

The Use of the Internet in Psychological Research: Comparison of Online and Offline Questionnaires

GIUSEPPE RIVA, Ph.D.,^{1,2} TIZIANA TERUZZI, M.S.,² and LUIGI ANOLLI, Ph.D.²

ABSTRACT

The Internet can be an effective medium for the posting, exchange, and collection of information in psychology-related research and data. The relative ease and inexpensiveness of creating and maintaining Web-based applications, associated with the simplicity of use via the graphic-user interface format of form-based surveys, can establish a new research frontier for the social and behavioral sciences. To explore the possible use of Internet tools in psychological research, this study compared Web-based assessment techniques with traditional paper-based methods of different measures of Internet attitudes and behaviors in an Italian sample. The collected data were analyzed to identify both differences between the two samples and in the psychometric characteristics of the questionnaires. Even if we found significant differences between the two samples in the Internet attitudes and behaviors, no relevant differences were found in the psychometric properties of the different questionnaires. This result, similar to the ones previously obtained in Web-based assessments of personality constructs, is even more interesting given the lack of control on the characteristics of the online sample. These findings suggest that, if sampling control and validity assessment is provided, Internet-based questionnaires can be a suitable alternative to more traditional paper-based measures.

INTRODUCTION

SEVERAL RESEARCHERS suggest that the Internet can be an effective medium for the collection and exchange of information in psychology-related research and data.¹⁻³ The Web approach is drawing a large interest in the scientific community because of the advantages it can give to experimental psychology. In a recent study, Musch and Reips⁴ tried to identify the possible pros of this approach by interviewing psychologists previously involved in Internet studies. The main advantages reported are the following:

- the large population access, that allows a greater external validity and the possibility of easily generalizing the obtained results;
- less experimental costs, and the possibility of avoiding all the hassles related to the use of laboratories: booking, limited space, the sharing of equipment, and so on;
- the possibility of providing the tools around the clock, without any time limitation;
- the completely voluntary participation, which usually improves respondents' motivation.

¹Applied Technology for Neuro-Psychology Laboratory, Istituto Auxologico Italiano, Milan, Italy

²Centre For Studies and Research In Communication Psychology, Università Cattolica del Sacro Cuore, Milan, Italy.

As recently noted by Miller and colleagues,⁵ "The Web may prove superior to paper, because it potentially provides increased accessibility; capability for dynamic and interactive forms, which eliminate the viewing of irrelevant questions; and customized feedback tailored to the content of the responses."

In general, the Internet's most attractive factor for psychological research is the opportunity of having a large worldwide pool of subjects always at one's fingertips: the Web offers both a large population range for conducting an experiment, but also the chance to reach a particular sample with specific features.

Since 1997 an increasing number of authors have tried to identify and better define the methodological process required for building an acceptable experimental design. John Krantz and colleagues⁶ started their online experiment on the determinants of female attractiveness, using a within-subjects design. This research was the first psychology Web experiment to be published in a scientific journal.

However, there are specific disadvantages in using the Internet for research:¹⁻⁷

- it is difficult to control the study environment because Web users have different types of hardware, software, and Internet connections. There is no way to ensure that everyone who participates in the experiment will receive exactly the same stimuli in terms of sound, color, or timing;
- study participants are usually unmonitored, so the researcher cannot be sure about the information collected. Members of electronic communities very often adopt false 'nickname' identities or gender switches, and openly accept them in others;
- people who participate in online experiments are self-selected and by no means random representatives of the general population. In particular, they are usually skewed toward the high end of the socioeconomic and educational spectrum;
- the creation of a Web-based assessment tool is not an easy task for a psychologist. It usually requires the development of different Web pages and the administration of the database in which the answers are stored.

As underlined by Buchanan,⁸ "There are clearly also a number of practical issues that should be considered by anyone seeking to use online tests and important theoretical issues that need to be resolved. In conclusion, online clinical tests are both desirable and dangerous. There is clearly great potential, but a lot of work must be done before this potential is realized. Only time and extensive research can tell us whether these instruments will become a useful tool in behavioral telehealth contexts."

The main challenge for Internet researchers now is how to define basic methodological constructs in order to gain better control over experimental internal and statistical validity.⁹⁻¹⁰

To explore the possible use of Internet tools in psychological research, this study compared Web-based assessment techniques with traditional paper-based methods of different measures of Internet attitude in a sample of Italian students: on one side we presented a paper-pencil questionnaire to a sample of college students, on the other side we developed an HTML page and online database, posting the same questionnaire on the Web.

In particular, we tried to identify:

- any difference in the responses collected;
- any difference in the psychometric properties of the assessment tools.

