Reips, U.-D. (1996, October). Experimenting in the World Wide Web. Paper presented at the 1996 Society for Computers in Psychology conference, Chicago.
Current e-mail address: ureips@genpsy.unizh.ch

Experimenting on the World Wide Web

Ulf-Dietrich Reips

University of Tübingen, Germany

Contact: Ulf-Dietrich Reips
Abteilung für Allgemeine und Entwicklungspsychologie
Psychologisches Institut
Friedrichstr. 21
D-72072 Tübingen
Germany

Phone: +49-7071-2978350

e-mail: ulf.reips@uni-tuebingen.de

Running head: WEB EXPERIMENTATION

Abstract

Web experiments, experiments conducted on the World Wide Web (WWW), are a new tool in experimental research. Web experiments differ fundamentally from laboratory and field experiments traditionally used in behavioral science. This article describes which hard- and software components are needed to set up a web experiment, and how this was realized in a “virtual psychology laboratory”. In addition, it contains a thorough methodological discussion of web experimentation which highlights advantages of this method. Web experiments offer for example (1) easy access to a geographically unlimited subject population, including subjects from very specific and previously inaccessible target populations; (2) bringing the experiment to the subject instead of the opposite; (3) high statistical power through high sample size while keeping a conventional a-level; and (4) reduced cost, because neither laboratory rooms nor experimenters are needed. Problematic aspects of web experiments are discussed as well, and solutions for minimizing them are offered.

The World Wide Web (WWW, or web), which is the graphical part of the internet, has been described as a source of information for psychologists (Kelley-Milburn & Milburn, 1995). A further application, which I will try to lay out here, is the use of the WWW for data collection, specifically, for conducting experiments on the WWW. As will be shown, the rapid development and wide popularity of the WWW offers the prospect of an entirely new class of methodology which might help overcome some of the fundamental limitations of psychological research at relatively low cost.

Traditionally, experimental methods can be categorized into laboratory and field experiments[1]. They come with certain theoretical and practical disadvantages which limit their power of explanation in a basic way. The web experiment, although limited in its own way, could be a method that leads the way out of the deadlock between an ecologically invalid situation (in the lab) and no control of potentially relevant factors (in the field). Additionally, it might allow for research in previously inaccessible areas and for validation of findings in the existing body of research.

What is a "Web Experiment"?

Since personal computers (PCs) became affordable and relatively easy to use, their use as devices in laboratory experiments has increased and evolved to a standard level. Usually the network features of these PCs were not used for experimental purposes, however, and if they were, then mostly in local area networks (LANs). Web experimentation at its core is the logical extension of laboratory experiments with PCs: A subject connects from her PC via the graphical interface of her WWW client software, a so-called web browser[2], to the laboratory PC or web server to participate in the experiment. The experiment runs much in the same way

as with a subject sitting in front of the laboratory PC, only that everything the monitor screen of the laboratory PC would show is instead transferred onto the monitor screen of the subject's PC which can be anywhere in the world. Any input from the subject, given via mouse clicks, text entries, or document requests, can be recorded by the web server and be responded to as predefined by the experimenter. Even response times can be recorded[3], along with information such as name and location of the subject's computer and type of web browser software used. With recent developments, even sound and video signals or mouse movements can be used as subject input.

Hard- and Software Requirements

A single PC with a permanent connection to the internet can be used as a "web server". It is recommended to have a second PC "mirror" the web experiment to allow for continuous online presence in case of technical problems with the first PC. Note that the nature of the bandwidth (data transfer capability, measured in kbit/s) of the connection to the next main internet node is generally much more important than the processing power of the server computer. A minimum bandwidth of 64 kbit/s is desirable if a web experiment contains mostly text and a few medium-sized graphics[4].

The subject needs a web browser such as Netscape Navigator (Netscape Communications Corporation, 1996). Most web browsers create a cache on the user's hard disk, in which all recently viewed files are saved. If the user requests a file again, for example a picture used in a web experiment on learning, then it will be loaded from the cache, allowing a much speedier display than when loading over the network.

The formatting of WWW documents is done in hypertext markup language (html). Html is platform independent, which means it can be viewed on any type of computer with any kind of operating system, as long as there is a web browser for that computer. WWW pages can be created and saved in html with standard word processing programs. For refinement and special functions a html editor program is required.

