Professional Psychology: Research and Practice
2002, Vol. 33, No. 2, 148–154

0735-7028/02/$5.00 DOI: 10.1037//0735-7028.33.2.148

Online Assessment: Desirable or Dangerous?

Tom Buchanan
University of Westminster

Internet-mediated psychological assessment procedures can play an important role in behavioral telehealth, but their use is not unproblematic. Possible uses of World Wide Web–based tests are discussed. Published empirical evaluations of Web-based personality tests indicate that they can be reliable and valid. However, evidence exists that Web-based versions of tests may not always measure the same constructs as their traditional antecedents: Equivalence cannot be assumed. Web-based clinical assessment seems viable, but there are potential difficulties with measurement of some constructs (particularly negative affect), as well as ethical considerations. While offering great potential, online tests of clinical constructs require stringent validation and cautious use.

The past several years have seen rapid development in the fields of computing and telecommunications. Perhaps the most visible changes have been in one of the areas where the two fields intersect and impact upon society: the "opening up" of the Internet, which has rapidly become an accepted and ubiquitous part of everyday life. This has had profound implications for the working practices of many people: among them, those with a professional stake in the theory and practice of psychology. Many aspects of our work are likely to be (or already have been) affected. These include both the dissemination of information related to mental health and the actual provision of mental health services through electronic media.

Electronic provision of mental health services is an aspect of *behavioral telehealth* (Nickelson, 1996), a term which describes the delivery, at a distance, of a range of services through a variety of possible media (e.g., telephone, fax, e-mail, interactive televideo, Internet chat rooms). The services delivered encompass a broad spectrum of activities related to online therapy (Jerome & Zaylor, 2000; Koocher & Morray, 2000; Laszlo, Esterman, & Zabko, 1999; Maheu & Gordon, 2000; VandenBos & Williams, 2000), ranging from replications of face-to-face therapeutic encounters via video conferencing technologies to "the use of telecommunications and information technology to provide access to behavioral health assessment, diagnosis" (Laszlo et al., 1999, p. 293). Maheu and Gordon (2000) have suggested that, given the wide range of activities covered by the term *behavioral telehealth*, online practitioners should use more specific terminology to describe their activities. Among the labels (and thereby activities) they suggest is "web-based psychological assessment" (Maheu & Gordon, 2000, p. 487). It is this last activity that is the focus of the current article.

Psychometric tests are regarded by many as an essential tool in clinical psychology. Meyer et al. (2001), in a compelling review of evidence for test utility, concluded that "clinicians who rely solely on interviews are prone to incomplete understandings" (p. 128). In addition to citing major assessment instruments (e.g., the Wechsler Adult Intelligence Scale—III, the Minnesota Multiphasic Personality Inventory—2), Anastasi and Urbina (1997) pointed to the use of "many brief questionnaires and rating scales for a quick evaluation of the multiplicity of problems encountered in practice" (p. 509) by clinicians and counselors. There is a wide repertoire of such instruments available for use as part of psychological assessment procedures. Evidence from surveys of usage (Camara, Nathan, & Puente, 2000; Watkins, Campbell, Nieberding, & Hallmark, 1995) indicates that a variety of measures are indeed used by clinicians who are performing assessments.

Given that these tools are useful when assessments are carried out in traditional settings, it is likely that they will also be useful when assessments are carried out through various technological media. There are indications that such assessments are in fact taking place. VandenBos and Williams (2000) surveyed the extent to which psychologists were providing various services through a number of technologies. They found that 15.1% of their respondents reported providing "psychological or neuropsychological assessment" through means described as "Other: Facsimiles, E-mail, etc." (Table 1, p. 491). This indicates that technologically mediated remote assessment is occurring. Further evidence comes from a survey of 56 individuals providing behavioral health services via the Internet (Maheu & Gordon, 2000). Of these, "68% normally provided on-line assessments" (p. 486). As the role played in our lives by the Internet continues to expand, more and more professionals are likely to want to perform Internet-mediated assessments.

