Should We Trust Web-Based Studies?

A Comparative Analysis of Six Preconceptions About Internet Questionnaires

Samuel D. Gosling and Simine Vazire — *University of Texas at Austin*
Sanjay Srivastava — *Stanford University*
Oliver P. John — *University of California, Berkeley*

*The rapid growth of the Internet provides a wealth of new research opportunities for psychologists. Internet data collection methods, with a focus on self-report questionnaires from self-selected samples, are evaluated and compared with traditional paper-and-pencil methods. Six preconceptions about Internet samples and data quality are evaluated by comparing a new large Internet sample (*N = *361,703) with a set of 510 published traditional samples. Internet samples are shown to be relatively diverse with respect to gender, socioeconomic status, geographic region, and age. Moreover, Internet findings generalize across presentation formats, are not adversely affected by nonserious or repeat responders, and are consistent with findings from traditional methods. It is concluded that Internet methods can contribute to many areas of psychology.*

Prior to the invention of the World Wide Web, Kiesler and Sproull (1986) discussed the possibility of using computers to collect data in the social sciences. Though optimistic about its potential, they warned, "Until such time as computers and networks spread throughout society, the electronic survey will probably be infeasible" (p. 403). Since then, the Web revolution of the 1990s has stimulated the massive interconnection of technologically advanced societies to a single computer network—the Internet—making it now possible to begin realizing the potential benefits envisioned by Kiesler and Sproull. Indeed, psychologists are now using the Internet for all kinds of research, extending well beyond survey data (Fraley, 2004); these uses range from gathering data on implicit associations (Nosek, Banaji, & Greenwald, 2002), revealed preferences (Rentfrow & Gosling, 2003a), and self-expression on the Internet (Vazire & Gosling, in press) to publicizing new research instruments (Goldberg, 2003).

There are several reasons why psychologists might be interested in collecting data from the Web (Buchanan & Smith, 1999; Schmidt, 1997). One major incentive is that Internet methods can provide access to samples beyond the reach of methods typically used in psychological research. A second potential benefit is the efficiency with which Internet data can be collected; computerized administration allows researchers to obtain sample sizes that far exceed those obtained with most traditional techniques. In addition, Internet methods offer a variety of more mundane, but practically significant, benefits such as dispensing with the need for data entry and being relatively inexpensive. However, these benefits cannot be realized until researchers have first evaluated whether this new technique compromises the quality of the data.

Although many researchers have begun using this new tool, its benefits and potential obstacles have gone largely unexamined. Previous researchers have addressed technical issues—the "how to" of Internet data collection (Birnbaum, 2001; Dillman, 1999; Dillman, Tortora, & Bowker, 1998; Fraley, 2004; Kieley, 1996; Morrow & McKee, 1998)—or have speculated about the pros and cons (Hewson, Laurent, & Vogel, 1996; Kraut et al., 2003; Michalak & Szabo, 1998; Schmidt, 1997). However, few have tested empirically the quality of data collected on the Internet. As a result, researchers trying to publish Internet-based studies have often found editors and reviewers to be, at best, skeptical of the quality of their data (e.g., Mezzacappa, 2000). The time has come for empirical analyses of the quality of Internet data.

Our analyses will focus on one kind of data now commonly collected on the Internet—questionnaire data gathered in self-selected Internet samples. We use the term *questionnaire* to refer to the self-report surveys, tests, and assessments widely used in psychology (e.g., personality tests). Many of the points we make also apply to other common psychological designs such as reaction-time experiments (see, e.g., McGraw, Tew, & Williams, 2000) administered to self-selected samples on the Internet. However, our analyses and conclusions about self-selected Internet samples are not meant to apply to the kinds of

Editor's note. William C. Howell served as action editor for this article.

Author's note. Samuel D. Gosling and Simine Vazire, Department of Psychology, University of Texas at Austin; Sanjay Srivastava, Department of Psychology, Stanford University; Oliver P. John, Department of Psychology, University of California, Berkeley.

Correspondence concerning this article should be addressed to Samuel D. Gosling, Department of Psychology, University of Texas at Austin, 108 East Dean Keaton Street, Austin, TX 78712. E-mail: gosling@psy.utexas.edu

 0003-066X/04/$12.00
Vol. 59, No. 2, 93–104 DOI: 10.1037/0003-066X.59.2.93

Samuel D. Gosling

assessments requiring representative samples, such as general surveys, opinion polls, or other research often used in sociology and other social sciences that rely on probabilistic sampling (but see Best, Krueger, Hubbarb, & Smith, 2001; Couper, 2000; Taylor, 2000).

Our goal is to evaluate six preconceptions that have been raised as likely limitations of Internet questionnaires. Our analyses draw on an Internet sample of questionnaire responses gathered in our own research, comparing them with a sample of traditional studies. We also draw on other findings relevant to each of the six preconceptions.

Six Preconceptions About Internet Data

It is prudent to be cautious about any new method. This is especially true for techniques that appear to offer many benefits while incurring few costs. Internet methods have certainly received their fair share of suspicion, ranging from nebulous expressions of distrust in departmental corridors and journal reviews to published concerns about the quality of Internet data (e.g., Azar, 2000; Birnbaum, 2004; Buchanan, 2000; Johnson, 2000, 2001; Reips, 2000; Skitka & Sargis, in press; Smith & Leigh, 1997). Concerns about Internet data have even found their way into popular outlets; for example, the eminent National Institute on Aging psychologist Paul Costa summarily described Internet data collection techniques as "flawed" (Peterson, 2003) when commenting on the results of a peer-reviewed Internet study. However, even if some caution is warranted, researchers should not go so far as to dismiss a new method without first evaluating their concerns. Our analyses focus on six frequently mentioned preconceptions about the quality of Internet data. These preconceptions, which are summarized in Table 1, should concern all researchers interested in using such data.

Preconception 1 is that Internet samples are not diverse. This is driven by the widely held idea that the Web is dominated by a rather narrow segment of society, such as "techies" or "nerds." According to the stereotype, Web users are young, White, upper middle-class men. Thus, one recurring concern of researchers (and reviewers) has been that Internet samples may overrepresent such demographics (e.g., Azar, 2000; Buchanan, 2000; Krantz & Dalal, 2000).