Essential software components for conducting experiments on the WWW are a web server program and Common Gateway Interface applications (CGIs). The web server delivers the web pages to the subjects and creates a log file of all document requests and other subject inputs. CGIs are needed for randomization, subject input driven responses, and server access to databases. The most simple setup of a web experiment is a CGI which randomly assigns subjects to treatment conditions and a web server which provides the experimental material (instructions, tasks, questionnaires, etc.) to the subjects (see Figure 1). There exist inexpensive "shareware" or even free applications of all needed software components. They can be found on the internet, for example at Bob Allison's Web Masters' Page (Allison, 1996), or by using one of the common WWW search engines[5].

Insert Figure 1 about here

A list of experiments currently conducted on the WWW can be found at the web site of the American Psychological Association (Krantz, 1995).

Methodological Issues

Representativeness

The internet user population is rapidly expanding. Various surveys show that all population parameters are converging with those of the general population (Graphics, Visualization, & Usability Center, 1996). Sometime in the near future the internet user population will probably be as representative of the general population as the population of television users. The powerful implication of this development for psychological research comes from particular characteristics of the internet, namely the almost complete loss of any importance of space and time. Communication is instantaneous between any two (or more) points on earth - a fact which has led to the term Global Village. This means that web experiments basically make any person from the general population as accessible to the researcher as the "usual subjects" from the campus cafeteria.

It is a well-known fact that most research in psychology is done with college students. If the major reason for not using subjects other than college students is the students' availability to the researcher, then web experiments should be considered as an alternative. Replication of past findings using web experiments will allow for generalization beyond the campus population.

Sampling issues remain as problematic in web experiments as in traditional research. Self-selection will be a serious methodological problem in those cases where aspects of the experiment are confounded with the subjects' decision to take part in the experiment. However, if it can be determined that confounding did not take place, then we might be able to generalize results from web experiments to a larger extent than in traditional research.

Presently, as long as there is not much data about the external validity of web experiments, the design type should be factorial to guarantee a maximum internal

validity. The selected area of research is recommended as being likely to be immune to sampling effects. Figure 2 shows an example design of a web experiment conducted at the author's Web Experimental Psychology Lab at the University of Tübingen, Germany (Reips, 1996), which follows these standards.

Insert Figure 2 about here

A second major aspect of representativeness is the intercultural issue: Participants in web experiments come from various cultures around the world. Therefore, web experiments offer a promising potential for cross-cultural studies.

To connect findings from research using the emerging web experiment method with the existing body of research, it will be necessary to have subsamples of subjects participate in the traditional laboratory setting. Whether there will be differences in findings between these subsamples and the WWW user samples, and to which features of traditional or web experimental methods they can be attributed, will be an interesting area of research in itself.

Ecological Validity

Ecological validity of laboratory experiments is generally low (Chapanis, 1970). Partially, this is due to the subject's perspective of being in an uncontrolled and unfamiliar situation.

Web experiments create a situation in which the experiment comes to the subject instead of the other way around. Subjects can be in their usual surrounding, let us say at their desktop at home or at work - the only requirement is that they have access to a computer connected to the WWW. The way current technical

development is headed, it will soon be common to connect to the WWW with wireless laptop and palmtop computers.

Participants in web experiments can freely choose at which time of day (or night) and on which day they wish to participate. Institutional regulations (e.g., lab hours that are limited to non-lunch daytime during weekdays) lose their limiting impact on research. It is possible that several subjects participate individually at the same time - there are no scheduling problems and some web server programs allow up to 1,000 simultaneous connections.

The situation in a web experiment is a special situation in itself, however. The dependency on computers and networks creates an experimental situation which limits generalization.

Motivation and Drop-out

With subjects being able to terminate their participation at any time during a web experiment, motivational issues become much more important than in laboratory experiments. On the one hand, in web experiments subjects will not stay in the experimental situation for reasons other than their true will, for example course credit considerations or “peer pressure” by the experimenter, thereby creating less error variance. On the other hand, a selective drop-out in one condition would make the experiment in itself worthless, but in turn would help in detecting unwanted confounding with such variables as dissimilar difficulty of the tasks in the treatment conditions. In short: Aside from ethical considerations, the complete voluntary character of participation allows for detection of artifacts, which are based on different motivational implications of the material used in the experimental conditions. Subjects in a less motivating or boring condition, who in a laboratory

experiment would have stayed due to, let's say, course credit considerations, will drop out of a web experiment with a much higher chance. In this case the laboratory experiment data would be contaminated by motivational confounding, the web experiment data would allow for detection of this effect.