Online Tests

Personality assessment instruments take various forms. It is likely that some (mainly projective tests, such as the Rorschach) will currently be difficult to use online (though one could imagine technological advances—e.g., in videoconferencing or virtual reality—making their use more practical). Others, though—mainly those adopting a questionnaire or checklist format—are eminently suited to the new medium. In such cases, the items of a traditional

TOM BUCHANAN received his PhD in psychology in 1996 from the University of Aberdeen. He is a senior lecturer in psychology at the University of Westminster, London, England.

CORRESPONDENCE CONCERNING THIS ARTICLE should be addressed to Tom Buchanan, Department of Psychology, University of Westminster, 309 Regent Street, London W1B 2UW, England. E-mail: tb99@iname.com

paper-and-pencil test are reproduced as a form on an Internet page. Using a Web browser such as Netscape or Internet Explorer, the respondent views and completes the form. The data from the completed form are then automatically transmitted to the test administrator. Instruments of this type are relatively easy to construct and are increasingly being used.

It is difficult to estimate the number of assessment-related Web sites and services currently available on the Internet other than to say that the number is large and increasing. The sites currently available vary considerably in content, quality, and function. Some are clearly commercially oriented. Others might be better classified as public services. Some (the majority) use questionnaire-based instruments. Others are moving toward a more graphics-based approach. Some are entertainment oriented. Others attempt to offer a service that is of some value. Some seek to adhere to high standards of professionalism and to spread good practice in psychological assessment. Others appear unprofessional, unconcerned with ethical and security issues, and may even present information intended to subvert the appropriate use of psychometric tests (e.g., for the purpose of manipulating a test outcome in the context of a custody case).

The motivation for development of many of these sites is easy to understand. There are a number of professional groups who potentially have much to gain from using online tests—for example, occupational psychologists offering psychometric test services to industry, educators, and researchers (Buchanan, 2000, 2001). For researchers interested in personality assessment, the Internet offers easy access to large numbers of participants, as well as a number of other advantages over traditional research methods (Buchanan, 2000; Reips, 2000). Easy access to (or for) large numbers of people may also be one of the advantages that the World Wide Web offers clinicians. Another advantage for clinicians may be that time and money can be saved by using Web-based tests—factors which Camara et al. (2000) identified as constraints upon the current use of tests by psychologists operating in managed-care contexts.

Potential Uses of Online Tests in Clinical Fields

There is growing interest in the provision of online mental health services, and a number of possible models for provision are starting to emerge (Laszlo et al., 1999) that resemble traditional therapeutic models to different extents. Given that traditional forms of psychological assessment and therapy often involve the use of psychometric tests to inform clinical diagnoses or assess the outcomes of therapeutic procedures, if online therapy becomes common, then online tests are likely be used in the same way. Indeed, a number of instruments purporting to measure various constructs (e.g., anxiety, depression, sexual jealousy, and relationship styles, among others) are already available through various "self-help" Internet sites that then provide feedback and advice on courses of action deemed appropriate by the Web site author. Such sites, displaying various degrees of professionalism, can easily be found by using any search engine such as Altavista, Excite, Google or Lycos.

The British National Health Service (NHS) provides an online "Healthcare Guide" (Banks, n.d.), which provides information on common medical symptoms. Part of this system allows users to investigate symptoms they are experiencing by means of a series of structured questions. On the basis of their answers, the system will suggest appropriate courses of action: For minor ailments (e.g., coughs and colds), it may suggest home treatments. For anything that is potentially more serious, it will advise telephoning a helpline staffed by nurses, contacting a health care professional, or summoning an ambulance. The system will also refer the user to information relating to their symptoms and possible solutions.