Preconception 2 is based on the stereotypical view of the Internet as a haven for the socially maladjusted, populated by social rejects and loners with no other outlet for social contact. This leads to the preconception that Internet users tend to be higher in depression and more socially isolated than are individuals who do not use the Internet. This preconception was reinforced by one high-profile study (Kraut et al., 1998) and sustained by the heavy media attention the findings received. As one newspaper article later noted, "Three years ago, Carnegie Mellon University researchers captured national headlines with findings that suggested the Internet, a tool designed to enhance communication, actually took people away from their families and caused them to feel more alone" (Mendenhall, 2001).

Preconception 3 is that the data obtained from Web sites are affected by the presentation format of the site (e.g., Azar, 2000; Bowker & Dillman, 2000; Dillman, Tortora, et al., 1998; Dillman, Tortora, Conradt, & Bowker, 1998). For example, is it the case that data obtained at a Web site that attracts people looking for self-insight would differ from data obtained at a Web site that attracts people looking for an amusing diversion?

Preconception 4 is that Web-questionnaire findings are adversely affected by nonserious responses (e.g., Azar, 2000; Buchanan, 2000). The great accessibility of the Internet makes Web-based questionnaires easy targets for nonserious responses, potentially undermining the validity of data gathered through this medium.

Preconception 5 is that findings are adversely affected by the anonymity afforded by Web questionnaires (e.g., Buchanan, 2000; Skitka & Sargis, in press). The anonymity afforded to individuals taking questionnaires on the Web inevitably leads to less control over who is taking the questionnaire than is afforded in traditional research contexts. For example, this reduction of control exposes Internet researchers to the possibility that individuals might complete the questionnaire multiple times (Buchanan, 2000; Johnson, 2001).

Preconception 6 is that Web-questionnaire findings are inconsistent with findings from traditional methods (e.g., Krantz & Dalal, 2000; Skitka & Sargis, in press). However, if Internet and traditional methods yield similar findings, it cannot be argued that Internet data are less valid than other data.

Comparative Data Sets

These six concerns are understandable, but is there any evidence to support them? We address each preconception

Simine Vazire

with empirical evidence, and where possible we provide a comparison between the samples, data, or findings from Internet methods and those from traditional methods. We also describe several steps researchers can take to improve the quality of Web-questionnaire data.

Our analyses compare a typical self-selected Internet sample with data gathered by more traditional means. Given our expertise in personality–social research, we shall use an Internet sample of questionnaire responses gathered in our own research on personality. What would be the most appropriate standard with which to compare these Internet data? To set the bar high for this new method, the comparison data should be the very best available data gathered by traditional means and should be drawn from the same field as the Internet data. Thus, our Internet sample will be compared with samples published in the top journal in personality–social psychology.

Web Data

The Web data come from visitors to a noncommercial, advertisement-free Web site, outofservice.com (see Srivastava, John, Gosling, & Potter, 2003, for details). The Web site contains personality measures as well as several games, quizzes, and questionnaires for entertainment purposes. Potential participants can find out about the site through several channels: It can be found with major search engines under key words like *personality tests*; it is listed on portal sites, such as Yahoo!, under their directories of personality tests; and individuals who have previously visited outofservice.com and signed up for its mailing list receive

Table 1
Six Preconceptions About Internet Methods

Preconception	Finding
1. Internet samples are not demographically diverse (e.g., Krantz & Dalal, 2000).	*Mixed.* Internet samples are *more* diverse than traditional samples in many domains (e.g., gender), though they are not completely representative of the population.
2. Internet samples are maladjusted, socially isolated, or depressed (e.g., Kraut et al., 1998).	*Myth.* Internet users do not differ from nonusers on markers of adjustment and depression.
3. Internet data do not generalize across presentation formats (e.g., Azar, 2000).	*Myth.* Internet findings replicated across two presentation formats of the Big Five Inventory.
4. Internet participants are unmotivated (e.g., Buchanan, 2000).	*Myth.* Internet methods provide means for motivating participants (e.g., feedback).
5. Internet data are compromised by anonymity of participants (e.g., Skitka & Sargis, in press).	*Fact.* However, Internet researchers can take steps to eliminate repeat responders.
6. Internet-based findings differ from those obtained with other methods (e.g., Krantz & Dalal, 2000).	*Myth?* Evidence so far suggests that Internet-based findings are consistent with findings based on traditional methods (e.g., on self-esteem, personality), but more data are needed.

Sanjay Srivastava

notification when a new questionnaire is added. As is common on the Internet, news of the site has also spread widely through informal channels such as emails or unsolicited links on other Web sites.

Two distinct Web pages were used, attracting a broad and diverse sample. One is entitled "All About You—A Guide to Your Personality" and attracts participants by appealing to their desire for self-insight. The second Web page is entitled "Find Your *Star Wars* Twin" and attracts participants by appealing to their desire for fun and entertainment. We consider below how these two Web pages differ, both in who they attract and in the substantive findings.

Most analyses are based on the first 361,703 participants to complete the questionnaire, but in some instances we also draw on Srivastava et al.'s (2003) study of personality development, which was based on a subsample of 132,515 participants (selected, for research purposes, to be comprised of participants between the ages of 21 and 60 years living in the United States or Canada).

Traditional Data

The traditional data were comprised of all studies published in a full year (2002; i.e., Volumes 82 and 83) of the premier empirical journal in personality–social psychology, the *Journal of Personality and Social Psychology* (*JPSP*). In terms of citation impact, *JPSP* is the highest ranked empirical journal in personality–social psychology (*Social Sciences Citation Index: Journal Citation Reports*, 2002), it has a 76% rejection rate, and in 2001, the latest year for which reports were available, received more manuscripts than any other American Psychological Association journal (American Psychological Association, 2003). Thus, *JPSP* provides a sample of the best published research using traditional methods. These are the kinds of studies that could potentially make use of Internet methods.

Internet methods are used for empirical studies of humans, so we culled from *JPSP* only empirical studies of humans. As a result, we excluded meta-analyses, theoretical articles, replies, computer simulations, analyses of literary passages, Internet-based studies, and nonhuman animal studies, yielding 510 samples from 156 articles. Estimates of the characteristics of traditional samples were obtained from trained coders who reviewed each article published in *JPSP* during 2002. Each sample in each *JPSP* article was coded for the following characteristics: sample size, gender composition, age composition, race composition, socioeconomic status composition, geographic location, and whether the sample was comprised of students.

JPSP publishes articles using either correlational or experimental designs. Given that the Internet data with which we intended to compare the *JPSP* data are based on a correlational design, we were concerned that a marked experimental bias in *JPSP* studies might distort our findings. Thus, we also coded all articles for whether they used experimental or correlational designs. Although there is a bias in favor of experimental designs in *JPSP*, the bias is rather modest (59% of studies are experimental, but only 27% of participants are from experimental studies). To ensure our comparisons are fair, we compared the Internet sample with the full set of 510 *JPSP* samples published in 2002 and also with the subset of 211 samples used in correlational designs.