Of course, web experiments allow for experimenter selected participation of subjects too. A number of subjects can be selected from a subject pool, for example psychology undergraduates at a certain university, and then be asked to participate in the web experiment. This would allow for control of the motivational effects of mandatory participation while keeping constant all other aspects. It will be interesting to see whether the results would differ from findings in research on the volunteer bias (Rosenthal and Rosnow, 1975).

To reduce drop-out in web experiments, a number of measures aside from pre-selecting subjects can be taken. These are, for example, rewards for participation, attractive web design, interesting cover stories, high data transmission speed, and a warm-up period with the same stimuli for all conditions.

Control

A question one might ask is "How do I know that a subject did not participate repeatedly?" Several measures can be taken to avoid data contamination. First, subjects should be asked to provide their electronic mail (e-mail) address or a phone number, which allows identification and might be used to check back with a random subsample of the subjects later. Secondly, one could analyze only the first data set from one computer address (the web server always writes the subjects' computer addresses in a log file). Thirdly, with random placement of subjects in the treatment conditions, any errors will be distributed evenly over the conditions. Generally, it

seems rather unlikely that people would try to "cheat" in web experiments, especially whenever participation requires considerable effort on the part of the subject.

To control for other possible artifacts, one can use a variety of techniques depending on the hypothesis tested. For example, in a learning experiment one might want to exclude data from subjects who took a break during the experiment. A break could be defined as an increase of more than a certain percentage of the average inter-trial time this subject needed. As mentioned above, many web servers log the exact times of document retrieval.

Technical variance (different computers, monitors, browsers, net connections) will contribute to error variance. Usually one wants to minimize error variance in controlled laboratory situations to be able to get to the "pure" relationship between independent and dependent variables. More error variance can also be an advantage, however, if it means replacing potentially undetected sources of systematic error with random error. Let us say something is wrong with a laboratory computer's monitor which interferes with the experiment - it will systematically influence the results. This could not happen in a web experiment, in which every subject is using a different monitor.

A further potential source of error is experimenter bias (Rosenthal, 1976). No experimenters are needed in web experiments.

Pragmatic and Financial Considerations

As has recently been mentioned in this journal (Erdfelder, Faul, & Buchner, 1996), power values in psychological studies remain sadly low. Erdfelder et al. remind us that in a study "there are only two ways to raise the power if the null hypothesis (H_0), the alternative hypothesis (H_1), and the test statistics have already

been specified: One must increase either the sample size N or the Type I error probability a." (p. 2). For traditional laboratory experiments, often there will be no other way than to increase a, due to pragmatic constraints such as time, lab space, and financial reasons. In web experiments, however, it is no problem at all to get to the "ideal" calculated sample size while keeping a conventional a-level. The subject pool is almost unlimited in size.

It has been mentioned before that web experiments run around the clock and allow for simultaneous access by a large number of subjects. This is only feasible because no experimenter needs to be present during the experiment. Once a web experiment is programmed and on the WWW, no salaries have to be paid for experimenters, and consequently, as has been mentioned before, experimenter bias is less likely as well. On the downside, subjects can not ask questions regarding the comprehension of the material. Therefore, web experiments have to be programmed to be clear and the instructions have to be easy to comprehend. Also, subjects should be encouraged to mail commentaries right after participating, to allow for clarification during the pilot phase and for further experiments.

Web experiments are financially attractive in even more ways: no rooms for laboratories are needed, and no bureaucracy regarding scheduling, insurance, and so forth. In its simplest form, web experiments can be run from an office PC, which may be used in other ways, too. On the other hand, some money for the transmission of the data has to be paid. Nevertheless, web experiments are much more cost effective than traditional experiments.

New Possibilities for the Process of Psychological Research

The new experimental method web experiment allows research in areas which were almost inaccessible for established methods. One example are studies with subjects from very specific target populations (e.g. persons age 20-30 with diabetes, an IQ higher than 130, and a traumatic crime experience). In the past it would have required an enormous effort to find and contact people who had the desired characteristics, and an even stronger effort to get all those people into the laboratory or at least interviewed. Through the worldwide access, studies like the one mentioned could be done almost as equally easily as with standard sample populations.