MATERIALS AND METHODS

Participants in the online version

Two hundred three online surveys were completed. One hundred four participants were males (51.2%), with 99 females (48.8%). The mean age of the sample was 23.8 (SD = 4.095), with a range from 16 to 41 years. All the subjects were recruited by means of newsgroups messages, search engines postings, and e-mail messages.

Participants in the offline version

The participants were 202 undergraduate students, recruited at a large Italian university by means of presentations during different

lessons and flyers. The average age of the sample was 22.96 ($SD = 1.57$), with a range from 21 to 28 years. The sample was composed of 128 males (63.4%) and 74 (36.6%) females.

Questionnaires

The questionnaire used in the survey is composed of three parts.

- Part one focused on conventional demographic variables (age, sex, school level) and variables indicative of both computer and Internet use.
- Part two consisted of the *Computer Use Survey* by Pratarelli et al.¹¹ This tool is composed by 74 true/false items dealing with a broad range of computer and Internet social and personal activities, also testing possible problematic behaviors.¹²
- The last part is composed by *The Internet Attitudes Survey* by Weiser,¹³ 20 Likert scaled items about personal/professional Internet use.

Procedures

The paper-based survey was administered at one time to entire classes throughout the day. Participants picked up a packet of paper-based self-report measures from the experimenter. All the participants were advised to complete the measures in its entirety, being as honest as possible and not to discuss their responses with others near them. On completion, at each time period, they deposited their

completed packet in a secure drop box controlled by the experimenter.

Participants completed the Web-based survey by accessing a designated Web site. This site was developed in HTML format using the Ms FrontPage 2000 editor, and hosted on a Microsoft NT Server with FrontPage extensions. On completion of the survey, participants were prompted to submit their data. On submission, the data were automatically entered into a tab-delimited format file and were no longer available to participants.

In both versions, participants were reminded that their responses were considered confidential and their participation was completely anonymous.

The statistical analyses were carried out using SPSS 10.1 for Windows.

RESULTS

Internet use

We first analyzed the two samples with respect to the Internet use variables. In general both samples achieved similar preferences: the most used Internet tools are e-mail and Web surfing, followed at a distance by chats and Usenet, MUD (Multi User Dungeons), and FTP (File Transfer Protocol). These choices underlined a particular preference towards tools oriented to interpersonal communication and information research (Tables 1–3). The data also showed that our samples do not make a large use of services like MUDs or FTP (see Tables 2 and 3), that have a relevant number of

TABLE 1. TIME SPENT EACH WEEK ON E-MAIL, NEWSGROUPS, WEB SURFING

Time	Subjects (%)					
	E-mail		Newsgroups		Web surfing	
	On-line	Off-line	On-line	Off-line	On-line	Off-line
Never use	1	13.9	82.8	86.6	2.5	16.3
1–2 hrs	42.9	33.7	13.8	9.4	37.4	26.7
3–5 hrs	25.1	31.2	3.4	2	28.6	35.6
6–9 hrs	21.7	19.8		2	24.6	18.3
More than 10 hrs	9.4	1.5			6.9	3

TABLE 2. TIME SPENT EACH WEEK ON MUDs

Time	Subjects (%)	
	Online	Offline
Never use	90.1	72.3
1–2 hrs	2.5	9.9
3–5 hrs	4.9	8.4
6–10 hrs	2.5	5.9
11–15 hrs		1.5

users in other countries, United States and northern Europe.¹⁴

We compared the two samples using chi-square analyses. We found significant differences in e-mail use (Chi-square = 18.1, d.f. = 4, $p < 0.001$) and in Web surfing (Chi-square = 13.7, d.f. = 4, $p < 0.008$). In particular, within the sample recruited online we have a greater use of the two tools and almost no subject ever used them. Instead, between 14 and 16% of the offline sample has never before experienced any Internet tool.

Analysis of the psychometric characteristics

Different exploratory factor analyses were performed to assess whether Web-based and paper-based questionnaires shared the same factor structure. We applied to both samples separately principal components of factor analyses with Varimax rotation (factor loadings cut-off: 0.25) on the scores obtained in the *Computer Use Survey* and in the *Internet Attitudes Survey*. The four analyses allowed us to identify six similar factors in the two samples.