Evaluating the Preconceptions

Preconception 1: Internet Samples Are Not Sufficiently Diverse

We address Preconception 1 by comparing the diversity of Web samples and traditional samples with respect to five important domains: gender, race, socioeconomic status, geographic region, and age. Table 2 presents a summary of the comparison between our Internet sample in the first data column and all of the *JPSP* traditional samples in the second data column. The third data column presents the statistics for only the 211 *JPSP* samples that used correlational designs. To provide a standard against which the Internet data can be compared, we begin each section below by examining the traditional samples published in this top journal.

Gender. The status quo with respect to gender composition of traditional samples is far from impressive; as Table 2 shows, an average of 71% of the participants in all traditional samples and 77% of participants in correlational studies are female. How do Internet samples compare?

Among the outofservice.com participants, only 57% of those who reported their gender were female, a gender discrepancy much smaller than that found in traditional samples. Although the stereotyped view is that far more men than women use the Internet, our findings are more consistent with recent data, which suggest that men and women are now using the Internet in equal numbers (Len-

Oliver P. John

hart et al., 2003). Indeed, depending on the subject matter, it is even possible to obtain samples that are comprised mostly of women; for example, in a sample of pet owners' personality descriptions of themselves and their pets, 83% of the 1,640 participants were women (Gosling & Bonnenburg, 1998).

Race. Race composition was reported for only 24% of all *JPSP* samples and 33% of samples from correlational studies, so we can draw only tentative conclusions about the racial diversity of traditional samples. Furthermore, given the premium placed on diversity, it seems likely that race was reported more often for diverse samples than for all-White samples. Indeed, a closer inspection of the samples revealed a subset of studies with almost no White participants (i.e., samples recruited on the basis of race). To reduce the bias introduced by these race-selected samples, we based our estimate of race composition on samples with at least 40% White participants. Using this method, we estimate that 80% of participants in traditional samples are White. How do Internet samples compare?

The outofservice.com participants reported their race as one of six categories: 26,048 (7.2%) were Asian; 9,133 (2.5%) were Black; 8,281 (2.3%) were Latino; 6,394 (1.8%) were Middle Eastern; 276,672 (76.5%) were White; and 13,668 (3.8%) indicated Other; 5.9% of participants declined to report their race. In contrast, the U.S. population is more diverse (e.g., 12.3% Black, 12.5% Latino; U.S. Census Bureau, 2000). These racial disparities in Internet use are consistent with a recent survey, which indicated

Table 2
Comparison of Traditional and Internet Sample Characteristics

Characteristic	Internet sample	*JPSP* samples in 2002: All traditional samples	*JPSP* samples in 2002: Correlational traditional samples
No. of participants	361,703	102,959	75,363
% of student samples	—[a]	85%	70%
% of samples reporting gender	—[a]	72%	80%
Avg. % female	57%	71%	77%
Avg. % male	43%	29%	23%
% of samples reporting race	—[a]	24%	33%
Avg. % White	77%	80%[b]	80%[b]
No. of non-Whites	83,192	14,949	14,006
% of samples reporting SES	—[a]	5%	10%
No. of non-U.S. participants	110,319	17,988	12,563
% of samples reporting age	—[a]	32%	54%
In student samples	—[a]	27%	49%
In nonstudent samples	—[a]	67%	71%
Mean age (in years)	24.3[c]	22.9[d]	25.1[d]

Note. Data for the Internet sample come from outofservice.com Web-questionnaire participants. Data for traditional samples come from analyses of samples in all articles published in one year (2002) of the *Journal of Personality and Social Psychology* (*JPSP*). The correlational samples are the subset of *JPSP* samples using a correlational design. All means for traditional samples are weighted by sample size. Avg. = average; SES = socioeconomic status.
[a] Not applicable to the Internet sample because it is a single sample. [b] To reduce the bias introduced by race-selected samples, these averages are based on samples with at least 40% White participants. [c] This average includes only participants between the ages of 11 and 100 years. [d] Because age was reported more often for nonstudent samples than for student samples and nonstudent samples tend to be older than student samples, this mean was calculated in two steps. First, mean age was calculated separately for student samples and nonstudent samples. Second, the overall mean was a weighted composite of the mean age of student samples (weighted by the total proportion of student samples) and the mean age of nonstudent samples (weighted by the total proportion of nonstudent samples).

that in the United States in 2002, 60% of Whites had access to the Internet, compared with 54% of Hispanics and 45% of Blacks. These differences were largely attributable to income disparities (Lenhart et al., 2003).

Although these differences are shrinking rapidly, the current racial disparities could pose a threat to the generalizability of results from Web questionnaires. However, it is important to note that although racial disparities exist in Internet samples, the large sample sizes obtained with Web questionnaires means that even fractions of percentages can reflect very large absolute numbers. For example, despite the small proportion of Latino participants in the outofservice.com sample (2.3%), there were still more than four times as many Latino participants (N = 8,281) in this Internet sample than in all the traditional samples in a whole year of *JPSP* combined (N = 2,061).

Socioeconomic status. A potential drawback to Internet samples is that people with lower incomes, people from rural areas, and people with less education are less likely to have access to the Internet than are high-income, urban, educated individuals. As noted in our discussion of race, there is evidence of income disparities in Internet use.

On the basis of the findings in Table 2, it is safe to say that traditional samples overrepresent highly educated individuals from high-income families. Fully 85% of traditional samples draw on university students for participants, including 70% of samples used in correlational studies, thus ensuring that participants are far more educated than the population at large, of which only 27% are college graduates (National Center for Education Statistics, 2000). This is consistent with historical trends, which show that social psychologists have drawn 70% or more of their research participants from student samples since at least the 1960s (Sears, 1986). Socioeconomic status information was reported for only 5% of all traditional samples and 10% of correlational studies, making it difficult to establish precisely the socioeconomic diversity of traditional samples. How do Internet samples compare?