Web experiments, which are publicly accessible and could remain on the WWW for documentation purposes, allow a clearer insight by other researchers into how the materials looked. Direct hypertext links to articles in online journals like Psycoloquy would provide the reader through a mouse click with the possibility to experience the described experiment from the perspective of a subject. Interactive editing of the materials will allow fast setup of web experiments to test alternative hypotheses. The data will be accessible online and open for reanalysis (of course after having taken care of protecting subjects' confidentiality through erasing or anonymizing personal information etc.). All stages of psychological research will become more public.

An Experimental Psychology Laboratory on the WWW

The "Web Experimental Psychology Lab" (Reips, 1995) was established in summer 1995. After an initial testing phase it was announced to several search engines on the WWW. From about the beginning of 1996 a substantial increase in the number of server accesses could be observed. Currently the number of hits is

around 400 per day, which translates into about 8 to 20 subjects with the current experimental setup. It seems easy to increase these numbers by placing more announcements to the WWW search engines.

The laboratory is bilingual: So far all experiments were run in English and in German. The entry page of the lab contains several slots with hyperlinks to the instructions of the experiments currently running. Care was taken in the construction of graphically appealing design of the web pages while keeping the files limited in size. A boring layout and long loading times are detrimental to the motivation to read web pages.

One experiment conducted in the Web Experimental Psychology Lab is an experiment on the acquisition of causal knowledge. Because this is a paper on the methodological implications of web experimentation I will solely describe the experiment's implications for this purpose. Figure 2 shows the design of the experiment. People who visited the web site and followed the hyperlink to this experiment were instructed about its scientific nature and the approximate time it would take for completion (20 to 40 minutes). Also, the first page contained a test item to find out whether the subject's web browser would allow for selection menus, and a file transfer protocol (ftp) hyperlink for immediate downloading of a suitable web browser. As a last "hurdle of commitment" subjects were asked to give their e-mail address or phone number.

Persons who decided to take part in the experiment were randomly distributed to one of two conditions for balancing of the names of items. Otherwise, all subjects were in the same learning condition during the first phase of the experiment. Therefore the expected drop-out of people checking in out of curiosity mostly

happened before the second, decisive, learning phase began. Always, submitting the response to a trial was followed by the display of a WWW page containing feedback and a new learning item. The pages were chosen by a CGI according to predetermined probabilistic rules.

Upon entering the second learning phase the subjects were randomly placed in one of the two experimental or the control conditions. As in the first learning phase, subjects were asked to go through at least 20 trials before entering the next phase. Learning items were displayed verbally and graphically. Due to the web browser's cache feature, repeated use of the same graphics allowed for instant display on the subject's computer monitor. The experiment was completed when subjects had answered two consecutive questionnaires regarding the learning material and one demographical questionnaire. This last web page also contained questions about the subject's computer equipment, the subject's perception of the connection speed, whether the subject took notes or used the web browser's "back" feature, and whether the subject participated alone. Upon completion subjects saw a "Thank you" note and were given the opportunity to send a comment via e-mail.

Between mid-January and the end of August 1996, 360 subjects participated in this experiment and entered the second learning phase. 61 of these were local subjects who were recruited in the usual way. 60 of the 299 "web subjects" indicated on the demographical questionnaire that they were "insiders". About one fifth of the remaining subjects claimed to be female (45), and about 30% (68) chose to participate in the English version of the experiment. The web subjects' age distribution was wider than the one of the local subjects (see Figure 3).

Insert Figure 3 about here

This, and the fact, that subjects from more than twenty different countries participated, is an indication of the WWW's wide user pool - a rich potential for data collection.

Discussion

Web experiments offer many advantages to the scientist that come in two sweet flavors: Some are methodologically, some are practically promising. Current psychological theories should be able to predict subject behavior in web experiments if they are valid.

Some of the possible disadvantages of web experiments can be avoided by taking appropriate measures, for example limiting data contamination from multiple participation by analyzing only the first data set from one computer address or limiting drop-out during the experimental phase by having subjects go through a warm-up period.