TABLE 3. TIME SPENT EACH WEEK ON CHAT ROOMS AND FTP

Time	Subjects (%)			
	Chat Rooms		FTP	
	On-line	Off-line	On-line	Off-line
Never use	42.4	36.1	55.5	57.9
1–2 hrs	28.6	33.2	36.5	36.6
2–5 hrs	22.7	23.8	6.4	4.5
5–15 hrs	5.9	6.4	0.5	0.5
More than 15 hrs	0.5	0.5	1.5	0.5

Four categories in the *Computer Use Survey* (P1, P2, P3, P4, explaining 31.69% of the total variance in the online sample and 31.92% in the offline sample):

- SUBSCALE P1—*Internet abuse* (Items included 3, 5, 6, 7, 9, 12, 20, 24, 25, 38, 41, 42, 43, 44, 55, 65, 66, 67, 68, 70; in both samples all these items have a factor loading > 0.25): This subscale held items which refer to physical diseases due to a hard-core Internet use, underlying relevant changes in eating habits, reduced sleep hours to stay logged on at night, being late at appointments, and neglecting one's social and family life.
- SUBSCALE P2—*Interpersonal conflicts* (Items included 27, 35, 40, 45, 46, 47, 48, 63; in both samples all these items have a factor loading > 0.25): This subscale underlined the presence of serious interpersonal conflicts. An inclination for establishing virtual relationships often emerges together with the inevitable choice of organizing the whole social life in function of the Internet.
- SUBSCALE P3—*Introversion/ Extroversion* (Items included 8, 12, 13, 15, 21, 27, 43, 54, 58, 72; in both samples all these items have a factor loading > 0.25): This subscale collected items related to introversion/extroversion dimension. In particular it focused on subject's behavior in establishing online relationships.
- SUBSCALE P4—*Unproblematic use of Internet* (Items included 1, 2, 6, 10, 16, 18, 28, 49, 53, 61, 62, 64, 70, 71; in both samples all these items have a factor loading > 0.25): This subscale isolated an apparent absence of problems with Internet use, including several items referring to a professional and instrumental net use.

Two categories in the *Internet Attitudes Survey* (W1, W2, explaining 38.53% of the total variance in the online sample and 41.93% in the offline sample):

- SUBSCALE W1—*Virtual relationship*: The intention of establishing virtual relationships is to have fun, identifying the

Internet as an alternative communicative medium.

- SUBSCALE W2—*Professional Internet use*: It conveys the idea of an instrumental Internet use, focusing on shopping online and searching for information.

Although this six-factor structure was clearly evident in both samples, four online subscales out of six loaded on items other than those included in the corresponding offline structure. This suggests that the structures, even if very similar, are not identical.

We then analyzed, using Cronbach's Alpha test, the level of internal reliability of both the two questionnaires and the six subscales. The two tests obtained a satisfactory Alpha level in both samples, with slightly lower values for the online sample:

- *Computer Use Survey*, online sample $\alpha = 0.75$; offline sample $\alpha = 0.83$;
- *Internet Attitudes Survey*, online sample $\alpha = 0.74$; offline sample $\alpha = 0.84$.

We found the same trend in the subscales analyses with levels of Alpha always higher than 0.5 in both samples:

- *P1*, on-line sample $\alpha = 0.88$; off-line sample $\alpha = 0.90$;
- *P2*, on-line sample $\alpha = 0.6$; off-line sample $\alpha = 0.75$;
- *P3*, on-line sample $\alpha = 0.56$; off-line sample $\alpha = 0.59$;
- *P4*, on-line sample $\alpha = 0.51$; off-line sample $\alpha = 0.59$;
- *W1*, on-line sample $\alpha = 0.85$; off-line sample $\alpha = 0.87$;
- *W2*, on-line sample $\alpha = 0.88$; off-line sample $\alpha = 0.90$.

Incidence of pathological behaviors

Finally we tried to verify, using the results obtained in the questionnaires, the incidence in two samples of the problematic behaviors related to Internet use. Unfortunately, the developers of the two questionnaires did not report any cut-off value in their papers. Given the lack of consensus over the "Internet addic-

tion" phenomenon, and the differences in the Web use we found between the samples, we analyzed the data using a classical psychometric approach: *z-scores*. Z-scores, which are calculated by subtracting the mean from the subject's score and dividing that answer by the standard deviation, measure the position of any point in a normal distribution in terms of its distance above and below the mean in units of standard deviations. In particular, for each sample we identified cut-off thresholds using one ($z = 1$) and two standard deviations ($z = 2$) over the mean of the totals for the given group:

- *Unproblematic Internet behavior* (Level 1) was defined as a z-score lower than 1;
- *At risk Internet behavior* (Level 2) was defined as a z-score included between 1 and 2 (excluded);
- *Problematic Internet behavior* (Level 3) was defined as a z-score equal or higher than 2.