In the outofservice.com questionnaire, a question about social class was added later during the data collection period, so this information was available for only a subset of the sample. Of these 116,800 participants, 1,323 (1.1%) reported being poor; 17,981 (15.4%) reported being working class; 6,405 (5.5%) reported being lower middle class; 53,669 (46.0%) reported being middle class; 34,105 (29.2%) reported being upper middle class; and 3,314 (2.8%) reported being upper class. This reasonably balanced social class distribution (22% below middle class, 32% above) suggests that although people in higher social classes were somewhat overrepresented, the sample included participants from a broad range of backgrounds. Thus, the barrier of socioeconomic background would seem to be only a limited one in terms of inclusion in Internet questionnaire studies. Indeed, Lebo (2000) noted, "The Internet is far from being a bastion of highly educated, well-paid users. While the vast majority of high education/high income people use the Internet, those with less education and lower incomes log on in impressive numbers" (p. 11). In one survey, 23% of adults with no high school degree reported using the Internet, and 38% of adults making less than $30,000 per year reported having Internet access (Lenhart et al., 2003).

Geographic region. Although traditional samples as a whole span a wide range of geographic regions (17% non-U.S. participants), any given traditional sample usually includes participants from only one locale, potentially compromising the generalizability of the findings. Furthermore, there is no evidence that traditional samples are themselves sampled in a representative way with respect to geographic region. How do Internet samples compare?

In order to address the possibility of a geographical sampling bias in Web samples, we analyzed the country and state of residence of the outofservice.com participants. Within the United States, participants represented all 50 states. Using this Internet sample, Rentfrow and Gosling (2003b) showed that the response rate from each state correlated almost perfectly (r = .98) with the state's total population, as indexed by the U.S. Census Bureau (2000); this suggests that participants were quite representative of the American population with respect to geographic region. Furthermore, the large sample sizes afforded by Web questionnaires make it possible to examine cultural and national differences despite the low percentages of participants from any particular country. For example, Table 2 shows that the outofservice.com sample had more than six times more non-U.S. participants (N = 110,319) than all of the traditional samples combined (N = 17,988). Thus, although the outofservice.com sample is comprised overwhelmingly of North Americans (69.5% of the participants are from the United States and 7.2% are from Canada), the sample does represent a breadth of geographic regions from around the world. In fact, 60 countries, from Albania (N = 368) to Zimbabwe (N = 110), are represented in this sample by at least 100 participants each. This geographic diversity is likely to become an even stronger advantage of Internet methods as accessibility to the Internet spreads throughout the world.

Age. It is widely believed that the Internet is used almost exclusively by young people and excludes older people. But do Internet studies fare worse than traditional methods, which rely heavily on college students?

In the traditional studies, age composition was reported for only 32% of the samples (and 54% of those using correlational designs), making it difficult to establish the mean age of traditional samples. Furthermore, age characteristics were reported much more frequently for samples comprised of nonstudents than for samples comprised of students. After taking into account these reporting biases (see Table 2 *Note*), we estimated the average age of participants in *JPSP* samples to be a mere 23 years. How do Internet samples compare?

In the outofservice.com dataset, of the 342,688 participants who reported their age (excluding those reporting ages below 11 or over 100 years), 44.6% (152,727) were 11 to 20 years old; 35.9% (123,046) were 21 to 30 years old; 11.7% (40,236) were 31 to 40 years old; 5.4% (18,498)

were 41 to 50 years old; 1.9% (6,432) were 51 to 60 years old; and 0.5% (1,749) were 61 to 100 years old. This distribution clearly shows that the sample was skewed toward a younger population. However, most traditional research recruiting participants from university subject pools excludes participants under 18 years of age; when we exclude people under 18 years of age from the Internet sample, the mean age increases from 24.3 to 27.6 years. Thus, the age bias in Internet samples is not as dramatic as one might imagine. Indeed, a study of 48,000 households sampled from the census pool showed that Internet access rates were similar for three age groups: 53% for 9–17-year-olds, 57% for 18–24-year-olds, and 55% for 25–49-year-olds; however, among the 50-and-over group, only 30% accessed the Internet (U.S. Department of Commerce, 2000). A smaller study of adults using more fine-grained age categories found similar results, with some indications that the lower rates of access in the census study's 50-and-over group was mainly attributable to adults 65 years and over: adults aged 50–64 years were almost three times as likely to have access as were adults 65 years and over (Lenhart et al., 2003). People of all ages also spend substantial amounts of time online. Among Internet users, 25–35-year-olds and 46–55-year-olds spend about the same amount of time online (11.3 hours per week and 10.3 hours per week, respectively; Lebo, 2000).

Another concern is whether Internet recruitment could produce age-related sampling biases. In their research on personality development, Srivastava et al. (2003) considered what age-related sampling biases would look like if they had occurred. A plausible scenario was that the Internet may be less familiar to older people, and thus older people may have to be higher in openness in order to seek out and participate in an Internet study. In fact, Srivastava et al. found a small drop in openness with age, not the increase one would predict if older adults needed to be especially open in order to participate. If there was any openness-related sampling bias, it was sufficiently weak that it did not overwhelm the age effects.

Srivastava et al. (2003) also tested another potential age-related sampling bias. If older people in the sample were unusually high in conscientiousness (because of selective mortality) or in openness (for the reasons described above), then the standard deviations for conscientiousness and openness should decrease with age because people at the low end of the continuum become less common as age increases. However, standard deviations did not decrease with age, providing more evidence that age-related sampling bias did not adversely affect the data.

Summary. The critical question is whether data obtained with a new technique—in this case, responses to Web questionnaires in self-selected samples—are at least as good as the viable alternatives currently used. We have shown here that although Internet samples are not representative of the population at large, they are generally more diverse than samples published in a highly selective psychology journal. Specifically, our comparisons suggest that Internet samples are more representative than traditional samples with respect to gender, socioeconomic status, geographic location, and age and are about as representative as traditional samples with respect to race. Indeed, our analyses of traditional samples are consistent with Krantz and Dalal's (2000) observation that "the overwhelming majority of traditional psychology studies make no effort whatsoever to ensure that the samples used are randomly selected (and, therefore, representative of the larger population being studied)" (p. 48).

Moreover, the wide accessibility of Web questionnaires makes them available to a large and broad audience. Physically handicapped, shy, and disorganized individuals with Internet access have as great a chance of being included as able-bodied, extraverted, and conscientious ones who might be overrepresented in community volunteer samples recruited by newspaper ads or the undergraduate student samples typical of much current psychological research. However, the other side of the same coin suggests that Internet methods may not be suitable for recruiting participants from certain special populations, such as the elderly, homeless, illiterate, or those living without electricity. Clearly, then, we do not suggest that Web samples are a good substitute for true random samples, but rather for the more common types of samples, such as undergraduates and volunteers, who typify much research in psychology. Thus, we are more concerned with the generalizability than the representativeness of the samples collected by Web questionnaires and with choosing the method best suited for the target population. Ideally, psychological data should be obtained in samples representative of the population to which the findings are to be generalized (e.g., Lerner, Gonzales, Small, & Fischhoff, 2003).