Interestingly, the symptoms dealt with by this system include stress and depression. It is easy to extrapolate from this and envision a similar system for mental health. Users of such a system might, for instance, complete a clinically focused personality inventory, mental health screening instrument, or symptom checklist (with, e.g., marker items drawn from the revised fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders).* This would then be automatically scored and feedback provided, directing them to follow a particular course of action (e.g., consult a therapist, follow a link to the Web site of a support group or information provider). Additionally, clients might take online tests at regular intervals as a measure of change or progress in therapy, in accord with Koocher and Morray's (2000) suggestion that "reasonable symptom monitoring and outcome assessment should be undertaken routinely" (p. 507). This is a suggestion that is entirely feasible within the boundaries of current technology, and instruments resembling this model already exist.

Strengths of Online Tests

I have argued (Buchanan, 1999) that one of the advantages of putting personality tests online is that it enables more people (e.g., members of the general public with an interest in psychology) to complete them than would otherwise be possible. This mirrors one of the arguments advanced in favor of online therapy or counseling (King & Moreggi, 1998; Laszlo et al, 1999): It permits people who might have previously been unable (due perhaps to geographical or financial circumstances) to access mental health services. Within the context of that form of behavioral telehealth, the ability to remotely assess clients one may never actually meet in a face-to-face context is obviously important.

It is not yet clear whether such a model—where client and therapist interact only electronically—will become the dominant paradigm in behavioral telehealth. An alternative is a model whereby electronic communication plays a supporting role (in the same way as call centers or other resources can be used to augment traditional services). In either case, there are other advantages to using online tests.

Among these is the possibility that people may be more candid when completing tests online. There is considerable evidence that people tend to disclose more information about themselves to computers than to other people (Davis, 1999; Joinson & Buchanan, 2001; Wallace, 1999). This, as Wallace discussed, could make it easier for people to "open up" to online therapists.

One potential reason for the apparent increased self-disclosure is that people are usually anonymous when completing online tests (Joinson, 1999). If a therapist and client are involved in an ongoing therapeutic liaison, this anonymity may be lost. Will the increased self-disclosure still be observed? There is evidence that—at least to some extent—it may. Joinson (1999) showed that

both anonymous and nonanonymous participants completed online tests in a less socially desirable manner (suggesting greater candor) than participants completing paper-and-pencil versions of the same test. It is possible that perceived anonymity is more important than real anonymity. This suggests both that further research in this area may be productive and that people completing clinical measures during online therapy are likely to be more open than they would be using paper-and-pencil tests, whether or not their identity is known to the therapist.

Web-Based Measurement of Personality

As with online psychological work in general (Krantz & Dalal, 2000; Ström, Pettersson, & Andersson, 2000), there is a requirement to establish the validity of online assessment instruments. All the potential advantages will mean nothing if the tests do not actually work. It has been suggested (Buchanan & Smith, 1999b) that there are a number of potential challenges to the validity of online personality tests. Those that are most important for clinical assessment probably revolve around lack of control in the testing situation and the possibility of extraneous (e.g., distraction, environmental cues, technical variability between different hardware and software configurations) or temporary (e.g., fatigue, intoxication) factors influencing responses. Factors such as language and cultural differences may also be important, as may interactions between the constructs being measured and the characteristics of the testing medium. For example, it might well be problematic to assess computer-related anxiety in this way.

A number of studies have been performed that permit comparison of online and traditional versions of the same personality test, either by inclusion of comparison conditions or reference to findings reported in the literature (Buchanan & Smith, 1999a, 1999b; Pasveer & Ellard, 1998; Schwarzer, Mueller, & Greenglass, 1999). These studies have demonstrated that online tests certainly can have acceptable psychometric properties and furthermore have shown a very close match between the different forms of the tests. However, although there is evidence that online and traditional versions of the same test can measure the same constructs, there is also evidence that the instruments are not always identical. This is especially apparent in cases in which tests have subscales intended to measure multiple constructs.