The distribution of the three levels within the subscales for both samples are reported in Table 4.

First, we verified in each sample any significant gender differences within the levels of the subscales. We found no differences in the on-line sample. In the offline sample, the only significant difference was found in subscale P2 (Chi-square: 18.42, d.f. = 2, $p < 0.0001$): this result indicates a greater presence of interpersonal conflicts in the male group.

Next, we checked the two samples for any differences between them. The results showed significant differences between the two samples in the P2 (Chi-square = 28.22, d.f. = 2, $p < 0.00001$), W1 (Chi-square = 8.27, d.f. = 2, $p < 0.016$), and W2 (Chi-square = 8.9, d.f. = 2, $p < 0.012$) subscales. In the W1 and W2 subscales we found a small percentage (2–4%) of problematic behaviors in the online sample only—more virtual relationships associated with a professional Internet use—even if the higher percentage of at risk behaviors is expressed by the offline group. In relation to subscale P2, we found more at risk and problematic behaviors related to interpersonal conflicts in the offline sample.

Finally, using log-linear analysis we searched for interaction effects between gen-

TABLE 4. INCIDENCE OF RISK LEVEL IN EACH SUBSCALE

Subscale	Risk level	Online			Offline		
		Men	Women	n	Men	Women	n
P1	3	5	6	11 (5.4%)	10	5	15 (7.4%)
	2	5	5	10	16	4	20
	1	89	93	182	102	65	167
P2	3	1	3	4 (2%)	7	0	9 (4.25%)
	2	14	15	29	121	2	123
	1	84	86	170	0	72	72
P3	3	5	1	6 (3%)	4	2	6 (3%)
	2	7	9	16	19	11	30
	1	87	94	183	105	61	166
P4	3	9	3	12 (5.9%)	9	1	10 (5%)
	2	16	7	23	13	3	16
	1	74	94	168	106	70	176
W1	3	8	0	8 (3.9%)	0	0	0
	2	21	9	30 (14.8%)	23	11	34 (16.8%)
	1	70	95	165	105	63	168
W2	3	3	1	4 (2%)	0	0	0
	2	17	4	21 (10.3%)	22	15	37 (18.3%)
	1	79	99	178	106	59	165

der and samples. No significant results were obtained.

DISCUSSION

In this study we explored the possible use of Internet tools in psychological research, by matching Web-based assessment techniques with traditional paper-based methods. In particular, we compared the results of a paper-pencil questionnaire submitted to a sample of college students, with the ones we obtained posting the same questionnaire on the Web.

We made a two level analysis. First we verified the validity of the online assessment instruments by comparing them with their paper-based counterpart. In fact, several researchers^{8–15} identified a number of potential challenges to the reliability of online tests:

- lack of control in the testing situation;
- the possibility of extraneous or temporary factors influencing responses;
- language and cultural differences;
- interactions between the constructs being measured and the characteristics of the testing medium.

Our data showed that completing two attitudes and behaviors questionnaires on the Web did not alter their psychometric characteristics. No significant differences were found in the factorial structure and internal reliability. In particular we found in both samples the following six subscales:

- SUBSCALE P1—*Internet abuse*;
- SUBSCALE P2—*Interpersonal conflicts*;
- SUBSCALE P3—*Introversion/ Extroversion*;
- SUBSCALE P4—*Unproblematic use of Internet*.
- SUBSCALE W1—*Virtual relationship*;
- SUBSCALE W2: *Professional Internet use*;

However, if the same structure was clearly evident in both samples, some online subscales loaded on items other than those included in the corresponding offline ones. This suggests that the structures, even if very similar, are not identical.

These findings show that Web-based data collection neither statistically enhance nor diminish the consistency of responses, nor compromise the integrity of the test, and are a suitable alternative to more traditional methods. This result, similar to the ones obtained in Web-based assessments of personality con-

structs¹⁵⁻¹⁸ and alcohol use,⁵ is even more interesting given the lack of control on the characteristics of the online sample.

Nevertheless the data also indicate that on- and offline versions of the same test can be equivalent but are not always identical. For this reason, it is probably better to assess again the validity of a traditional assessment instrument when it is used online.

Further, we compared the results obtained by the two samples searching for significant differences. In previous studies other researchers¹⁹⁻²⁰ found that technologically advanced assessment methods—audio, computer, and video—produced higher rates of risk disclosure. However in our study the results do not confirm this trend. Even if we found more problematic virtual relationships associated with a professional Internet use in the online sample, the offline sample showed more at risk and problematic behaviors related to interpersonal conflicts. Moreover, given the topics of the two questionnaires—Internet attitudes and behavior—a possible explanation for the results of the online sample may be the interaction between the constructs being tested and the medium used to test them.