Finally, it is worth noting that few cost-effective methods permit researchers to assess truly representative samples. Indeed, unlike opinion-type surveys where representativeness is crucial, representativeness may not always be the highest priority in psychological research (Mook, 1983). As one researcher acknowledged, "No one has ever gotten a random sample in the lab" (Krantz; as quoted in Azar, 2000). No standard method can capture a truly representative sample; even methods that are meant to do so, such as random-digit dialing, include sampling biases because only a small percentage of contacted individuals actually participate (Chang & Krosnick, 2003).

Preconception 2: Internet Samples Are Unusually Maladjusted

Is the Internet a haven for the socially maladjusted and thus an unrepresentative source of participants? Internet users have stereotypically been depicted as socially isolated computer nerds or social-misfit hackers. These stereotypes were bolstered by one early study of 73 households (Kraut et al., 1998), which reported a significant increase in depression and a decrease in social contact after the households got Internet access. These findings elicited substantial interest from the media, and with over 230 citations (Institute for Scientific Information), also garnered a great deal of attention from other researchers.

However, conclusions about Internet users based on this early portrayal would be premature. A much less

publicized follow-up of this same sample and an additional sample, however, indicated that the depression effect was not reliable and that Internet users are in fact not unusually maladjusted (Kraut et al., 2002). Other recent data also suggest that Internet users are largely similar to nonusers in their adjustment. A recent phone survey found that 72% of Internet users had visited with a friend or relative on a given day, compared with 60% of nonusers (Lenhart, 2000). Another study (Lebo, 2000) found that there is little difference in hours per week spent socializing with friends (9.7 hours for users, 9.9 hours for nonusers) or participating in clubs and volunteer organizations (2.4 hours for users, 2.0 hours for nonusers). In addition, when we compared the 21-year-olds in the Internet sample with 21-year-old participants from three university subject pools, we found no significant differences in neuroticism or introversion, the traits associated with depression and social isolation, respectively. Thus, there is little support for the belief that Internet users are unusually maladjusted.

Preconception 3: Internet Findings Do Not Generalize Across Presentation Formats

Very little is known about the impact of different administration formats in traditional research. For example, when administering a questionnaire, researchers freely chose between Scantron forms and paper-and-pencil tests, or between administering the questionnaire to individuals versus groups, with little concern for the differences among administration formats. Despite the lack of regard for presentation effects in traditional methods, they have been of concern to Internet researchers (Bowker & Dillman, 2000; Dillman, Tortora, et al., 1998; Dillman, Tortora, Conradt, et al., 1998).

Anyone who has browsed the Web will know that Web questionnaires come in a wide variety of styles, with some appealing to more serious motives, such as gaining insight into one's own behavior, whereas others seem to offer little more than an amusing diversion. Furthermore, software and hardware differences among participants mean that not every participant will see the exact same presentation of a questionnaire. Do differences in presentation formats affect the findings?

Although we cannot provide a definitive test of this question, the outofservice.com data do permit a preliminary examination. The outofservice.com data were obtained using two versions of the same Big Five Inventory (John & Srivastava, 1999), the All About You version and the *Star Wars* version. The two sites drew somewhat different profiles of participants. All About You drew 66% women, whereas *Star Wars* drew 39% women. All About You participants were on average about two years older than *Star Wars* participants. Controlling for gender, the greatest between-sites difference on the Big Five personality dimensions was for openness; *Star Wars* participants tended to be slightly more open to experience (partial $r = .10$). These analyses suggest that the Web sites are drawing on somewhat different kinds of participants. In spite of these differences, though, Srivastava et al.'s (2003) regression analyses clearly replicated across the two Web formats. This suggests that despite the sampling differences across the two formats, presentation format did not significantly affect the nature or quality of the results. Although social scientists concerned with probabilistic sampling have found that formatting effects can influence response rate (Bowker & Dillman, 2000; Dillman, Tortora, Conradt, et al., 1998), there is as of yet no evidence that they affect the content of people's responses. Of course, even small differences in presentation format can have serious negative consequences for certain experiments (e.g., sensation and perception experiments), but for most questionnaire research, presentation effects do not seem to jeopardize the quality of the data.

Preconception 4: Internet Participants Are Not Sufficiently Motivated (to Take the Study Seriously and Respond Meaningfully)

The validity of any research methodology relying on volunteers is contingent on the ability and willingness of volunteers to provide meaningful responses. This is particularly true for Web questionnaires because their great accessibility makes them easy targets for nonserious responses. However, the issue of participant motivation and responsiveness is an important one for both traditional and Internet methods. Paper-and-pencil measures are presumably just as susceptible to faking or dishonest responses as Web-based measures. Furthermore, it has long been recognized that students participating in psychological research are more likely than nonstudents to be suspicious of the research (Argyris, 1968; Jourard, 1968; Orne, 1962) and to have hostile feelings toward the experimenter (Jackson & Pollard, 1966; Jourard, 1968), suggesting that traditional student samples have their own difficulties with respect to participant motivation.

It is possible to test whether Web-questionnaire data are adversely affected by nonresponsiveness (unmotivated or noninterpretable responses). One method advocated by Johnson (2001) is to screen each submission for markers of nonresponsiveness such as long strings of identical responses. Another method is to examine scale reliabilities and discriminant validities (John & Benet-Martinez, 2000). Problems with administering questionnaires (e.g., random or otherwise unreliable responses) would be reflected in diminished coefficient alpha reliabilities. Attempts to self-enhance for the sake of receiving positive feedback should result in increased scale intercorrelations because people would not discriminate among traits and would simply rate themselves positively on all traits. To test these possibilities, we compared the alpha reliabilities and scale intercorrelations of Big Five Inventory data collected using paper-and-pencil and Internet methods.

The outofservice.com reliabilities were very similar to those obtained using traditional paper-and-pencil measures (John & Srivastava, 1999); reliabilities across methods were within two hundredths of a point of each other for each trait. Also, the discriminant intercorrelations among the Big Five Inventory scales in the Internet data (absolute mean $r = .16$) were at least as good as those obtained using traditional methods ($r = .20$; John & Srivastava, 1999).