The advantages of having a large world wide pool of subjects always right at one's fingertips are extremely attractive. It seems possible to overcome some of the objections against the traditional experiment. Data from the emerging field of web experimentation might help us to be able to generalize much better than ever across demographic, cultural and numerical boundaries - and this while saving money.

A final point for anyone hesitating to explore the rich possibilities of web experimentation: An experiment programmed for the WWW can always be used in a local laboratory too.

Figure Captions

Figure 1. The basic setup of a web experiment.

Figure 2. The flow chart of a factorial web experiment on causal probability learning.

Figure 3. Age distribution of web subjects and local subjects.

References

Allison, B. (1996). Web master's page [On-line]. Available: http://gagme.wwa.com/~boba/masters1.html

Bortz, J. & Döring, N. (1995). Forschungsmethoden und Evaluation [Research methods and evaluation]. Berlin, Germany: Springer.

Chapanis, A. (1970). The relevance of laboratory studies to practical situations. In D. P. Schultz (Ed.), The science of psychology: Critical reflections. New York: Appleton Century Crofts.

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

Graphics, Visualization, & Usability Center (1996). GVU's WWW user surveys [On-line]. Available: http://www.cc.gatech.edu/gvu/user_surveys

iWORLD (1996). BrowserWatch [On-line]. Available: http://browserwatch.iworld.com/browsers.shtml

Kelley-Milburn, D., & Milburn, M. A. (1995). Cyberpsych: Resources for psychologists on the internet. Psychological Science, 6, 203-209.

Krantz, J. H. (1995). American Psychological Society: Psychological research on the net [On-line]. Available: http://psych.hanover.edu/APS/exponnet.html

Netscape Communications Corporation (1996). Netscape Navigator 3.0 [Computer software]. Available FTP: Hostname: ftp5.netscape.com Directory: 3.0/ (Apple Macintosh version ends with .hqx, Windows 3.1 version starts with n16, Windows 95 or NT version starts with n32).

Reips, U.-D. (1995). The Web's Experimental Psychology Lab [On-line]. Available: http://www.uni-tuebingen.de/uni/sii/Ulf/Lab/WebExpPsyLab.html

Rosenthal, R. (1976). Experimenter effects in behavioral research. New York: Appleton Century Crofts.

Rosenthal, R. & Rosnow, R. L. (Eds.) (1975). The volunteer subject. New York: Wiley.

Schubert, T., & Waldzus, S. (1996). Sozialpsychologische Werkstatt Jena [Social psychological place of work Jena][On-line]. Available: http://www.uni-jena.de/~ssw/zugang.htm

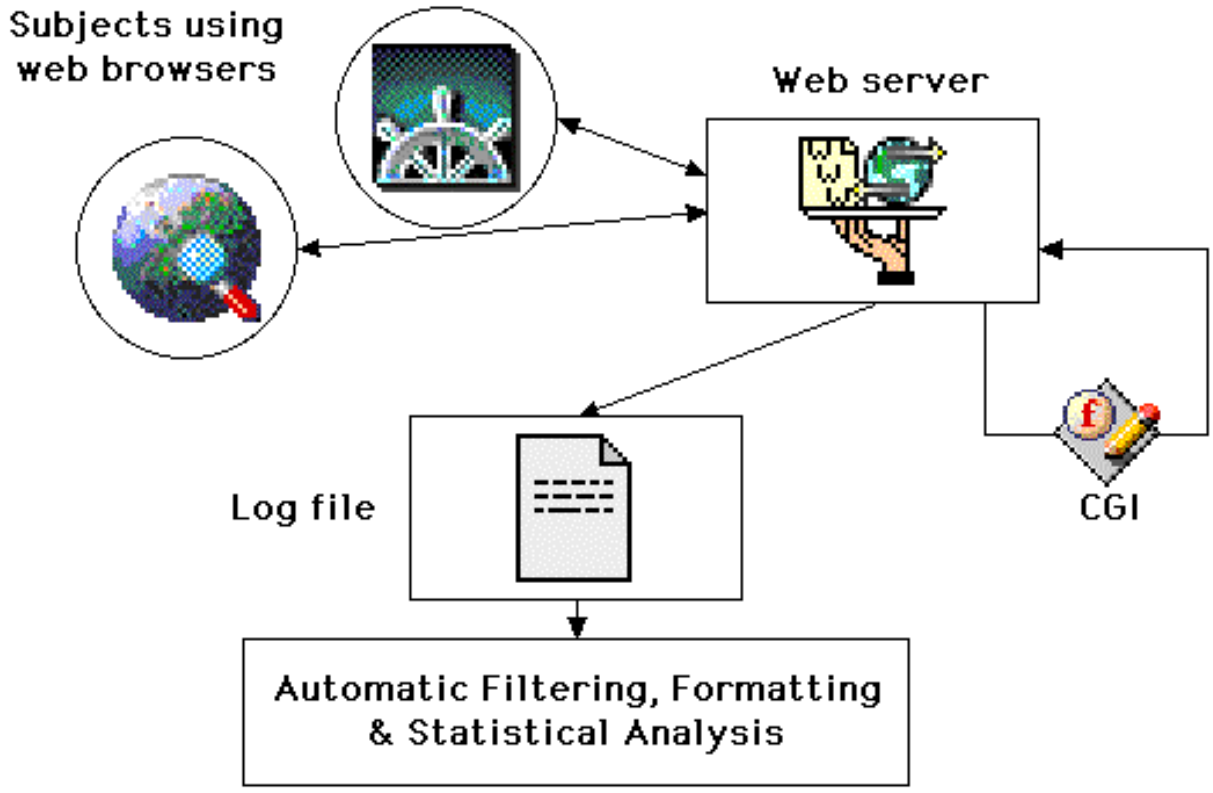
Subjects using
web browsers
Web server
Log file
CGI
Automatic Filtering, Formatting
& Statistical Analysis

