to experimental conditions. Different from the usual practice in many offline studies, demographic questions by default are asked at the *beginning* of a web experiment, because this procedural sequence has been shown to produce less dropout and better data quality (Frick et al. 2001), and asking contact information early reduces the ethical and procedural problem that participants who drop before providing the information can't be contacted. WEXTOR also provides for optional assignment to levels of quasi-experimental factors, optional client-side response time measurement, randomly generated continuous user IDs for enhanced multiple submission control, and it automatically includes the meta tags described in this chapter. In addition, WEXTOR implements the warm-up technique for dropout control mentioned above and provides for interactive creation of dependent measures and materials. The hosting facility supports download of the original log file (raw data) and a data file ready for import to Excel or SPSS.

WEXTOR's current version may be used freely for educational and non-commercial uses. Its web address http://psych-wextor.unizh.ch/wextor/.

Pre-testing a web experiment

Internet-based experiments offer so many advantages in comparison with laboratory experiments (Reips 2000, 2002c) that some of the challenges may easily be overlooked. In Internet-based studies the experimenter's control over the participant setting is much reduced. Often, it is not possible or desirous to communicate with the participants during the experiment. As a result, there may be questions or misunderstandings with some participants regarding the experimental materials or procedure that can not be corrected, and technical errors may go undetected. (Chapter 29, this volume). Consequently it pays off to invest much time and effort into pre-testing the Internet-based experiment.

Experiments generated with WEXTOR offer the usability advantages of self-containment of the experiment files and a one-to-one correspondence of files and folders to screens and conditions.

Procedures to apply in pre-testing

Pre-testing begins with repeated tests by the experimenter. These should be conducted with different web browsers, and possibly from different computers. Most important are checks of the correspondence between surface and deep structure of the web pages: Is a response to a radio button recorded as intended and consistent throughout the study materials, e.g. will a ‘Yes’ response to the question ‘Do you wear glasses?’ always be recorded as ‘1’, even if the order of responses is varied?

Pre-tests are best to be continued with a small number of friends and acquaintances who will not hesitate in helping the experimenter in improving the study. Other possible procedures include unobtrusive observation of volunteers in the lab and trying the experiment on different web browsers and on a number of computers with a variety of monitors. The final stage of pre-testing should always be a limited ‘live’ online test.

(p. 382) What to discover and how to prevent

Some issues are frequently discovered during pre-testing, others may go undetected until the data from test trials are downloaded and analysed.

If text fields or text areas are used as input devices on web pages, it may sometimes be forgotten to limit the number of characters that can be filled into these spaces. In the case of asking for a percentage it makes sense to limit the number of digits to three. Otherwise, the data file may later contain entries saying 'one hundred' or '58.99999' or even a ten page excerpt from a googled dissertation on the 'understanding of the probability format'. Such entries cause a lot of work and possibly methodological issues during analysis, and may in some cases even crash the Web server. The content of a text field can be limited by using the maxlength tag 'e.g. writing 'maxlength=3' for limiting input to three characters'.

Web experimenters are forced to make some critical decisions while planning their experiments. It is necessary to decide if a certain functionality is worth the additional technologies, potential methodological disadvantages and the psychological impact that is associated with it. For example, Schwarz and Reips (2001) found certain Javascript routines to dramatically increase the dropout in a web experiment by comparing a non-Javascript version with an otherwise identical version of the experiment that used these routines. Furthermore, Buchanan and Reips (2001) showed that sampling may get biased depending on technology used in Internet-based research. In their web-based personality study (also see Chapter 28 this volume) they observed respondents to indicate a higher level of education in the non-Javascript version, probably indicating a higher awareness of Internet security problems in more educated persons. They also observed personality differences between people accessing the study on a Mac versus a PC.

A problem that can occur if one codes the web pages used in an Internet-based experiment by hand: the value used for missing data is the same as a code used for real data (Birnbaum and Reips 2005). For example, the authors describe a survey on the Web in which the participants were asked to identify their nationalities. In this case, the same code value (99) was assigned to India as to the preselected 'missing' value. If not warned, the researcher might have concluded that there had been a large number of participants from India.

Recruitment

Internet-based studies appeal to many people, because they are curious about research or they see it as a way to learn something about their psychological self, as in self-tests in magazines. It has been found that material incentives are not of highest importance when recruiting participants (Göritz 2006), but small incentives do have an impact on dropout and data quality (Frick et al. 2001).