These analyses suggest that the Web-questionnaire data were not especially affected by random or otherwise unreliable responses, nor by attempts to self-enhance for the sake of receiving positive feedback. This is consistent with other studies showing that the reliability and factor structure of other psychological constructs are similar for paper-and-pencil and Web versions (Buchanan & Smith, 1999; Johnson, 2000).

Web questionnaires also have the benefit of drawing from self-selected samples. Previous research has shown that participants from self-selected samples provide clearer, more complete responses than participants who are not self-selected volunteers such as undergraduate psychology students (Pettit, 2002; Walsh, Kiesler, Sproull, & Hesse, 1992). There is also evidence that participants engage in less socially desirable responding and survey satisficing when responding to a Web questionnaire than to a paper-and-pencil questionnaire (Kiesler & Sproull, 1986; Richman, Kiesler, Weisband, & Drasgow, 1999) or a telephone interview (Chang & Krosnick, 2003). Furthermore, Web questionnaires provide a unique advantage for motivating participants to respond seriously: appealing to people's desire for self-insight by providing interesting, immediate feedback. Participants are motivated to answer seriously to receive accurate feedback about their personality. This unique advantage is made possible by the automated data entry and scoring permitted by Web questionnaires. These same features also save time and money in recruitment and data entry.

Preconception 5: The Anonymity Provided by Web Questionnaires Compromises the Integrity of the Data

Although many traditional methods take steps to ensure participants' confidentiality, few can claim to provide complete anonymity. When completing a questionnaire using traditional methods, participants typically show up to a laboratory, hand in their completed questionnaire to an experimenter, and have their data entered by hand—all steps that reduce the anonymity of their responses. Web questionnaires, in contrast, allow participants to complete the questionnaire alone, without ever coming into contact with an experimenter, and eliminate the need for data entry. These characteristics allow researchers to address questions that would be difficult or impossible to address with traditional methods. For example, participants may feel more comfortable disclosing personal information in a Web questionnaire than in a less anonymous setting such as a research lab (Levine, Ancill, & Roberts, 1989; Locke & Gilbert, 1995). Indeed, reporting of stigmatized health, drug-related, and sexual behaviors has been shown to increase as anonymity increases (Turner et al., 1998).

However, providing anonymity is often seen as a trade-off that leaves researchers open to repeat responders (individuals who complete the questionnaire multiple times; Buchanan, 2000; Johnson, 2001). Detecting repeat responders would be easy if participants provided unique identifying information, such as names or social security numbers, but asking for such intrusive information would substantially reduce confidentiality and thus the number of participants. Fortunately, Web questionnaires allow a number of steps to minimize the effects of repeat responders. In our own research, we have used three major strategies: reducing the motivation to respond multiple times, using a proxy method for identifying participants, and directly asking participants whether they have completed the questionnaire before.

A major motivation for participants to respond multiple times is to see a range of possible feedback (e.g., how their personality scores would look if they answered the questions differently). Therefore, our first strategy was to give participants a direct link to all of the possible feedback options to allow them to satisfy their curiosity. Our second strategy was to identify repeat responders using the Internet protocol (IP) addresses that the Web server logs with each completed questionnaire. A single IP address can be associated with multiple responses submitted during a single session, such as by individuals taking the test again but changing their answers to see how the feedback changes. Thus, we eliminated repeated responses from the same individual at a single IP address. To avoid eliminating responses from different individuals using the same computer (e.g., roommates or people using public computer labs), we matched consecutive responses from the same IP address on several key demographic characteristics (e.g., gender, age, ethnicity) and when such a match was detected, we retained only the first response. Johnson (2001) offers another solution for detecting repeat responders: He suggests comparing the entire set of item responses in consecutive entries to identify duplicate or near-duplicate entries.

Some repeat responses could remain in the sample even after taking these steps to eliminate them. For example, participants might revisit the questionnaire to see whether their scores change over time. Thus, our third strategy was to add a question asking participants whether they had completed the questionnaire before. Only 3.4% responded that they had completed the questionnaire before. Most important, analyses showed that repeat responding (as identified by the question) did not change the findings in Srivastava et al.'s (2003) study on personality development. When repeat responding is of great concern, researchers can always take additional precautions such as requiring participants to provide a valid email address where they receive authorization to complete the questionnaire (Johnson, 2001).

Preconception 6: Internet Findings Are Not Consistent With Findings From Traditional Methods

Across a range of topics, evidence is accumulating that effects obtained using Internet methods are typically consistent with the effects from studies using traditional methods. Cross-method consistencies have been demonstrated for such constructs as self-monitoring (Buchanan & Smith, 1999), reaction-time studies (McGraw et al., 2000), and self-esteem (Robins, Trzesniewski, Tracy, Gosling, & Potter, 2002).

To directly test cross-method replicability, Srivastava et al. (2003) recently compared age and gender effects in their Web-based study on personality development with those found in personality development studies using traditional samples. The aim was straightforward: If the same questions were asked of both kinds of data, would the same answers emerge? Their findings show that age effects in the Internet data set and two traditional data sets were remarkably consistent for at least four of the Big Five personality dimensions, in spite of different instruments, different sampling methods, different cultures, and different age groups being studied. Similarly, the gender effects in the Internet data set were consistent with effects found in traditional data sets (Benet-Martinez & John, 1998; Costa, Terracciano, & McCrae, 2001; Feingold, 1994; Goldberg, Sweeney, Merenda, & Hughes, 1998).

These studies contribute to the growing body of evidence that psychological findings obtained using Web samples are consistent with findings obtained using traditional methods. Similar findings have been obtained with other psychological constructs, ranging from self-trust to sexuality, using both self-selected and assigned Web samples (Buchanan & Smith, 1999; Foster, Campbell, & Twenge, 2003; Johnson, 2000; McGraw et al., 2000; Pasveer & Ellard, 1998; Smith & Leigh, 1997). Nevertheless, this question has yet to be resolved conclusively, and more evidence is needed before we can be sure that the two methodologies are consistent. Of course, if the two methods do yield inconsistent findings, it should not be concluded automatically that the Internet method is the inaccurate one. Indeed, the real-world generalizability of findings from traditional methods often goes untested; for example, in the experimental social psychological literature, it has recently been noted, "a number of presumed social psychological truisms in fact have limited generalizability outside of the lab" (Skitka & Sargis, in press).