For example, Woolhouse and Myers (1999) compared traditional and Web-based versions of a new personality inventory. They found that the reliabilities of the subscales in both versions of their instrument were acceptable. However, they also detected some differences in the factor structures of the two forms of the instrument: Some items loaded on different subscales in the different versions.

Similarly, another study (Buchanan, Goldberg, & Johnson, 1999) evaluated an online version of a five-factor personality inventory (Goldberg, 1999). This instrument addresses the domain constructs of the Five-Factor Model (Costa & McCrae, 1992), a widely used and influential model of the structure of personality. An online version of the scale was completed by 2,448 participants, who also supplied demographic information and answered self-report questions about the frequency of a number of social behaviors. Again, scale reliabilities were acceptable, and most of the inventory's items loaded on the expected factors. However, in some cases items were found to load on the "wrong" factors or to have strong loadings on more than one factor. This necessitated dropping some items to improve the degree to which the inventory's scales addressed the intended constructs. Further analyses indicated that correlations between scale scores and self-reports of personality-relevant behaviors were as expected and followed patterns documented in the literature, providing some degree of preliminary evidence of the validity of the scales. Despite this, the fact remains that the Web version of the inventory did not appear to have identical psychometric properties to the instrument on which it was based.

More recently, Johnson (2000) has reported a similar exercise with a Web-mediated version of another five-factor personality inventory. This instrument addresses both the domain constructs (as measured by the instrument used previously [Buchanan et al., 1999]) and the "facet" subscales (more tightly focused constructs within each of the personality domains described by the five-factor structure) of the Five-Factor Model. Although the expected five-factor structure was clearly evident in Johnson's data, a (small) minority of these subscales loaded on factors other than those expected. On the basis of these findings, Johnson suggested that his Web-mediated version of the inventory did, broadly speaking, have the expected factor structure but that revisions were required to improve its measurement of the facet-level constructs. The important implication of these studies is that on- and offline versions of the same test can be equivalent but are not always identical. One cannot take the psychometric properties of online instruments for granted, even if they are direct translations of traditional instruments.

These findings suggest that online tests can be satisfactory and valid measures of the intended constructs. However, the differences found also suggest that whenever a traditional assessment instrument is converted for use online, the validity of the new form must be assessed prior to its use. If a test used for online diagnosis or assessment of therapeutic outcomes is not psychometrically satisfactory, then the implications for the client may be very serious. If a screening instrument detects the wrong problem, this may lead to an inappropriate course of treatment being planned. Worse, if used to monitor therapeutic progress, an invalid instrument might give the wrong picture of the client's state of health. The most drastic example that one could imagine here would be an instrument that purported to detect suicidal ideation but did not actually do so. It is clear that for instruments with clinical applications, validation is absolutely crucial.

Web-Based Measurement of Clinical Constructs

To date, only a few published studies report attempts to assess the validity of clinically focused online tests. In an early example, Stones and Perry (1997) demonstrated the potential for online assessment with clinical populations. They constructed an online questionnaire, based on a checklist that respondents used to report symptoms associated with panic attacks. Patterns of responses and reports of comorbidity with other disorders were similar to those reported in the literature based on traditional clinical samples. Although these data are not based on a test per se, they demonstrate that marker items for psychological conditions can be used online.

Similarly, Cutter (1999) has shown that it is possible to collect data on suicidal behavior and its correlates via a Web page. Although Cutter stressed that his current data do not permit prediction of which individuals will go on to kill themselves, such efforts may eventually lead to development of techniques to identify those most at risk. There are indications that the Web may prove more powerful than current tools for identifying those in danger and steering them toward appropriate counseling. The possibility that the Internet will offer increased access to services (Laszlo et al., 1999) means that many more people might be reached (or able to seek help) through the Internet than through traditional means. There is also evidence that suggests online services might be better, as well as more accessible, than traditional services in this area. For example, Levine, Ancill, and Roberts (1998) found that computerized measures were better predictors of suicidality than were interviews with clinicians. The Samaritans, a suicide-prevention organization based in the United Kingdom, has reported that people contacting them by e-mail were twice as likely to report suicidal feelings as were telephone callers (Gibson, 1999).