How to find participants for Internet-based experiments? Of course, one may always contact family and friends. However, these people often have particular knowledge and motivations towards the experimenter, they may bias the results. As such, this group is ideal for pre-testing the experiment (see above), because they will likely tend to talk freely about the technical problems they encountered or the way they (mis)understood the instructions.

Web portals and lists

Free recruitment is available from portals and lists for web experimentation. The largest such sites are ‘Psychological Research on the Net’ by John Krantz (http://psych.hanover.edu/research/exponnet.html) and the criterion-searchable web experiment list shown in Figure 24.4 (http://genpsylab-wexlist.unizh.ch/) (Reips and Lengler 2005). Entries can be requested via a web form, each entry will be checked by an administrator who is knowledgeable in Internet-based research. The highly awarded Web Experimental Psychology Lab (http://tinyurl.com/dwcpx [Reips 2001], founded in 1995, is the oldest website for recruitment of participants and resembles a virtual laboratory. More web experiments can be found at the Online Social Psychology Studies site by Scott Plous (http://www.socialpsychology.org/expts.htm).

Mailing lists, forums, and newsgroups

Figure 24.4 The Web Experiment List (Reips and Lengler 2005), entry page.

One very effective way of recruiting participants are emails to mailing lists or messages to forums and newsgroups of people who don't mind (p. 383) receiving messages with information about studies for participation (e.g. mailing lists for psychology students). At a conference on decision-making (SPUDM, Zürich 2003), the author heard an interesting paper on the first day of the conference, and decided to replicate that study overnight. He was able to include the results in his talk on the second day of the conference, in order to demonstrate the efficiency of web-based research (see Reips 2006b). Within eight hours, complete data sets from 162 participants (compared to 64 in the original study) were recorded in the Web experiment, most of which were recruited via three mailing lists.

Recruiting via mailing lists, forums or newsgroups should be done with care, because some of the readers/recipients of the recruitment message may see it as spam. Therefore, it is best to ask the group's moderator or administrator for permission to send the message or even convince the moderator to endorse it and send it.

Participant pools and online panels

Often, researchers are keen on knowing who is part of the population they are sampling from. Well-kept participant pools or online panels are a way of approximating this goal, and potentially having hundreds or thousands of motivated participants ready for participation (see Göritz 2006 and Chapter 30 this volume; Reips, 2000). However, there is discussion about panel-lists getting used to or even burnt out on the panel. They may know how to work their way through studies without paying much attention to it or even abuse the panel, in particular if they receive incentives for being a member. Also, panels likely become targets for manipulation attempts, if they are used in market research (Geissler *et al.* 2005).

(p. 384) Search engines and banners

Meta tags (see Planning of web experiments) can be used for recruitment via search engines. This type of recruitment is only suitable in long-term studies, though, because it usually takes several weeks for the search engines to become aware of the new web pages and list them. Two meta tags to be used are shown in the following example of a fictitious study targeted at people with protanopia, a particular subtype of red–green colour blindness.

```
<HTML>

<HEAD>

<META NAME=‘keywords’ CONTENT=‘protanopia, red-green colour blindness, survey, experiences, psychology, research>

<META NAME=‘description’ CONTENT=‘research psychologists invite people with protanopia (a particular variant of red-green colour blindness) to complete a survey of their experiences that we hope will contribute to understanding this condition’

<TITLE>Survey of persons with protanopia</TITLE>

</HEAD>

<BODY>

(Further information on the study and a link to the study would be placed here)

</BODY>
```

Compared with other methods of recruitment, search engines are usually much less effective. This also applies to banner advertisements, which recently are often linked with search engines. Because banners usually represent commercial advertising, a banner ad for a scientific study is hard to distinguish from a deceptive link to a commercial message. Tuten *et al.* (2000) found banner advertisement to be ineffective in web research.

Offline recruitment

An alternative to any type of Internet-based recruitment is traditional recruitment offline. If an experiment is conducted both in the laboratory and on the Web, random allocation to the mode condition is essential for acceptable mode comparisons.