Summary

Table 1 summarizes the six preconceptions addressed in this article and contrasts them with the actual findings from our comparative analyses of traditional and Internet methods. Our analyses suggest that the samples gathered using Internet methods are at least as diverse as many of the samples already used in psychological research and are not unusually maladjusted. Internet samples are certainly not representative or even random samples of the general population, but neither are traditional samples in psychology. Moreover, the large sample sizes afforded by the Internet mean that even small proportions of participants (e.g., Latinos) are represented by large absolute numbers.

Our analyses also suggest that the data provided by Internet methods are of at least as good quality as those provided by traditional paper-and-pencil methods. This is evidenced by the finding that Web-questionnaire results generalize across presentation formats, do not appear to be tainted by false data or repeat responders, and are, so far, consistent with results from traditional methods. In short, the data collected from Internet methods are not as flawed as is commonly believed.

A final reason for encouraging researchers to use Internet methods when applicable is that doing so helps "give away" psychology to the public (Miller, 1969). By involving a large and broad population in their research and by giving interesting feedback to participants, researchers can use the Internet to stimulate public interest in psychological research.

The primary goal of our article has been to evaluate some widespread preconceptions about Internet data, with a particular focus on data from self-selected Internet samples. Our findings challenge the view that traditional methods are inherently superior to Internet methods. We have highlighted some of the strengths of Web questionnaires, such as large and diverse samples and motivated respondents, many of which also apply to other designs administered on the Internet (e.g., reaction-time experiments). Furthermore, we have demonstrated that traditional methods have their own weaknesses, such as overreliance on student samples and lack of anonymity.

At the same time, we do not want to suggest that Internet data are free from methodological constraints. Like every method, the Internet has certain drawbacks, such as the lack of control over the participant's environment and susceptibility to fake responses. Nor do we think Internet methods should replace traditional methods in all instances. Instead, there is room for both, and researchers should select whichever method suits their particular research goals. As with all research, the best studies will seek convergence across multiple methods.

In order to reap the benefits of Internet methods, the field of psychology needs to address the widespread distrust of Internet data. Our analyses show that many objections against Internet data are unfounded and that Internet methods provide many important advantages over traditional methods. Internet methods do have their own unique obstacles, such as repeat responders, but these obstacles can be overcome. As with any new methodology, caution is justified, but as we have shown, simple wholesale rejection of Internet data is not warranted. We urge psychology researchers to add the Internet as one more tool in their arsenal of research methods.

REFERENCES

American Psychological Association. (2003). Summary report of journal operations 2002. *American Psychologist, 58,* 663–664.

Argyris, C. (1968). Issues in evaluating laboratory education. *Industrial Relations, 8,* 28–40.

Azar, B. (2000). A web of research: They're fun, they're fast, and they save money, but do Web experiments yield quality results? *Monitor on Psychology, 31,* 42–47.

Benet-Martinez, V., & John, O. P. (1998). Los Cinco Grandes across cultures and ethnic groups: Multitrait–multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology, 75,* 729–750.

Best, S. J., Krueger, B., Hubbarb, C., & Smith, A. (2001). An assessment of the generalizability of Internet surveys. *Social Science Computer Review, 19,* 131–145.

Birnbaum, M. H. (2001). *Introduction to behavioral research on the Internet.* Upper Saddle River, NJ: Prentice Hall.

Birnbaum, M. H. (2004). Methodological and ethical issues in conducting social psychological research via the Internet. In C. Sansone, C. C.

Morf, & A. T. Panter (Eds.), *Handbook of methods in social psychology* (pp. 359–382). Thousand Oaks, CA: Sage.

Bowker, D., & Dillman, D. A. (2000). *An experimental evaluation of left and right oriented screens for Web questionnaires.* Retrieved July 6, 2003, from http://survey.sesrc.wsu.edu/dillman/papers/AAPORpaper00.pdf

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

Chang, L. C., & Krosnick, J. A. (2003). *National survey via RDD telephone interviewing vs. the Internet: Comparing sample representativeness and response quality.* Manuscript submitted for publication.

Costa, P. T., Jr., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology, 81,* 322–331.

Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly, 64,* 464–494.

Dillman, D. A. (1999). *Mail and Internet surveys: The tailored design method.* Hoboken, NJ: Wiley.

Dillman, D. A., Tortora, R. D., & Bowker, D. (1998). Principles for constructing Web surveys (SESRC Tech. Rep. No. 98-50). Retrieved July 6, 2003, from http://survey.sesrc.wsu.edu/dillman/papers/websurveyppr.pdf

Dillman, D. A., Tortora, R. D., Conradt, P. P., & Bowker, D. (1998, August). *Influence of plain vs. fancy design on response rates for Web surveys.* Paper presented at the Joint Statistical Meetings. Retrieved July 6, 2003, from http://survey.sesrc.wsu.edu/dillman/papers/asa98ppr.pdf

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116,* 429–456.

Foster, J. D., Campbell, W. K., & Twenge, J. M. (2003). Individual differences in narcissism: Inflated self-views across the lifespan and around the world. *Journal of Research in Personality, 37,* 469–486.

Fraley, R. C. (2004). *How to conduct behavioral research over the Internet: A beginner's guide to HTML and CGI/Perl.* New York: Guilford Press.

Goldberg, L. R. (2003). *International personality item pool.* Retrieved April 1, 2003, from http://ipip.ori.org/new_home.htm

Goldberg, L. R., Sweeney, D., Merenda, P. F., & Hughes, J. E. (1998). Demographic variables and personality: The effects of gender, age, education, and ethnic/racial status on self-descriptions of personality attributes. *Personality and Individual Differences, 24,* 393–403.

Gosling, S. D., & Bonnenburg, A. V. (1998). An integrative approach to personality research in anthro-zoology: Ratings of six species of pets and their owners. *Anthrozoos, 11,* 148–156.

Hewson, C. M., Laurent, D., & Vogel, C. M. (1996). Proper methodologies for psychological and sociological studies conducted via the Internet. *Behavior Research Methods, Instruments, & Computers, 28,* 186–191.

Institute for Scientific Information (ISI). (n.d.). *Social sciences citation index.* Retrieved September 2, 2003, from ISI database.

Jackson, C. W., & Pollard, J. C. (1966). Some nondeprivation variables which influence the "effects" of experimental sensory deprivation. *Journal of Abnormal Psychology, 71,* 383–388.

John, O. P., & Benet-Martinez, V. (2000). Measurement: Reliability, construct validation, and scale construction. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 339–369). New York: Cambridge University Press.

John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–139). New York: Guilford Press.