Perhaps most important for current purposes, Davis (1999) compared Web-based and paper-and-pencil versions of a scale measuring self-focused negative thought (the Ruminative Responses Scale; Nolen-Hoeksema & Morrow, 1991). He found that although Web respondents reported higher levels, the psychometric properties of Web and traditional versions of the measure were equivalent. This finding of equivalence is consistent with those reported previously for nonclinical personality inventories. Davis furthermore found that women scored higher than men on the Web version of the scale. This finding is consistent with evidence that women tend to report more depression than men and to report more rumination (Nolen-Hoeksema, 1987), and it could be interpreted as evidence that the Web scale has some validity as a measure of ruminative responses.

These studies suggest that online clinical assessment may well be viable. However, it is plain that more work needs to be done to establish the psychometric properties and utility of clinically oriented online tests. There are also other issues that require examination.

Potential Problems for Web-Based Assessment

One phenomenon likely to present problems for certain online tests is the possibility of interactions between the construct being tested and the medium used to test it. For example, Tseng, Tiplady, Macleod, and Wright (1998) have shown that levels of computer anxiety can affect participants' responses in different ways depending on the technology used to gather data (PC vs. pen-based personal digital assistant). Joinson (1999) has argued that certain individual differences may affect the way people respond to online tests. This may well be relevant for clinical instruments, especially those that measure negative affect.

It has also been shown that online respondents tend to report higher levels of negative affect than participants filling out paper-and-pencil questionnaires (Davis, 1999). Is this due to enhanced self-disclosure, or perhaps to the well-publicized suggestion (Kraut et al., 1998) that Internet users tend to be more depressed than nonusers? Although the evidence for the latter suggestion is not clear (McKenna & Bargh, 2000), if it is the case, then there may be implications for the provision of online mental health care.

Evidence pertaining to this issue comes from the work of Lambert, Senior, Phillips, and David (2000), who recruited people suffering symptoms of depersonalization disorder either through the Internet (via a bulletin board) or through referrals to their research unit by general practitioners or psychiatrists. All individuals who volunteered to participate in the study were sent a (paper-and-pencil) battery of questionnaires, including the Beck Depression Inventory (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). Those participants recruited through referrals had significantly higher depression scores than those recruited through the Internet, despite the fact that they were similar in terms of both demographic variables and depersonalization symptoms. If Internet users really are more depressed in general, one would expect the Internet group to score higher on the paper-and-pencil depression scale. The reverse was true. Lambert et al. (2000) attributed the difference to the fact that more depressed people may be more likely to be referred to a specialist unit. Although this argument certainly makes sense, the finding does serve to emphasize that the picture of differences between online and offline populations, in terms of both existence and measurement of negative affect, is far from clear.

There are also professional and ethical issues that need to be addressed. There are many nonprofessionals interested in online assessment and mental health services. Some of these well-intentioned but untrained individuals have created "tests" to measure and give feedback on various psychological conditions. Some of the problems arising from these amateur efforts, however well-intentioned, are discussed in an earlier study (Buchanan & Smith, 1999b). The main objection is that in most cases there is no evidence that these instruments have any degree of reliability or validity. If people who complete these tests take the feedback seriously, there is potential for harm to be done.

Even with tests that are psychometrically acceptable, there are dangers. Goritz, Batinic, Goersch, and Moser (1999) have shown that even relatively minor manipulations can produce measurable emotional changes in participants in online experiments. One can easily imagine the potential effect of giving people information that might have serious implications for their well-being. To give potentially distressing feedback, in the absence of any proper follow-up or counseling, would clearly be unacceptable.