In general we did not find large differences in response sets of online participants compared with those of participants who completed a paper survey, even if the online sample was not controlled.

In conclusion, our finding suggests that Internet-based assessment can be a suitable alternative to more traditional paper-based measures. However, there are different practical issues, mainly related to sampling control and validity assessment, that should be considered by psychologists interested in using online tests.

ACKNOWLEDGMENTS

The authors would like to thank Fondazione “Piero, Pietro e Giovanni Ferrero” (Alba, Italy) for the support given to this study. The present work was also supported by the Commission of the European Communities (CEC), in particular by the IST programme (Project

VEPSY UPDATED, IST-2000-25323—<http://www.cybertherapy.info>).

REFERENCES

1. Birnbaum, M.H. (Ed.). (2000). *Psychological experiments on the Internet*. San Diego: Academic Press.
2. Byers, C. (2001). Interactive assessment: An approach to enhance teaching and learning. *Journal of Interactive Learning Research* 12:359-374.
3. Riva, G., & Galimberti, C. (Eds.). (2001). *Towards CyberPsychology: Mind, Cognition and Society in the Internet Age*. Amsterdam: IOS Press. Online: <http://www.emergingcommunication.com/volume2.html>.
4. Musch, J., & Reips, U.D. (2000). A Brief History of Web Experimenting. In: M.H. Birnbaum (Ed.), *Psychological Experiments on the Internet* San Diego: Academic Press, pp. 61-87.
5. Miller, E.T., Neal, D.J., Roberts, L.J., Baer, J.S., Cressler, S.O., Metrik, J., & Marlatt, G.A. (2002). Test—retest reliability of alcohol measures: Is there a difference between Internet-based assessment and traditional methods? *Psychology of Addictive Behaviors* 16:56-63.
6. 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.
7. Azar, B. (2000). A Web of research. *Monitor on Psychology*, 31:42-44.
8. Buchanan, T. (2002). Online assessment: Desirable or dangerous? *Professional Psychology: Research and Practice*, 33:148-154.
9. Riva, G. (2001). The mind over the Web: The quest for the definition of a method for Internet research. *CyberPsychology and Behavior*, 4:7-16.
10. Riva, G. (2001). The mind in the Web: Psychology in the Internet age. *CyberPsychology & Behavior* 4:1-6.
11. Pratarelli, M.E., Browne, B.L., & Johnson, K. (1999). The bits and bytes of computer/Internet addiction: A factor analytic approach. *Behavior Research Methods, Instruments & Computers* 31:305-314.
12. Pratarelli, M.E., & Browne, B.L. (2002). Confirmatory factor analysis of Internet use and addiction. *Cyberpsychology & Behavior* 5:53-64.
13. Weiser, E.B. (2000). Gender differences in Internet use patterns and Internet application preferences: A two-sample comparison. *CyberPsychology & Behavior* 3:167-177.
14. Council, U.S.I. (2000). *State of the Internet 2000*. ITTA Inc. Available: <http://usic.wslogic.com/intro.html> [2002, 8 July].
15. 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.

16. Pasveer, K.A., & Ellard, J.H. (1998). The making of a personality inventory: Help from the WWW. *Behavior Research Methods, Instruments & Computers*, 30:309–313.
17. Smith, M.A., & Leigh, B. (1997). Virtual subjects: Using the Internet as an alternative source of subjects and research environment. *Behavior Research Methods, Instruments & Computers*, 29:496–505.
18. Schwarzer, R., Mueller, J., & Greenglass, E. (1999). Assessment of perceived self-efficacy on the Internet: Data collection in cyberspace. *Anxiety, Stress and Coping* 12:145–161.
19. Gerbert, B., Bronstone, A., Pantilat, S., McPhee, S., Allerton, M., & Moe, J. (1999). When asked, patients tell: Disclosure of sensitive health-risk behaviors. *Medical Care* 37:104–111.
20. Turner, C.F., Ku, L., Rogers, S.M., Lindberg, L.D., Pleck, J.H., & Sonenstein, F.L. (1998). Adolescent sex-

ual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science* 280:867–873.

Address reprint requests to:

Prof. Giuseppe Riva, Ph.D.
 Centre For Studies and Research In
 Communication Psychology
 Università Cattolica del Sacro Cuore
 Largo Gemelli 1
 20123 Milan
 Italy

E-mail: auxo.psylab@auxologico.it

Copyright of CyberPsychology & Behavior is the property of Mary Ann Liebert, Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.