Figure 1.

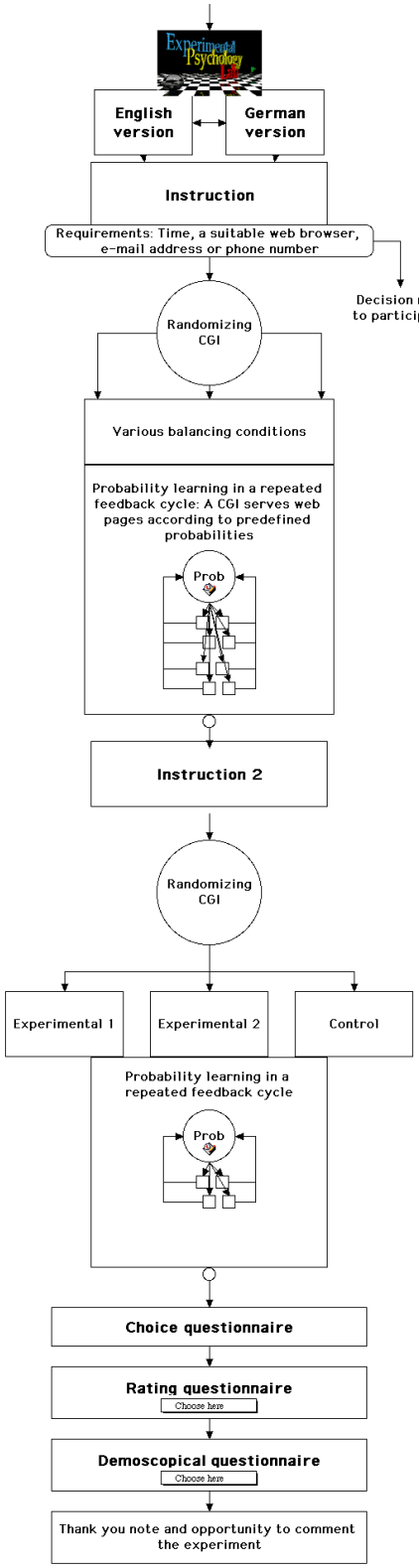
English version
German version
Instruction
Requirements: Time, a suitable web browser, e-mail address or phone number
Decision not to participate
Randomizing CGI
Various balancing conditions
Probability learning in a repeated feedback cycle: A CGI serves web pages according to predefined probabilities
Prob
Instruction 2
Randomizing CGI
Experimental 1
Experimental 2
Control
Probability learning in a repeated feedback cycle
Prob
Choice questionnaire
Rating questionnaire
Choose here
Demoscopical questionnaire
Choose here
Thank you note and opportunity to comment the experiment

Figure 2.