Multiple site entry technique

Depending on the chosen recruitment strategy the sample will be more or less heterogeneous and more or less known. In some cases, experimenters may prefer a homogeneous sample, for example to reduce error variance, or to better be able to compare results with past research (*ca*. 80 per cent of samples in psychological research are made up from young psychology undergraduates: Reips 2000). In other cases, sample composition may simply be not important to experimenters, or only as a distant boundary condition to generalizability. A heuristic strategy that can be used to estimate the influence of sampling and self-selection effects on results is the *multiple site entry technique* (Reips, 2000, 2002c). One implements the technique by advertising the Web experiment in different locations on the Internet or offline. Each link is formed differently, to later allow the separation of sources during data analysis. For example, the URL of the Web experiment could be amended by '… index.html?source=studlist' in the version that is sent to a student mailing list, and by '… index.html?source=sport' if a link is placed on a website dealing with sports. In the resulting data file there will be a column 'source', in which each participant has one entry 'studlist' or 'sport' or else.

Monitoring a web experiment

Once a web experiment is up and running and participants have been recruited it is important to monitor the study. For example it is necessary to note one's own test accesses to the website in order not to let these contaminate the data. Ways of doing this are:

- ♦ Record ones own IP address each time one logs into the experiment, unless one has a static IP
- ♦ Record the times of one's test accesses
- ♦ Whenever one logs on, one enters 'TEST' or similar into one of the text fields on the web pages belonging to the Web experiment.

With all of these methods, the critical data sets can be identified during preparation of data analysis.

Free tracking services are available for monitoring site traffic, and with a little bit of installation (p. 385) work of freeware this can be done on one's web server too. Monitoring site traffic allows the researcher to detect sudden changes in the frequency of visits to the Web experiment. For example, a press note or a link from a highly frequented site may increase number of visits, and technical problems may be the reason for a sudden decrease in traffic. Detailed monitoring of site traffic allows further diagnoses. Via the so-called 'referer' information it can be determined how much traffic came from which website linking to the experiment. Tracking services sometimes offer visualization of location of participants. Figure 24.5 shows a participant's location, visiting time, provider, type of connection, operating system, type of web browser, screen resolution, screen colour depth and Javascript status on a map, here near the White House in Washington, DC. At maximum resolution of the zoomable map the picture would show two buildings.

Data analysis

Data handling and analysis in Internet-based experimenting requires particular attention to three phenomena: *requests and entries not coming from human participants, technical variance*, and *high rates of non-response*. These are never or only rarely seen in offline studies and may therefore easily be overlooked by experimenters untrained in Internet-based research.

Inclusion criteria

In psychological offline studies it would be regarded as highly unusual to consider data entries coming from other than human participants. At most, an experimenter needs to be aware of test trials by staff members and take appropriate steps to keep those from contaminating the data. In Internet-based experimenting, the latter remains an issue, and the former requires attention as well. In many web-based studies the majority of unique requests and entries do not come from human participants, but from bots (spiders, crawlers or robots)[7] mentioned above, including bots from search engines or those looking for specific exploits in a web server. The prevalence of bot requests has increased from *ca*. 6 per cent in 1996 to *ca*.65 per cent in 2005, and the number of different types of robots has increased almost eightfold (Reips, 2007).

Visits generated by bots, test visits by experimenters, and other undesirable entries in the data file need systematic attention before the analysis process, resulting in the following check list of inclusion criteria:

1. no visits by experimenters
2. no visits generated by bots
3. no multiple visits from same IP address (e.g. exclude every entry after first one)
4. if identification of log entries by one participant depends on IP addresses only, then it is recommended to exclude entries from Internet providers with dynamic IP addressing (Reips 2002c)

5. no ‘snoopers’, only include those that make it beyond the experiment's start page (or are assigned to an experimental condition)

6. only those that answer a certain number/percentage of items or those items that are essential to the study.

The respective entries are marked and not included in the main data analysis. Because the number of cases is often substantial, it should become a standard procedure to report figures for types of in- or exclusions at the beginning of the Results section in articles about Internet-based research (Reips 2002c).

Checking for technical variance

Figure 24.5 Visualization of participant location (left of information box) and additional information offered by tracking services for website traffic.