It is worth noting while on the subject of ethics that the ethical guidelines of the American Psychological Association (1992) may have implications for the use of online tests. Of particular relevance is Standard 2.02(a), which requires that psychologists who use tests "do so in a manner and for purposes that are appropriate in light of the research on or evidence of the usefulness and proper application of the techniques" (p. 1603). There are obviously still important questions to be answered about the usefulness and proper application of online tests. Do we know enough about Web-based assessments to be sure that they are ready for use?

Implications

Consideration of these issues suggests that online tests measuring clinical constructs or conditions should at present be used only under two sets of circumstances. The first is for research projects

in which no potentially threatening feedback is given to participants. In this situation, one could advise respondents that release of test scores is not possible for ethical reasons and suggest that they consult a qualified professional if worried or seeking diagnoses about their mental or emotional health. This is an extension of the procedures used with "normal" online personality tests and is essential if valid instruments are to be developed for use in other situations.

The second situation is within the context of online therapy or counseling of various types. The advantage of this closed, relatively controlled situation for use of online tests is that the scores of an identifiable individual will be interpreted, and feedback supplied, by a qualified professional who can provide appropriate counseling while disclosing the information. At this time, such use should probably be tentative and exploratory. Before online diagnostic instruments become standard tools or can be used as mass screening instruments freely available on the Web, further work is required to establish their validity and consider other factors.

The requirement of validity is of utmost importance for tests that have such clear potential to affect a person's well-being and to influence the course of therapy. It is obvious that no psychologist should countenance the use in a therapeutic context of an online test whose validity has not been satisfactorily demonstrated independently of any paper-and-pencil versions.

Other factors may also present problems. Some revolve around the fact that one is relying on a technological medium to gather potentially sensitive data from an individual in a remote location. That technological medium, as anyone who has used it will know, is far from perfect.

There is potential for breakdowns in connectivity that may interfere with the assessment process. If assessment is being done asynchronously (e.g., via a Web page which clients may complete at their leisure), this may not be a problem. If synchronous assessment (e.g., where a test is administered during a videoconferencing session) is attempted, the process may be severely disrupted.

Unless one uses a secure server and encrypted communications, there is potential for communications to be intercepted by a third party. Although it is technically possible for confidentiality to be compromised in this way, it remains to be seen whether it will prove to be something that matters in reality. The risk is probably less than that of "old fashioned" surveillance devices being placed in a therapist's office.

The possibility of anonymous communication has been advanced as a factor that may lead people (in situations in which they can be anonymous, such as the NHS Direct Healthcare Guide mentioned previously) to disclose more about themselves. In other situations, it may be a problem: The therapist may have no way of knowing that the persons he or she is dealing with are who they say they are. Ways of establishing the identity of clients being assessed within the context of ongoing treatment are clearly important to determine.

Another problem relates to the use of norms or cutting scores, which play a major role in clinical assessment. To understand the meaning of a test score, one must understand how it relates to the scores one would expect other members of the population to achieve. In the absence of such information, a single score is meaningless. It has been argued above that one cannot simply take an established test, convert it for use on the Internet, and expect it to be psychometrically equivalent. The same is true of norms. Even setting aside situations in which online versions of tests must be revised for psychometric reasons—in which case established norms clearly become useless—there are reasons to suspect that norms may change when tests are placed online. For example, some studies have demonstrated that data derived from Web questionnaires are more variable than data derived from their paper-and-pencil equivalents (Buchanan & Smith, 1999a; Pasveer & Ellard, 1998). This clearly has implications for population distributions of scores. Other factors hypothesized to influence responses to online questionnaires (increased self-disclosure, possibly higher negative affect) might operate to shift mean scores up or down compared with paper-and-pencil assessments of the same construct in the same population. It therefore appears that previously published norms are not appropriate for use with online versions of established tests—their use could lead to misjudgments about the psychological states or personality characteristics of test takers. As part of the validation procedure, new norms should be established for every instrument used online and every population with which it is used in practice.