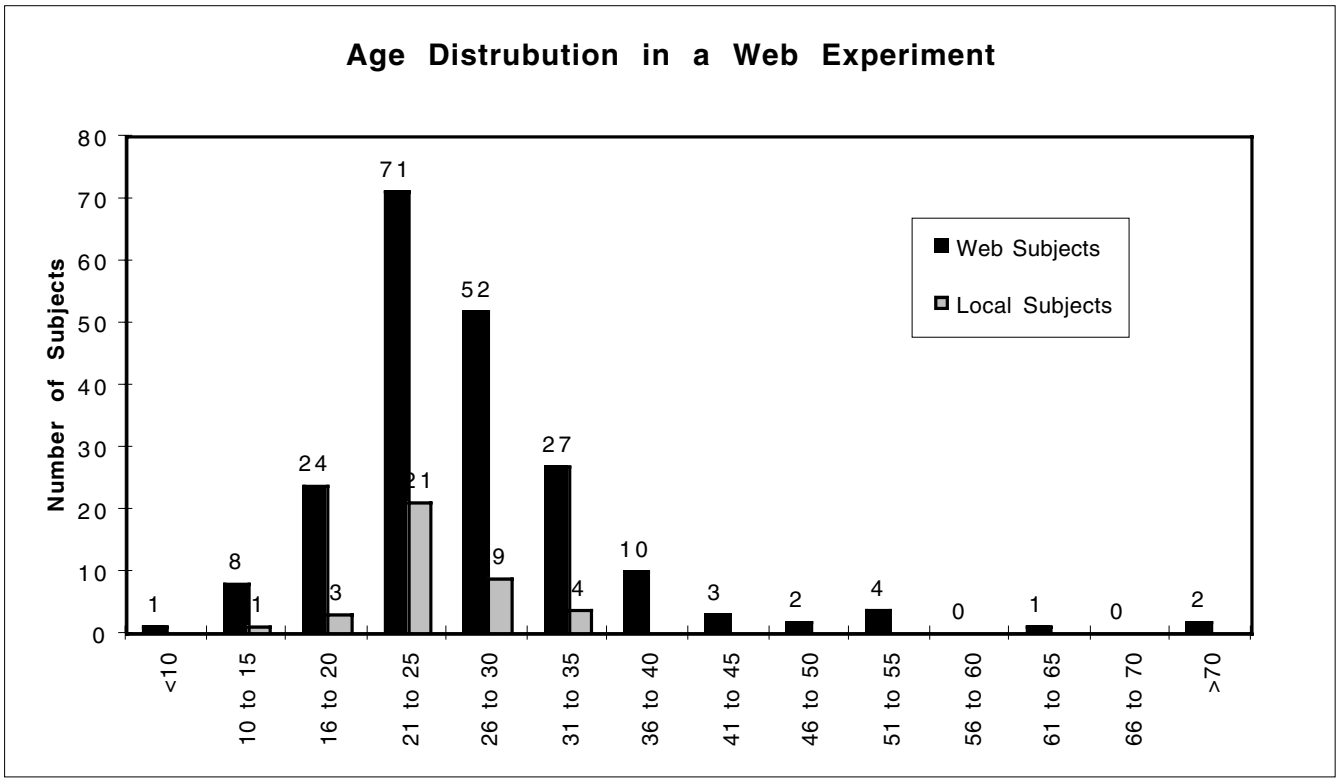

Figure 3.

---

[1]Although field experiments are ideal for maximizing internal and external validity (Bortz and Döring, 1995, p. 57), they are rarely conducted in psychological research. This appears to be due to the relatively high expenditure of field research.

[2]An overview of web browsers can be found (by using one) on the Browserwatch page (iWORLD, 1996). The most widely used web browser can also be obtained via file transfer protocol (ftp) (Netscape, 1996).

[3]Response times in web experiments are prone to error variance from different computers, computer setups, and network lag. However, these error sources are controllable to a certain extent by taking baseline measurements and by programming the experiments as Java applets, stand-alone WWW programs, which can be downloaded by the subject and sent back after data collection. For an example, see Schubert and Waldzus (1996).

[4]For very large pictures, sound, and video data a higher bandwidth is recommended to allow for simultaneous logins by a number of subjects.

---

[5]WWW search engines are web sites with large and permanently updated databases of WWW pages and one or more high speed computers, which answer search requests for terms and concepts.