Technical variance is built into Internet-based research (Reips, 2000; Schmidt, Chapter 29, this volume) and may influence results if it is overlooked (Reips 2002b). Conducting experiments over the Internet involves a worldwide network of cables, routers (computers at cross-points of Internet traffic), plugs, and participants use a wide range of computers, monitors, keyboards, mice, speakers or ear phones, web browsers and net connections. Technical variance is not necessarily of disadvantage, its impact depends on the study design (Reips 2000, 2002b, 2002c, see also Chapter 25 and 29, this volume). In an (p. 386) (p. 387) Internet-based experiment tight experimental control and the associated dependency on a fully functional and unbiased equipment available in a laboratory are replaced by random technical variance, and therefore generalizability of the resulting effects is increased, if the error noise allows for the detection of effects (Reips, 2000). In support of the trade-off, Krantz (2000, 2001) has shown effects of technical variance in computer monitors and alerted the scientific community to how choice of mouse may affect response timing in psychological studies.

Incompatibilities with particular technologies may systematically bias the sample. Buchanan and Reips (2001) showed that technology used in the implementation of an Internet-based study may interact with demography or personality. In their study they showed that the average education level was higher in a web-based test if no JavaScript was used, and that Mac users scored significantly higher on Openness than PC users. Consequently, one of the first standard procedures in data analysis in Internet-based research is to compare dropout and values on dependent variables for the different types of technologies the HTTP protocol provides information for. These are the type of operating system, type and version of web browser and connection speed. Additional information can be collected via JavaScript or Java, e.g. JavaScript compatibility, screen resolution, colour depth. These control analyses for technical variance can be combined in one analysis step with those from the multiple site entry technique (see above).

Dropout analysis

Internet-based research in general shows high rates of non-response. These may therefore be used as dependent variables (Reips 2000, 2002b, 2002c).

A dropout analysis is typically carried out with all included data sets plus those mentioned in inclusion criterion 6 above. The point of a dropout analysis is to see where and how many participants left the study (e.g. for use as a dependent variable) and to find out whether the dropout may be systematic. If dropout coincides with an experimental manipulation, then motivational confounding may be at work and the study is severely compromised (Reips 2000, 2002b), see Figure 24.3 above.

What dropout rate to expect? Musch and Reips (2000) found that the average dropout rate in Internet-based experiments is 34 per cent (median = 35 per cent), with a range from 1–87 per cent. The large range suggests that there are other motivational factors that influence dropout, some of which were since examined empirically. Musch and Reips showed that the completion rate of Internet-based experiments was 86 per cent if some form of reward (individual payments or lottery prizes) was offered, up from 55 per cent without such rewards. In a follow-up experiment Frick *et al.* (2001) tested the influence of two variables on dropout: placement of assessment of personal information and information about financial incentives for participation on dropout rates. They found the following dropout rates: 5.7 per cent for participants who were asked personal information and information about financial incentives at the beginning of the experiment, 21.9 per cent in the condition with both at the end. Dropout rates in the other conditions were in-between (13.2 and 14.9 per cent).

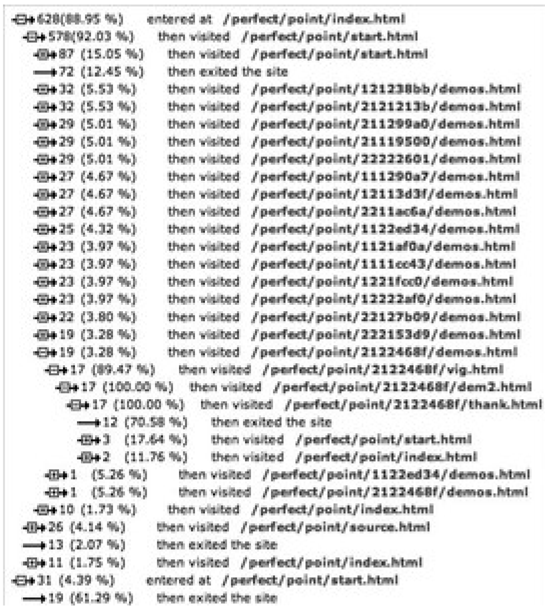

Figure 24.6 A dropout tree created with Scientific LogAnalyzer (Reips and Stieger 2004).