Use with different populations also raises questions. A possible reason for the finding (Buchanan et al., 1999; Johnson, 2000) that items in online versions of personality tests can load on the "wrong" factors relates to linguistic or cultural differences in the interpretation of items (e.g., the adjectives *mad* and *mean*, both of which are items in at least one published assessment instrument, will be interpreted very differently by English speakers in the United States and the United Kingdom). With the possibility that therapists may be assessing or treating people from outside their immediate locale, these factors must be borne in mind.

In addition to issues that require attention for practical and ethical reasons (e.g., data security), there are probably two key areas that require future work from a psychological perspective. One priority is to improve our understanding of the relationship between negative affect (especially depression) and Internet use and to assess the extent to which it may influence online assessment of various clinical constructs. It may be that some will be more suited than others to assessment by these means. Another priority is to test the assertion (Buchanan, 2000, 2001) that responses to online questionnaires will be characterized by greater self-disclosure (the "candor hypothesis"). There is a body of evidence from studies conducted with stand-alone computers, and of "online behavior" in more general terms, that suggests this is indeed the case. However, the principal assertion does not yet appear to have been put to empirical test.

Evaluations of online tests to date have indicated that they can be valid and useful instruments. However, there is evidence that their psychometric properties cannot be taken for granted. When an established, traditional test is converted for use on the Internet, one cannot simply assume that it will remain an adequate measure of the same constructs or that the published norms will be usable with the new version of the instrument. 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.

References

American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47,* 1597–1611.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.

Banks, I. (n.d.). *The NHS Direct Healthcare Guide.* Retrieved October 1st, 2001, from NHS Web site: http://www.healthcareguide.nhsdirect.nhs.uk

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571.

Buchanan, T. (1999, October). *Online personality assessment: Equivalence of traditional and WWW personality measures.* Paper presented at the meeting of German Online Research '99, Nürnberg, Germany.

Buchanan, T. (2000). Potential of the Internet for personality research. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 121–140). San Diego, CA: Academic Press.

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

Buchanan, T., Goldberg, L. R., & Johnson, J. A. (1999, November). *WWW personality assessment: Evaluation of an online Five Factor Inventory*. Paper presented at the meeting of the Society for Computers in Psychology, Los Angeles, CA.

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

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

Camara, W. J., Nathan, J. S., & Puente, A. E. (2000). Psychological test usage: Implications for professional psychology. *Professional Psychology: Research and Practice, 31,* 141–154.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual.* Odessa, FL: Psychological Assessment Resources.

Cutter, F. (1999, November). *Collecting suicide data from a Web page*. Paper presented at the meeting of the Society for Computers in Psychology, Los Angeles, CA.

Davis, R. N. (1999). Web-based administration of a personality questionnaire: Comparison with traditional methods. *Behavior Research Methods, Instruments, & Computers, 31,* 572–577.

Gibson, E. (1999, November 4). Online help can save lives. Emma Gibson finds a wealth of support groups on the Internet. *The Guardian,* p. 13.

Goldberg, L. R. (1999). *International Personality Item Pool: A scientific collaboratory for the development of advanced measures of personality and other individual differences.* Retrieved April 6, 2001, from http://ipip.ori.org/ipip

Goritz, A., Batinic, B., Goersch, A., & Moser, K. (1999, October). *Induzierbarkeit von Stimmunslangen uber das WWW.* Paper presented at the meeting of German Online Research '99, Nürnberg, Germany.

Jerome, L. W., & Zaylor, C. (2000). Cyberspace: Creating a therapeutic environment for telehealth applications. *Professional Psychology: Research and Practice, 31,* 478–483.

Johnson, J. A. (2000, March). *Web-based personality assessment.* Poster session presented at the 71st Annual Meeting of the Eastern Psychological Association, Baltimore, MD.