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

Johnson, J. A. (2001, May). *Screening massively large data sets for non-responsiveness in Web-based personality inventories.* Invited talk to the joint Bielefeld-Groningen Personality Research Group, University of Groningen, The Netherlands.

Jourard, S. M. (1968). *Disclosing man to himself.* Princeton, NJ: Van Nostrand.

Kieley, J. M. (1996). CGI Scripts: Gateways to World-Wide Web power. *Behavioral Research Methods, Instruments, & Computers, 28,* 165–169.

Kiesler, S., & Sproull, L. S. (1986). Response effects in the electronic survey. *Public Opinion Quarterly, 50,* 402–413.

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., Kiesler, S., Boneva, B., Cummings, J., Helgeson, V., & Crawford, A. (2002). Internet paradox revisited. *Journal of Social Issues, 58,* 49–74.

Kraut, R. E., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2003). *Psychological research online: Opportunities and challenges.* Report prepared for the American Psychological Association's Taskforce on the Internet and Psychological Research. Retrieved August 28, 2003, from http://www-2.cs.cmu.edu/afs/cs.cmu.edu/user/kraut/www/RKraut.site.files/articles/kraut03-PsychologicalResearchOnline.pdf

Kraut, R., Patterson, V., Lundmark, M., 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.

Lebo, H. (2000). *The UCLA Internet report: Surveying the digital future.* Retrieved July 6, 2003, from University of California, Los Angeles, Center for Communication Policy Web site: http://www.ccp.ucla.edu/UCLA-Internet-Report-2000.pdf

Lenhart, A. (2000). *Who's not online: 57% of those without Internet access say they do not plan to log on.* Retrieved July 6, 2003, from Pew Internet & American Life Project Web site: http://www.pewinternet.org/reports/toc.asp?Report=21

Lenhart, A., Horrigan, J., Rainie, L., Allen, K., Boyce, A., Madden, M., & O'Grady, E. (2003). The ever-shifting Internet population: A new look at Internet access and the digital divide. Retrieved July 6, 2003, from Pew Internet & American Life Project Web site: http://www.pewinternet.org/reports/toc.asp?Report=88

Lerner, J. S., Gonzales, R. M., Small, D. A., & Fischhoff, B. (2003). Effects of fear and anger on perceived risk of terrorism: A national field experiment. *Psychological Science, 14,* 144–150.

Levine, S., Ancill, R. J., & Roberts, A. P. (1989). Assessment of suicide risk by computer-delivered self-rating questionnaire: Preliminary findings. *Acta Psychiatrica Scandinavica, 80,* 216–220.

Locke, S. D., & Gilbert, B. O. (1995). Method of psychological assessment, self-disclosure, and experiential differences: A study of computer, questionnaire, and interview assessment formats. *Journal of Social Behavior & Personality, 10,* 255–263.

McGraw, K. O., Tew, M. D., & Williams, J. E. (2000). The integrity of Web-delivered experiments: Can you trust the data? *Psychological Science, 11,* 502–506.

Mendenhall, D. (2001, August 22). Web doesn't promote isolation, study says. Retrieved July 6, 2003, from *San Francisco Chronicle* Web site: http://sfgate.com/cgi-bin/article.cgi?file=/chronicle/archive/2001/08/22/BU98992.DTL

Mezzacappa, E. (2000). Letter to the Editor. *APS Observer, 13,* 10.

Michalak, E. E., & Szabo, A. (1998). Guidelines for Internet research: An update. *European Psychologist, 3,* 70–75.

Miller, G. A. (1969). Psychology as a means of promoting human welfare. *American Psychologist, 24,* 1063–1075.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist, 38,* 379–387.

Morrow, R. H., & McKee, A. J. (1998). CGI scripts: A strategy for between-subjects experimental group assignment on the World-Wide Web. *Behavior Research Methods, Instruments, & Computers, 30,* 306–308.

National Center for Education Statistics. (2000). *Digest of education statistics 2000.* Washington, DC: U.S. Department of Education.

Nosek, B. A., Banaji, M., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration Web site. *Group Dynamics, 6,* 101–115.

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

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.

Peterson, K. S. (2003, May 11). Controversial study says personality evolves with age. Retrieved July 6, 2003, from *USA Today* Web site: www.usatoday.com/news/health/2003-05-11-personality-age_x.htm

Pettit, F. A. (2002). A comparison of World-Wide Web and paper-and-pencil personality questionnaires. *Behavior Research Methods, Instruments, & Computers, 34,* 50–54.

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.

Rentfrow, P. J., & Gosling, S. D. (2003a). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology, 84,* 1236–1256.

Rentfrow, P. J., & Gosling, S. D. (2003b). *The personality of America: An interplay of social and physical environments.* Unpublished manuscript.

Richman, W. L., Kiesler, S., Weisband, S., & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology, 84,* 754–775.

Robins, R. W., Trzesniewski, K. H., Tracy, J. L., Gosling, S. D., & Potter, J. (2002). Global self-esteem across the life span. *Psychology and Aging, 17,* 423–434.

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

Sears, D. O. (1986). College sophomores in the lab: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology, 51,* 515–530.

Skitka, L. J., & Sargis, E. G. (in press). Social psychological research and the Internet: The promise and the perils of a new methodological frontier. In Y. Amichai-Hamburger (Ed.), *The social net: The social psychology of the Internet.* Oxford, England: Oxford University Press.

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.

Social sciences citation index: Journal citation reports. (2002). Philadelphia: Institute for Scientific Information.

Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology, 84,* 1041–1053.

Taylor, H. (2000). Does Internet research work? Comparing online survey results with telephone survey. *International Journal of Market Research, 42,* 51–63.

Turner, C. F., Ku, L., Rogers, S. M., Lindberg, L. D., Pleck, J. H., & Stonenstein, F. L. (1998, May 8). Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science, 280,* 867–873.

U.S. Census Bureau. (2000). Population (Section 1). In *Statistical abstracts of the United States* (pp. 1–62). Washington, DC: U.S. Department of Commerce.

U.S. Department of Commerce. (2000). Falling through the net: Toward digital inclusion. Retrieved July 6, 2003, from http://search.ntia.doc.gov/pdf/fttn00.pdf

Vazire, S., & Gosling, S. D. (in press). E-perceptions: Personality impressions based on personal Web sites. *Journal of Personality and Social Psychology.*

Walsh, J. P., Kiesler, S., Sproull, L. S., & Hesse, B. W. (1992). Self-selected and randomly selected respondents in a computer network survey. *Public Opinion Quarterly, 56,* 241–244.