Dropout can be analysed and reported in four ways that differ in information: overall, as rates, and curves, and as trees (Reips 2002b, Reips and Stieger 2004). *Overall* dropout is the percentage of total participants in a study that did not complete it. Usually this information will not suffice in manuscripts submitted to good journals where editors and reviewers are aware of the high rates of non-response in Internet-based research. Dropout *rates* provide summary information by experimental condition, just as in the example above this paragraph. Often this level of information will suffice in reports. Figure 24.3 shows the visually most informative way of reporting dropout, as *curves*. Dropout curves provide information by condition about dropout over time and—if these are marked—in relation to particular events during the experiment. Finally, dropout can also be reported in a *tree* format, see Figure 24.6. Such dropout trees are generated by the Scientific LogAnalyzer (Reips and Stieger 2004), online they can interactively be manipulated by mouse-clicking the '−'− and '+'−squares to expand and collapse the tree's branches. A branch shows the absolute numbers of participants, and their proportion relative to the mother branch, that chose to visit the web page. **(p. 388)**

Archiving of web experiments

One of the chief advantages of Internet-based experimenting is the public visibility and availability of the materials. Reips (2002c) defines public archiving as one of sixteen standards for Internet-based experimenting: 'The experimental materials should be kept available on the Internet, as they will often give a much better impression of what was done than any verbal description could convey.' (p. 254). Archived web experiments may serve readers of publications as helpful illustration, given that the web address is provided in the publication. Furthermore, archived web experiments may be used as teaching materials for students in psychology and related sciences. Finally, materials that are available online can be used by researchers to easily create variants of experiments, significantly speeding up the research cycle (Reips 2000; Birnbaum 2004).

Internet-based experiments can be archived in the Web experiment list that was earlier described as a website for the recruitment of participants. The site contains a search function that will generate lists of web experiments by field, (p. 389) status (active versus archived), time, language (English or German), and combinations thereof.

Conclusion

It has been noted that Internet-based experimenting is neither similar to nor different from laboratory experiments (Reips 2002c). There are a number of particular methodological issues that may be general or associated with the tasks of planning, generating, pre-testing, recruitment, monitoring, analysis, and archiving of an experiment. Some of the issues are related to technology and its proper use (see Chapter 29, this volume), some to experimental design (see Chapter 25, this volume) others are due to the psychological effects of various aspects of the vast social sphere that continues to build up: the Internet.

Hopefully, this chapter provides a helpful collection of techniques, recommendations, and pointers to those who will soon conduct their first Internet-based experiment, and to those who continue to learn about research on the Internet.

References

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

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

Birnbaum, M. H. and 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, NJ: Erlbaum.

Bohner, G., Danner, U. N., Siebler, F. and 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*(2), 149–151.

Bosnjak, M. (2001). Participation in non-restricted web surveys: a typology and explanatory model for item non-response. In U.-D. Reips, and M. Bosnjak, (eds), *Dimensions of Internet Science* (pp. 193–208). Lengerich: Pabst Science.

Brenner, V. (2002). Generalizability issues in internet-based survey research: implications for the internet addiction controversy. In B. Batinic, U.-D. Reips and M. Bosnjak (eds.), *Online social sciences* (pp. 93–114). Seattle, WA: Hogrefe and Huber.

Buchanan, T. and Reips, U.-D. (2001). Platform-dependent biases in online research: do Mac users really think different? In K. J. Jonas, P. Breuer, B. Schauenburg and M. Boos (eds), *Perspectives on internet research: Concepts and methods*. Retrieved 27 December 2001 from **http:\\server3.uni-psych.gwdg.de/gor/contrib/buchanan-tom**.

Buchanan, T. and Smith, J. L. (1999). Research on the internet: validation of a world-wide web mediated personality scale. *Behavior Research Methods Instruments and Computers 31*(4), 565–571.

Coomber, R. (1997). Using the Internet for survey research. *Sociological Research Online 2*(2), 14–23.

Erlanger, D., Feldman, D., Kutner, K., Kaushik, T., Kroger, H., Festa, J., *et al*. (2003). Development and validation of a web-based neuropsychological test protocol for sports-related return-to-play decision-making. *Archives of Clinical Neuropsychology 18*(3), 293–316.

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

Geissler, H., Loewe, G., Göritz, A., Irmer, C., Baierl, M., von Heesen, B., Schwabl, T., Wagner, F. and Knapp, F. (2005). *Current relevance and future of online access panels*. Panel discussion at 7th General Online Research (GOR) Conference, March 21–23, Zürich Switzerland. Available at **http://gor.de/gor05/podiumsdiskussion_en.htm**.

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

Horswill, M. S. and 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.

Joinson, A. (2003). *Understanding the psychology of Internet behaviour: Virtual worlds, real lives*. Houndmills: Palgrave Macmillan.