Joinson, A. N. (1999). Anonymity, disinhibition and social desirability on the Internet. *Behavior Research Methods, Instruments, and Computers, 31, 433–438.*

Joinson, A. N., & Buchanan, T. (2001). Doing educational psychology research on the Web. In C. Wolfe (Ed.), *Teaching and learning on the World Wide Web* (pp. 221–242). San Diego, CA: Academic Press.

King, S. A., & Moreggi, D. (1998). Internet therapy and self help groups—The pros and cons. In J. Gackenbach (Ed.), *Psychology and the Internet: Intrapersonal, interpersonal and transpersonal implications* (pp. 77–109). San Diego, CA: Academic Press.

Koocher, G. P., & Morray, E. (2000). Regulation of telepsychology: A survey of state attorneys general. *Professional Psychology: Research and Practice, 31,* 503–508.

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

Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukophadhyay, T., & Scherlis, W. (1998). Internet paradox: A social technology that reduces social involvement and psychological well-being? *American Psychologist, 53,* 1017–1031.

Lambert, M. V., Senior, C., Phillips, M. L., & David, A. S. (2000). Depersonalization in cyberspace. *The Journal of Nervous and Mental Disease, 188,* 764–771.

Laszlo, J. V., Esterman, G., & Zabko, S. (1999): Therapy over the Internet? Theory, research and finances. *CyberPsychology & Behavior, 2,* 293–307.

Levine, S., Ancill, R. J., & Roberts, A. P. (1988). Assessment of suicide risk by computer-delivered self-rating questionnaire: Preliminary findings. *Acta Psychiatrica Scandinavica, 80,* 216–220.

Maheu, M. M., & Gordon, B. L. (2000). Counseling and therapy on the Internet. *Professional Psychology: Research and Practice, 31,* 484–489.

McKenna, K. Y. A., & Bargh, J. A. (2000). Plan 9 from cyberspace: The implications of the Internet for personality and social psychology. *Personality and Social Psychology Review, 4,* 57–75.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., et al. (2001). Psychological testing and assessment: A review of evidence and issues. *American Psychologist, 56,* 128–165.

Nickelson, D. W. (1996). Behavioral telehealth: Emerging practice, research and policy opportunities. *Behavioral Sciences and the Law, 14,* 443–457.

Nolen-Hoeksema, S. (1987). Sex differences in unipolar depression: Evidence and theory. *Psychological Bulletin, 101,* 258–282.

Nolen-Hoeksema, S., & Morrow, J. (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: The 1989 Loma Prieta earthquake. *Journal of Personality and Social Psychology, 61,* 115–121.

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.

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

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.

Stones, A., & Perry, D. (1997). Survey questionnaire data on panic attacks gathered using the World Wide Web. *Depression and Anxiety, 6,* 86–87.

Ström, L., Pettersson, R., & Andersson, G. (2000). A controlled trial of self-help treatment of recurrent headache conducted via the Internet. *Journal of Consulting and Clinical Psychology, 68,* 722–727.

Tseng, H. M., Tiplady, B., Macleod, H. A., & Wright, P. W. (1998). Computer anxiety: A comparison of pen-based personal digital assistants, conventional computer and paper assessment of mood and performance. *British Journal of Psychology, 89,* 599–610.

VandenBos, G. R., & Williams, S. (2000). The Internet versus the tele-

phone: What is telehealth, anyway? *Professional Psychology: Research and Practice, 31,* 490–492.

Wallace, P. (1999). *The psychology of the Internet*. Cambridge, England: Cambridge University Press.

Watkins, C. E., Jr., Campbell, V. L., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice, 26,* 54–60.

Woolhouse, L., & Myers, S. (1999, September). *Factors affecting sample make-up: Results from an Internet-based personality questionnaire.* Paper presented at the 1999 British Psychological Society Social Psychology Section Conference, Lancaster, UK.

Received April 16, 2001
Revision received November 11, 2001
Accepted December 17, 2001 ■