5 The Potential of Electronic Communication and Information Technologies as Research Tools: Promise and Perils for the Future of Communication Research

NICHOLAS A. PALOMARES
University of California, Davis

ANDREW J. FLANAGIN
University of California, Santa Barbara

Scholars across several disciplines praise the methodological advantages afforded to research endeavors by the use of electronic communication and information technologies. At the same time, however, scholars note several disadvantages stemming from the use of these tools. To assess the utility of electronic communication and information technologies as research tools, we synthesize work on the methodological issues surrounding the use of these technologies for the future of communication research. Our overarching goal is to review prior and current perspectives in order to appropriately assess and critique the use of electronic tools in research pursuits while informing future applications of these tools in the field. Throughout the chapter we articulate the argument that sound theoretical and methodological practices determine the appropriateness of the application of any particular tool. Accordingly, our focus is on the specific methodological concerns that arise with the use of electronic technologies in the conduct of research. We organize our discussion into six major sections: sampling issues; data integrity concerns; including the reliability and validity of research design and measures; the potential afforded by electronic communication and information technologies; contemporary ethical considerations; an explicit assessment of the future of communication research in view of the application of electronic technologies; and a concluding section that establishes the contribution of our review and briefly outlines outstanding issues.

AUTHORS’ NOTE: An abbreviated version of this chapter was presented at the annual conference of the National Communication Association, November, 2004, in Chicago, Illinois. The authors thank Miriam Metzger and Ron Rice for their thoughtful comments and suggestions on an earlier version of the chapter.
INTRODUCTION

Scholars across several disciplines praise the methodological advantages provided by electronic communication and information technologies. Expounded benefits include minimizing time spent on mundane tasks such as data entry and participant scheduling (Ilieva, Baron, & Healey, 2002; Reips, 2000, 2002a; Schmidt, 1997; Stanton & Rogelberg, 2002), the ability to acquire large samples quickly (Birnbaum, 2000, 2001; Christians & Chen, 2004; Krantz & Dalal, 2000), and the capacity for novel data capture using technologies such as automated internet data mining (Burton & Walther, 2001) and web-based observational cameras (Stanton & Rogelberg, 2002). At the same time, however, scholars have noted several disadvantages stemming from the use of electronic technologies, such as a limited ability to acquire representative samples (Brenner, 2002), greater infringement on participants’ privacy (Elgesem, 2002), and an increased likelihood of participant fraud (Madge & O’Connor, 2002). Indeed, even the same methods are viewed both positively and negatively. For example, participant self-selection has been framed both as an advantage for (Reips, 2000) and as an obstacle to (Mustanski, 2001; Sills & Song, 2002) sound research.

Nonetheless, the use of electronic communication and information technologies in the conduct of research has increased rapidly over the past decade with the growth of the internet and Web, and many researchers affirm continued use of these tools (Musche & Reips, 2000). In fact, the application of these technologies is evident in the field of communication (e.g., Eveland, Cortese, Park, & Dunwoody, 2004; Flanagan & Metzger, 2003a; Kalyanaraman & Sundar, 2003; Lee, 2004; Lee & Nass, 2002; Lemus, Seibold, Flanagan, & Metzger, 2004; Palomares, 2004; Tidwell & Walther, 2002), and increasing usage seems inevitable in the future. In spite of remarkable growth in the use of these tools in the conduct of research, academic inquiry into their usage has not kept pace. To assess the utility of using electronic communication and information technologies as research tools, in this chapter we synthesize and critically assess work on the methodological issues surrounding the use of these technologies for the future of communication research.

Conceptual Boundaries and Overview

Research and reviews focusing on contemporary technologies range widely in both the terms used to describe these tools [e.g., information and communication technologies (Lievrouw et al., 2001), electronic media (Culnan & Markus, 1987), and advanced information technologies (Huber, 1990)], and their definitions. Based on the ubiquity of specific tools and their prevalence in research endeavors, we focus on electronic communication and information technologies, defined as electronic tools that facilitate communication and information capture, transmission, processing, and storage among people who typically are not co-located. Among the most central electronic communication and information tools used for research today are the internet (i.e., interconnected computers operating as an infrastructure for digital global communication), the Web (i.e., the system of computer servers using a graphical user interface accessed via the internet that includes documents and other files interconnected by hyperlinks), and electronic mail (email). In addition, other internet-based technologies (e.g., newsgroups, chat functions, web logs, and multiuser dungeons, or MUDs) and non-internet-based technologies (e.g., virtual reality, active badges, and mobile telephones) are included in this review when appropriate and informative.

We emphasize research methods that have taken significant advantage of electronic communication and information technologies and to which scholars have dedicated considerable empirical work and discussion. Accordingly, primary attention is devoted to survey research, experimental research, and content analysis, with less attention on other methodologies, such as interaction analysis and focus groups. In this manner, discussion of some research tools is more limited, yet reflects the current use of electronic tools in the conduct of scientific research.