Krantz, J. H. (2000). Tell me, what did you see? The stimulus on computers. *Behavior Research Methods, Instruments and Computers 32*, 221–229.

Krantz, J. H. (2001). Stimulus delivery on the web: what can be presented when calibration isn't possible. In U.-D. Reips, and M. Bosnjak, (eds.), *Dimensions of internet science* (pp. 113–130). Lengerich: Pabst Science.

Krantz, J. H. and Dalal, R. S. (2000). Validity of web-based psychological research. In M. H. Birnbaum (ed.), *Psychology experiments on the internet* (pp. 35–60). San Diego, CA: Academic Press.

Laugwitz, B. (2001). A web-experiment on colour harmony principles applied to computer user interface design. In U.-D. Reips and M. Bosnjak (eds), *Dimensions of internet science* (pp. 131–145). Lengerich: Pabst Science.

Musch, J. and 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.

(p. 390) 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.

Plant, R. R., Hammond, N. and Whitehouse, T. (2003). How choice of mouse may affect response timing in psychological studies. *Behavior Research Methods, Instruments and Computers 35*(2), 276–284.

Reips, U.-D. (2007). Collecting data in surfer's paradise: internet-mediated research yesterday, now and tomorrow. Manuscript submitted for publication.

Reips, U.-D. (2006a). Internet-basierte Methoden [Internet-based methods]. In F. Petermann and M. Eid (Hrsg.), *Handbuch der Psychologischen Diagnostik* (pp. 218–225). 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 and M. Bosnjak (eds.), *Online social sciences* (pp. 69–79). Seattle, WA: Hogrefe and Huber.

Reips, U.-D. (2002b). Internet-based psychological experimenting: five dos and five don'ts. *Social Science Computer Review 20*(3), 241–249.

Reips, U.-D. (2002c). Standards for internet-based experimenting. *Experimental Psychology 49*(4), 243–256.

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

Reips, U.-D. (2006b). Web-based methods. In M. Eid and E. Diener (eds.), *Handbook of multimethod measurement in psychology* (pp. 73–85). Washington, DC: American Psychological Association.

Reips, U.-D. and 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. and 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. and Stieger, S. (2004). Scientific log analyzer: a web-based tool for analyses of server log files in psychological research. *Behavior Research Methods, Instruments and Computers 36*(2), 304–311.

Rodgers, J., Buchanan, T., Scholey, A. B., Heffernan, T. M., Ling, J. and 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*(8), 619–625.

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. (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.

Schulte-Mecklenbeck, M. and Neun, M. (2005). WebDiP: a tool for information search experiments on the worldwide web. *Behavior Research Methods 37*(2), 293–300.

Schwarz, S. and Reips, U.-D. (2001). CGI versus Javascript: a web experiment on the reversed hindsight bias. In U.-D. Reips and M. Bosnjak (eds.), *Dimensions of internet science* (pp. 75–90). Lengerich: Pabst Science.

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

Notes:

(1) The term 'Internet-based experiment' is largely synonymous with the widely used terms 'Web experiment', 'Internet experiment', 'on(-)line experiment', 'web-based experiment', and 'WWW experiment', because experiments using Internet services other than the WWW (such as email, Telnet, ICQ, Gopher, FTP etc.) are rarely conducted. Because most experiments on the Internet are based on web pages the term 'web experiment' is used most often in this chapter.

(2) http://genpsylab-wexlist.unizh.ch/.

(3) Currently, some of these may pass the review process, because many reviewers are no exemption to the situation, which in many ways parallels the time when statistical testing was introduced to research.

(4) In routinely checking requests for enlistment of studies on the Web experiment list, the Web survey list, and the Web Experimental Psychology Lab the author frequently encounters instances of these configuration errors. Most recently, he was able to download a data set that included personal information from 65,000 participants in a study. These data, including email addresses, were also accessible to spammers.

(5) Mac OS X is currently, and has been, particularly comfortable and safe for setting up Internet-based research.

(6) The source code of a web page can easily be viewed by choosing the respective option on the 'View' menu in most web browsers.

(7) Examples are msnbot, googlebot, spider.search.ch, npbot, f-bot testpilot, jetbot, turnitinbot, psbot, HenryTheMirageRobot, Gaisbot, Nutch-bot, Exabot, WISEnutbot, itbot.

Ulf-Dietrich Reips

Ulf-Dietrich Reips, Department of Psychology, University of Zurich.