Our overarching goal is to review current research perspectives in order to appropriately assess and critique the use of electronic technologies in research endeavors while informing future applications of these tools in communication research. Throughout the chapter we articulate the argument that sound methodological practices and theoretical concerns determine the rigor and appropriateness of the application of any particular tool(s). Clearly, contemporary technologies offer appealing advantages that have enormous utility for communication scholars. The attractiveness of these tools, however, also creates potential for misuse and abuse and often relies on substantial new skills and resources. Scholars, in other words, may realize benefits when using these technologies, while they may neglect important drawbacks that can potentially surface.

A secondary goal of this chapter, therefore, is to engender a simultaneous appreciation of the utility and potential pitfalls scholars face when invoking electronic communication and information technologies in their research. In this manner, we aim to foster an informed awareness of the methodological issues accompanying the use of electronic technologies for research pursuits. Highlighting ambiguities and unresolved issues, we also hope to spark discussion and continued research on the promise and peril of using these technologies as research tools within and outside the field of communication, so that future research maintains a high level of methodological scrutiny and rigor.

We structure this chapter according to crucial considerations manifest across research methods, rather than by technologies that can aid research endeavors. By so doing, we systematize the literature around major methodological issues integral to most types of research and highlight the bases of sound methods, regardless of the tools invoked in the conduct of research. This structure serves to frame methodological issues in a way that is highly heuristic to scholars' future research endeavors. That is, scholars can use our review by extrapolating the methodological issues we discuss onto specific forms of a wide range of research activities. Accordingly, we organize our discussion into six major sections, focusing on
sampling issues; data integrity concerns, including the reliability and validity of research design and measures; the potential afforded by electronic communication and information technologies; ethical considerations; an explicit assessment of the future of communication research in view of the application of electronic technologies; and a concluding section that establishes the contribution of our review and briefly outlines outstanding issues.

SAMPLING

Sampling—the procedure by which researchers acquire their units of analysis—is a primary concern of all sound research (e.g., surveys, experiments, and content analyses). Researchers must weigh several alternatives when determining the best sampling method, including decisions regarding who should be solicited, the necessary number of participants, and the type of sampling required (e.g., probability versus nonprobability). Internet technologies, at times, appear to simplify sampling issues. For instance, different time frames may generate different response rates or different kinds of respondents. Indeed, faster responses from participants can be obtained via the internet (Christians & Chen, 2004; Reips, 2002a): An experiment that randomly assigned individuals to complete the same questionnaire via either email or traditional mail found that email questionnaires were completed and returned in almost half the time (9.22 days versus 16.43; Tuten, Urban, & Bosnjak, 2002), a finding corroborated by a review of 11 studies showing that email surveys were returned in an average of 5.59 days, compared to 12.21 days for mail surveys (Ilieva et al., 2002). Ultimately, reductions in response time can result in differences in response rates. In addition, the Internet often affords larger sample sizes compared to non-internet samples, due to decreased costs per respondent (Birnbaum, 2000; Krantz & Dalal, 2000; Reips, 2000), which might yield different kinds of respondents. Nonetheless, in spite of some apparent advantages, such affordances should be framed by an informed awareness of relevant methodological considerations, as detailed next.

Online Participant Recruitment

One of the first considerations for using Internet technologies to collect responses from online participants is how to inform individuals about the research study. Several strategies exist. Announcing the study in Internet newsgroups (e.g., alt.use.net.surveys) and online communities (such as MUDs), submitting the study’s website to search engines (e.g., google.com or yahoo.com) and web portals that list online studies (e.g., the American Psychological Society’s listing: http://psych.hanover.edu/Research/exponnet.html), using specialized metatags (i.e., descriptors of a particular webpage used for indexing by search engines), sending email solicitations, submitting messages to an email listserv (such as the Communication Research and Theory Network accessible at http://lists1.cac.psu.edu/cgi-bin/wa?A0=CRTNET), and other methods have all been suggested and used as means of publicizing studies and soliciting participants (e.g., Batagelj & Vehovar, 1998; Birnbaum, 2000; Buchanan & Smith, 1999; Krantz & Dalal, 2000; Reips, 2001, 2002b; Williams & Robson, 2004).

Even though they can prove useful, these methods for recruiting participants also pose problems. For example, these methods do not guarantee that all participants who know about the study actually participate (Vehovar, Batagelj, Manfreda, & Zaletel, 2002). In fact, the proportion of those who know about a study and those who choose to participate is difficult to determine (Couper, 2000). The use of web portals, newsgroups, and search engines, for example, does not enable a researcher to determine who read information about a particular study but chose not to participate. There are methods to determine the number of individuals who visit a study’s website and choose not to participate (Werner, 2002); however, measuring those who knew about the study but did not visit the study’s website remains virtually impossible.

Alternatively, with email solicitations, participant response rates can be assessed based on the number of email messages sent and the number of individuals who participated in the study. Using email solicitations, however, requires a list of current and accurate email addresses, and assumes that individuals check their email accounts on a regular basis. In addition, participants who forward email messages to others, such as their friends or colleagues, can cause problems for determining response rates and adhering to target populations. Overall, although the efficiencies of Internet-based tools for sampling are appealing, they are accompanied by their own difficulties.

Online Response Rates

Despite such problems with participant recruitment and the accurate determination of response rates, a great deal of research has focused on accurately determining these rates for online surveys. Reviews of online surveys show response rates of around 30–40%. Cook, Heath, and Thompson (2000), for example, reviewed 68 online surveys and found that the average response rate was 40%. Comley (2000), in a review of 42 Web-based pop-up surveys, showed that response rates ranged from 9% to 48%, with an average of 27%. In each case, these reviews included only surveys in which a specific number of potential participants were available, thus allowing for response rates to be determined highly accurately. Comparisons of online versus mail surveys, however, seem to demonstrate a slightly higher response rate for mail surveys. For example, a review of 10 studies by Ilieva et al. (2002) found that the average response rate for online surveys was 39%, as compared to 46% for mail surveys (for a review with similar conclusions, see Cho & LaRose, 1999).

Two experiments have directly tested response rate differences between online and traditional mail questionnaires (Couper, Blair, & Triplett, 1999; Truell,
Bartlett, & Alexander, 2002). Both experiments randomly assigned individuals to participate in one of two modes of data collection (i.e., traditional mail questionnaire or online questionnaire). Even though Truell et al. found no significant difference between online (51%) and mail (53%) response rates, Couper et al. noted a 71% response rate for the mail questionnaire condition and 43% for the online questionnaire condition—a 28% difference. Potential explanations for this discrepancy include participants’ frequency of email use and differences in question numbering between studies. Specifically, not all participants in the Couper et al. study used their employer-assigned email addresses regularly, perhaps exaggerating the difference in response rates between mail and online modes. In addition, a confound in the numbering of questions for Couper et al. might explain differences between the two studies, at least in part: Whereas Truell et al. numbered items identically in both conditions, Couper et al.’s email questionnaire contained 14 subsections with 94 items (numbered from 1–94), but for the mail questionnaire items within each subsection were numbered separately. A potential result is that participants in the email condition of Couper et al.’s study may have been deterred from responding by the large number of questions more so than those in the mail condition. Further research is necessary to resolve such discrepancies; however, these results raise the possibility that subtle factors can influence online response rates. To date, scholars have focused on three main sources of influence related to online response rates—researcher-based factors, participant-based factors, and technical factors—each of which we cover in turn.

Researcher-based factors. Researchers have examined several factors within their control that can influence response rates online. Similar to non–Internet-based research, prenotifications, personalized letters, and follow-up contacts and reminders are associated with higher response rates, according to qualitative and quantitative reviews (Cook et al., 2000; Sheehan, 2001; Tuten et al., 2002). Another method to increase response rates is a promise to share results upon study completion, which is relatively fast, easy, and inexpensive via the Internet (Bosnjak & Baticic, 2002; Kleinman, 2004). Using a multimode technique for data collection also increases response rates for online research (Dillman, Phelps, Tortora, Swift, Kohrell, & Berck, 2001; Ilieva et al., 2002; Schaefer & Dillman, 1998). Using this technique, a complete data collection strategy via one mode (e.g., email) is implemented and, after a short amount of time, is followed by another attempt to collect data in a different mode (e.g., mail).

Mixed results regarding response rates exist for other researcher-based factors. Response rate data regarding survey length have been viewed as equivocal by Sheehan (2001) and as a nonfactor by Bosnjak and Baticic (2002), who reported that online participants indicated a willingness to spend at least 10 minutes on a given study. Even though some research has shown no relationship between incentives and response rates (e.g., Frick, Bachtiger, & Reips, 2001), a meta-analysis of 68 online surveys (Cook et al., 2000) revealed that incentives had a negative relation to response rates, perhaps, as the authors speculate, because researchers anticipated low response and, therefore, felt the need to offer incentives. A meta-analysis of mail surveys revealed that up-front incentives (i.e., before survey completion) increased responses by between 8% (for nonmonetary rewards) and 19% (for monetary rewards), whereas incentives contingent upon survey completion did not have an effect (Church, 1993). Perhaps, then, incentives should be given to participants prior to the completion of online studies as well.

Participant-based factors. The effects of individual characteristics also have been examined with regard to online response rates. Participants’ curiosity and interest in the research process (Bosnjak & Baticic, 2002), education level (Cook et al., 2000), experience with research participation, and computer literacy levels (Vehovar et al., 2002) are positively associated with participation in research studies. Another important participant-based factor is the research issue’s importance (i.e., timeliness, salience) to participants (Cook et al., 2000; Sheehan, 2001; Vehovar et al., 2002). The more interest participants have in the topic, the higher the response rates for online studies. Finally, assurances of confidentiality and anonymity result in higher response rates (Couper, 2000; Couper et al., 1999; Mustanski, 2001; Sassenberg & Kreutz, 2002). Additional individual difference factors that are not currently addressed in the literature but might relate to response rates are the amount of free time that potential participants have, the amount of time they spend on the Internet, concerns with online privacy and security, and, perhaps, even participants’ outspokenness or eagerness to share ideas online.

Technical factors. Finally, technical factors may affect participants’ response rates online. Low modem speeds (or other bandwidth issues), incompatible Web browser versions, internet service provider connection time costs, unreliable connections, and other technical factors have been proposed to decrease response rates (Couper, 2000). A decrease in response rates due to internet service provider costs may be exaggerated in places where flat rates and unlimited connection time for internet access providers are rare, where recipients are charged for receipt of messages, or where internet access is not relatively affordable (e.g., China). In addition, email address errors, firewalls that inhibit message delivery, and anti-spam software all hinder the research process through the resultant nondelivery of messages to potential participants (see, for example, Vehovar et al., 2002). Due to its potential importance in understanding the most efficient use of internet-based technologies for participant solicitation, further research on technical factors is required.

Generalizability

Generalizability refers to the extent to which one is able to infer accurately from a sample to a larger, specified population (Singleton & Straits, 2005). There is a long-standing critique of social scientific research for being too focused on student samples and, therefore, unable to generalize accurately to larger, more diverse populations (see, for example, Smart, 1966). Some scholars, on the other hand, do not consider convenience samples of students to be problematic because certain phenomena span across human behavior and therefore will endure across
all types of samples (see, for example, Birnbaum, 2001). Nevertheless, researchers currently have proposed the internet as a tool for sampling nonstudent populations in order to gain more generalizable (i.e., representative) samples (e.g., Buchanan, 2000). Getting a truly representative sample of a large population via the internet, however, is problematic (e.g., Best & Krueger, 2002; Brenner, 2002; Stanton & Rogelberg, 2002).  

For a sample to be representative, every unit in the target population (i.e., those to whom one wishes to generalize) must have an equal chance of being included in the sample (Babbie, 1998; Cook & Campbell, 1979; Singleton & Straits, 2005). To achieve this, a reliable and exhaustive list of the target population must exist. Using such lists, means of random selection (such as random digit dialing) are used to select potential respondents, and follow-up techniques (such as incentives) are often employed to gain compliance. In the absence of such ideal lists, researchers have relied on means of communication, such as the telephone, that include a high proportion of the target population (for example, current estimates for telephone saturation in the U.S. are 93%; Best & Krueger, 2002).  

Internet-based technologies, however, do not have the equivalent of a telephone directory, nor are webpage uniform resource locators (URLs), email addresses, or other means of identification sufficiently standardized to enable effective means of random selection. Consequently, the frame population (i.e., those for whom one has contact information) of internet users diverges significantly from the target population of internet users (Couper, 2000). Even if random selection of internet users were possible, current internet saturation (both direct and indirect) in the U.S. is approximately 70% (Internet World Stats, 2004), suggesting that a large portion of the population would remain inaccessible to research through this means. These considerations highlight the digital divide that separates those with access to advanced technologies from those who lack such access (bridges.org, 2003). Moreover, internet users are significantly different from non-users on many demographic, communication, expertise, and attitudinal variables (see Katz & Rice, 2002).  

Limitations to generalizability are even more pronounced at the global level (Krantz & Dalal, 2000). Countries across the globe vary drastically in terms of their internet saturation levels and may vary based on economic, social, and cultural factors. The United States Central Intelligence Agency (2003) estimates, for example, that Turkey has almost twice as many internet users as Iran, despite demographically comparable populations. Discrepancies like this can adversely affect a researcher’s ability to make accurate generalizations to a global target population.

Internet Versus Non-Internet Sample Comparisons

Researchers have compared demographic characteristics of internet-based samples to non–internet-based samples as a means to assess generalizability. Much of this research, however, has been limited to comparisons between internet samples and non-internet college samples. Nevertheless, compared to college samples, internet samples tend to be more representative of the U.S. as a whole (Bailey, Foote, & Throckmorton, 2000; Birnbaum, 2000; Reips, 2000), yielding older participants (Bailey et al., 2000; Birnbaum, 2000) and providing more equal sex composition (Birnbaum, 2000; sex composition was nearly equal for internet samples, whereas college samples were composed of 73% females). Research also suggests, however, that internet samples are less ethnically diverse than college samples (Bailey et al., 2000), although examination of individuals from various cultures is possible via the internet if they are sought out actively (Reips, 2000). Substantiating these findings, a review of over 30 studies that used internet samples (Krantz & Dalal, 2000) concluded that (a) there is no consensus regarding the sex composition of internet samples, but there is a small trend toward sex equal distributions; (b) the large majority of internet samples is ethnically composed of White participants with a tendency to contain North Americans, particularly from the U.S.; and (c) the mean and median ages across studies ranged between 26 and 35 years. Internet samples, thus, are fairly representative of the general U.S. population regarding a small set of factors; however, crucial differences exist that call for a need to conduct future research, especially comparing internet samples with non-internet samples composed of nonstudent participants. Katz & Rice (2002) for example, compare demographics between their national random telephone samples and U.S. census data in order to assess the representativeness of their samples.

Improving the Representativeness of Internet Samples

Scholars have suggested and critiqued means to improve the representativeness of internet samples. For example, Best and Krueger (2002) reviewed two common solutions for nonrepresentative samples: weighting and the form-resistant correlation hypothesis. First, data from internet respondents can be weighted by the incidence of known characteristics in the target population using, for instance, census data. Weighting, however, must be conducted with variables that are highly correlated with the key variables under investigation.

The second option is to assert the form-resistant correlation hypothesis by stating that a nonrepresentative sample may shift the means of certain variables but the correlations between variables remain the same. By this logic, the variables measured with internet samples simply shift their positions along their respective axes while maintaining the same correlation among variables. In this manner, although the means of variables measured using internet samples may be unrepresentative of larger populations, the associations between these variables are representative.

The authors critiqued both methods for potentially ignoring the fact that internet and non-internet samples could differ in terms of factors that mediate or moderate responses to the variables measured. Best and Krueger (2002) demonstrated and confirmed this claim empirically, finding that their email and telephone samples were different and that weighting or using the form-resistant correlation hypothesis...
did not increase the generalizability of the email sample's results. Similarly, Vehovar, Manfreda, and Batagelj (1999) confirmed the limited ability to increase representativeness of internet samples via weighting and Katz and Rice (2002) chose not to weight data because in their assessment costs typically outweigh benefits. In light of these findings, a relevant issue to consider is under what circumstances researchers should use internet samples, an issue we address in the next section.

The Utility of Sampling via the Internet

Despite these participant recruitment, response rate, and representativeness limitations, the internet can be a useful tool for acquiring samples from which accurate generalizations are made. Couper (2000) offered several suggestions to obtain representative samples on the internet. Target populations can be limited to those for whom a list of email addresses is available, thus providing an inclusive frame population from which a probability sample can be drawn. Such a sample would be representative of the individuals included in the list of email addresses. For example, if an email list of every employee from a particular organization were available, then a probability sample could be drawn, and based on those data generalizations could be made to the organization as a whole. The sample of employees from the organization, therefore, would be representative of the organization's employees (although not representative of other organizations).

As another possibility to use the internet to acquire representative samples, random digit dialing of U.S. citizens could be used in conjunction with web or email surveys, such that once called, participants would be asked for their email addresses or to visit a website (e.g., HarrisInteractive.com, 2004, currently employs this technique). This method could cut down personnel costs because telephone calls would be significantly shorter than traditional telephone surveys. In addition, experiments could be conducted with representative samples using this telephone-internet hybrid method. Such a technique, however, would not include those without access to internet technologies. Nonetheless, these methods suggest that there are ways to access samples via the internet and still reap the benefits of the technology while increasing the ability to make accurate generalizations.

Probability sampling is not always the best option (see Shapiro, 2002; Singleton & Straits, 2005). Rather, some research projects call for sampling methods (e.g., purposive, quota sampling) that internet-based tools are well equipped to provide. Indeed, the internet is useful for acquiring samples of individuals from specialized or hard-to-reach populations (e.g., Fortune 500 or volunteer organizations) through purposive sampling (Birnbaum, 2000; Christians & Chen, 2004; Duffy, 2002; Madge & O’Connor, 2002; Mustanski, 2001; Reips, 2002a; Schmidt, 1997; Tuten et al. 2002). Moreover, as noted earlier, internet samples enable researchers who are heavily reliant on university student samples to gain access to a broader range of participants. In this manner, the internet provides a means to access nonstudent samples, potentially bolstering the generalizability of results (see Shapiro, 2002).

Finally, research that focuses on the characteristics and behaviors of only internet users need not attempt to represent the wider population (i.e., those without internet access), although sampling issues still apply if generalizing to all internet users is a research goal. Strategies to capture internet user data include surveys, tracking of online behaviors, and diary methods that record internet usage. In addition, creative strategies for accessing relatively new internet users also exist. For example, in research by Kraut et al. (1998), families received hardware, software, and internet connections in return for access to their internet activities and survey and interview responses. Such quid pro quo arrangements, although expensive, enable researchers to capture extremely comprehensive data on internet use. Nevertheless results obtained in this manner may suffer from limitations to their generalizability.

Unobtrusive Data Obtained via the Internet

A final benefit of internet-based research tools is the ability to acquire data not attained via direct interaction with individuals. Unobtrusive data readily obtained online include (but are not limited to) online postings of text messages, online advertisements, information on websites, online news articles, Web-server activity logs, newsgroup archives, organizational documents, and other units of analysis that do not include human participation in a research study. Rice (1990) discusses the use of computer systems as unobtrusive monitors of data on communication usage, flows, and content, and discusses related issues of sampling, validity, reliability, and ethics. Ramírez, Walther, Burgoon, and Sunnafrank (2002) discuss such “extractive” data collection strategies, noting that ample social information can be obtained in this manner. Analyses of these data could occur through content analysis (Krippendorff, 1980, 2004; Neuendorf, 2002; see, for example, Kleinnman, 2004), network analysis of relations (Rice, 1990; Rice & Shook, 1990), or inferential statistics (Flanagan & Spivey, 2004). Limitations to these data collection methods may include sophisticated technical knowledge required to obtain or extract data in some cases and ethical and legal concerns (discussed at greater length later in this chapter). Potential problems regarding the use of messages posted online (e.g., Usenet messages, chat room postings, or online communities) are a lack of knowledge about the demographic characteristics of message senders; interpreting communicators’ intended meanings, which can present more difficulty when conversations are text-only (Williams & Robson, 2004); and more effort establishing researcher visibility and intimacy in ethnographic research (LeBesco, 2004).

Scholars have begun to realize the potential of unobtrusive data collection online, and in other venues pre-dating the Web (Burton & Walther, 2001; Foot, Schneider, Dougherty, Xenos, & Larsen, 2003; Kleinnman, 2002; Krippendorff, 2004; Majchrzak, Rice, Malhotra, King, & Ba, 2000; McLaughlin, Goldberg, Ellison,
Lucas, 1999; Mitra & Cohen, 1999; Neuendorf, 2002; Rice, 1982, 1990; Rice & Shook, 1990; Rössler, 2002; Stanton & Rogelberg, 2002). Of course, with this potential come myriad issues for researchers to consider. For example, how might researchers systematically sample from among the many millions of websites available (whether seeking a random, purposive, or quota sample), which of the thousands of newsgroups (that date back to 1981 with about 800 million posts via http://groups.google.com/) should one sample, and how does one account for the dynamic nature of the internet, which is constantly undergoing changes (cf., Krippendorff, 2004; Neuendorf, 2002)? Relatedly, scholars conducting computer-aided content analyses to automatically extract certain words or phrases from electronic texts must be aware of potential problems inherent in the programs. For example, computer programs can have difficulty counting typographical errors, misspelled words, words that are not uniformly hyphenated within and across texts (e.g., email versus email), numbers in numeral versus alphabetic format (e.g., 5 versus five), words differentially conjugated (e.g., run, ran, running), words that are typographically similar but semantically different (e.g., to play a game, to play an instrument, to perform in a play, to play with a toy, to play versus being serious), and words that differ typographically yet are semantically similar (e.g., empty, unfilled, vacant). Krippendorff (2004) presented a thorough discussion of potential problems in this domain, as well as information on computer programs and other means that can potentially resolve such issues.

In sum, although internet-based tools have great potential for aiding research endeavors, researchers must be cognizant of the implications of their sampling method and select the one that is most appropriate for their research goals.

DATA INTEGRITY

Data integrity is a crucial issue when using electronic communication and information technologies as research tools. In short, data integrity refers to the quality of data and the associated results. Data should be both reliable and valid. That is, data should be measured in a way that yields consistent results (i.e., reliable) and should be measured by means that accurately reflect the concept under study (i.e., valid). Even though reliability and validity are core concerns of all research, electronically based methods can create unique problems in these regards. In this section, we focus on some of the major threats to data integrity posed by the use of electronic tools for conducting all types of research. First, we discuss threats to the validity of results, including social desirability, self-selection biases, participant drop-out and mortality, and other threats to validity. Next, we focus on the reliability and validity of concept measurement via technologies. Then, we discuss factors related to item response length and quality. Finally, we consider issues regarding data interdependence based on measuring and observing nonindependent human participants and other units of analysis.

Threats to the Validity of Results

The issue of validity of results focuses on the accuracy of the results obtained. Valid relationships between variables result when the relation does not stem from an unspecified variable, a methodological flaw, or other confound. Several factors can threaten result validity, including social desirability, demand characteristics, participant drop-out and mortality, and other issues, as discussed next.

Social desirability. Social desirability refers to the tendency of research participants to respond to questions or behave in ways that are normatively appropriate (i.e., socially desired), although the term is used under many guises (see, for a review, Richman, Kiesler, Weisband, & Drasgow, 1999). Social desirability can affect the extent to which research results obtained accurately reflect participants’ true perceptions and behaviors. Evidence on social desirability in the electronic environment has been mixed. Some research has found less social desirability for online compared to paper-and-pencil studies (Kiesler & Sproull, 1986), whereas other research showed no difference across these modes (Pettit, 2002). Computerized testing has also been shown to lead to more social desirability than paper-and-pencil testing; however, this effect was moderated by computer experience, such that when computer experience was great and questionnaire completion was via computer, individuals had higher social desirability scores than when computer experience was low and survey completion was via computer (Finegan & Allen, 1994). In addition, in a meta-analysis of 61 studies comparing computerized, paper-and-pencil, and face-to-face interviews, Richman et al. (1999) demonstrated that whereas computerized testing versus face-to-face interviews showed lower social desirability effects, several moderating variables were also important, including study design, percentage of women in the sample, mean age of the sample, and type of subject population. Overall, results suggest that social desirability is a complex phenomenon that does not seem to differ simply as a main effect of data collection mode.

An important factor for social desirability is participants’ desire for impression management, as demonstrated in several studies. For example, Wilkerson, Nagao, and Martin (2002) found that the mode of administration (computerized versus paper-and-pencil) for a specific questionnaire did not matter, but that the purpose or context of the questionnaire did. Specifically, a job screening context showed higher social desirability effects than did a consumer opinion context. Similarly, Epstein, Klinkenberg, Wiley, and McKinley (2001) found that those completing an Internet questionnaire, with peers nearby who could potentially see their responses on the computer screens, rated same-sex targets lower in attractiveness (compared to those completing a paper-and-pencil questionnaire with no overlooking peers). They speculated that this effect was perhaps due to fear of ridicule from peers, although they did not directly assess this claim. Furthermore, research indicates that individuals can treat computers much like they treat other individuals (see, for a review, Reeves & Nass, 1996), suggesting that impression management efforts may be at work when interacting with computers as social agents in
the administration of questionnaires. Thus, impression management is one potential reason for social desirability in online environments.

A related factor is participants’ sense of anonymity. Anonymity reduces identifiability, leading to reduced social desirability bias because respondents are less likely to be held accountable for behaviors and attitudes that may not correspond with societal norms (Sassenberg & Kreutz, 2002). In fact, participants responding to a questionnaire online have been shown to exhibit the least amount of social desirability bias (regardless of anonymity or lack thereof), whereas those responding by paper-and-pencil anonymously gave more socially desirable responses, and those responding via paper-and-pencil nonanonymously showed even more social desirability in their responses (Joinson, 1999).

In their review of relevant research, Knapp and Kirk (2003) concluded that differences in social desirability biases between data collected online and offline may stem from the fact that confidential data collection techniques (through which participants possibly can be identified) are often employed in online data collection efforts, as opposed to anonymous techniques (through which participants cannot be identified). To assess differences in social desirability, they conducted an experiment in which data collection mode (paper-and-pencil mail questionnaire, online questionnaire, and automated touch-tone phone questionnaire) was randomly assigned to participants who answered the same set of questions ranging in topic sensitivity. In all conditions, a private setting for questionnaire completion was supplied and anonymous response methods were used. They found no significant differences in the social desirability of responses among the three groups, suggesting that anonymity (versus confidentiality) is a crucial factor mitigating social desirability response biases. Based on these findings, although anonymity is a key factor, there is also evidence that data collection mode is important in explaining social desirability response effects, with advantages to electronic means under certain conditions, as compared to other modes (e.g., paper-and-pencil questionnaire conducted in person rather than via the mail). Moreover, some recent research suggests that individuals are better able and more willing to express their “true” selves freely in a computer-mediated context, thus electronic data collection may be superior at reducing socially desirable responses and behaviors (Bargh, McKenna, & Fitzsimmons, 2002).

Certain factors affect the existence of socially desirable responses, but they are not inherent to online or computerized data collection methods. Rather, these factors may be related to these data collection modes (e.g., computer-mediated communication tends to offer reduced social cues that can increase anonymity). Future research, therefore, should focus on determining which features of electronic communication and information technologies influence social desirability effects most markedly and the circumstances under which these factors are most likely to occur.

Self-selection bias. Participant self-selection can bias research results because individuals who choose to participate in a particular research project can be different than those who choose not to participate. In this manner, those who self-select to participate may have distinctive characteristics that influence (or bias) results. Depending on the sampling method invoked, this issue may be exaggerated in online studies because online research participants may have to seek out research opportunities more actively than with other forms of data collection (e.g., random digit dialing or student participation pools; Mustanski, 2001; Sills & Song, 2002). For instance, in a sequential sample collected online (whereby every 1000th person who visited MSNBC.com was invited to participate), less online sexual activity and online compulsive behavior were found than in a convenience sample of participants who volunteered for the same topical questionnaire through news and word of mouth reports (Cooper, Scherer, & Mathy, 2001). This finding suggests that individuals are more likely to volunteer for an online study that addresses a topic that is salient to them. Thus, as mentioned earlier, when an issue is important to participants, data integrity may be jeopardized due to self-selection biases.

Reips (2000), nonetheless, has argued that for online experiments, self-selection can be better than less voluntary participation (e.g., college students participating for mandatory course credit). Online participants, he reasoned, are more motivated to participate and provide more thoughtful answers because they can terminate their participation more freely and do not feel forced to take part in research projects. Reips (2000, 2002a) also suggested ways to test if self-selection is a problem for Web-based studies. For example, researchers can use various entry (or “splash”) webpages that are advertised in different online locations, all of which lead to the actual study webpage. Reips asserts that results from the different groups of participants can be compared and, if no differences emerge, then self-selection is not an issue. This strategy, however, does not account for the possibility that participants may still differ from nonparticipants, even among those with access to the internet. Importantly, however, self-selection creates greater biases for survey research than for experimental designs, since the random assignment of subjects to conditions helps reduce the effects of self-selection because individuals with varying degrees of motivation to participate become equally distributed across experimental conditions (Cook & Campbell, 1979; Piper, 1998). Nonetheless, high motivation among participants overall may influence results regardless of the methods employed.

Participant dropout or mortality. The third major threat to the validity of online research results is participant drop-out or mortality, which occurs when participants fail to complete their participation in a study, either within one survey or over time in longitudinal panel designs (Knapp & Heidingsfelder, 2001; Piper, 1998). Participant dropout is of particular concern when the premature departure of participants is nonrandom, indicating that some characteristic(s) of those who drop out is meaningful to both their decision to terminate their participation and to relevant independent or dependent variables (Cook & Campbell, 1979). The effect of nonrandom departure is that participants exhibiting certain characteristics
are underrepresented in the data collection effort and, therefore, results cannot account for the influence of that characteristic or are influenced disproportionately by participants who lack that characteristic.

Causes of participant dropout take two general forms. First, dropouts can be technically induced, such that electronic communication and information technologies somehow prevent certain participants from continuing with a study (due, for example, to Web browser incompatibilities, slow loading pages or other bandwidth problems; Knapp & Heidingsfelder, 2001; Reips, 2000). Second, dropouts can be due to participant-based reasons, such as lack of incentives to continue with a long study, cognitive fatigue, the existence of sensitive topics, a lack of obligation to continue, and other reasons (Frick et al., 2001; Knapp & Heidingsfelder, 2001). For example, Frick et al. (2001) found that although a monetary incentive did not increase initial willingness to participate, it did decrease participants’ dropout rates.

Random dropout is typically not a cause for major concern. Dropout is, nonetheless, undesirable and several solutions for reducing dropout rates have been suggested. Reips (2002a) offered three remedies: (a) Researchers can give potential participants information about what the study will entail so that they can accurately assess their willingness to participate before starting; (b) researchers can ask participants to estimate the probability that they are likely to complete the entire study and then only examine data from highly motivated individuals; and (c) researchers can use a warm-up technique, which asks participants a set of questions that is not going to be analyzed, so that only highly motivated individuals participate in the actual study. Even though these solutions may reduce dropout rates for online studies, they also introduce motivational confounds, such that examining only highly motivated participants may induce lower participation rates or biased responses, as discussed earlier.

Additional suggestions to reduce dropout rates include requesting personal information (e.g., age, sex) prior to the study’s major tasks and measures, rather than following them, in order to guard against dropout of those who may not wish to share personal information (Frick et al., 2001). Further, participation on the weekend (rather than during the week) tended to reduce dropout rates, perhaps due to reduced time constraints, although asking for participants’ email addresses prior to a study increased dropout by 12%, perhaps due to confidentiality concerns (O’Neil & Penrod, 2001). Nonetheless, some features of electronic communication and information tools may help to alleviate dropout rates, given, for example, the reduced cost and increased ease of sending participation reminders via email and automated prompts to participants if they stop responding. In addition, although giving participants a study progress indicator has been suggested to decrease dropout rates (Vehovar et al., 2002), no effect was demonstrated in an experiment that manipulated progress indicator for an online questionnaire (Couper, Traugott, & Lamias, 2001). This particular finding perhaps resulted because webpage download time was greater for the progress indicator condition, which lasted 3–4 minutes longer to complete than the no progress indicator condition. Additional research, without feature-based confounds, is necessary to determine reasons and solutions for participant mortality.

Additional threats to validity. Demand characteristics, such as gender-based expectations revealed by a researcher’s sex, are also potential threats to data validity. Scholars have proposed, however, that demand characteristics cause less concern for research using electronic technologies, especially for online research, due to a decrease in the amount of researcher-participant interaction stemming from the ability to automate procedures (Hewson, Laurent, & Vogel, 1996; Piper, 1998; Reips, 2000; Stanton & Rodelberg, 2002). Furthermore, in online research a reduction in evaluation apprehension may result, again possibly due to less researcher-participant interaction (Piper, 1998). Nevertheless, participants at times interact with computers as if they were human (see Reeves & Nass, 1996), which could result in the preservation of these problems. No known data, however, speak to any of these claims.

Maintaining appropriate control over technical factors and participants can be problematic in online studies as well. For example, computer software and hardware specifications can vary from participant to participant (Couper, 2000; Krantz, 2001; Mustanski, 2001; Schmidt, 1997), which might result in unintended variation in experimental stimuli or questionnaire presentation. In addition, contamination effects for online experimental research can create a loss of control because participants can more readily interact with each other and divulge information about the study and its conditions (Buchanan, 2000; Piper, 1998), especially if participants are recruited via newsgroup postings or other means where they may already be in close contact with one another.

Problems of control can be addressed by several means. First, random assignment for experimental research guards against participant control issues to some degree. Moreover, technical specification inconsistencies can be offset by measuring participants’ computer configurations, either unobtrusively through computer scripts or directly by asking participants, and then statistically controlling for these variations. Also, technical inconsistencies can be minimized by using client-side (as opposed to server-side) programs that are downloaded by participants and that enable data to be captured and sent to researchers via participants’ computers in fixed ways that are not reliant on participants’ browsers, bandwidth, or computer configurations (see Reips, 2002a, for a discussion of these technical issues). However, client-side solutions may be inhibited by users’ willingness to download software to their personal computers, in view of computer virus and spyware risks.

A final threat to the validity of results is that individuals can participate in the same study multiple times, which is more likely to occur in online data collection efforts (compared to other means) because of the automated nature of much online research (Buchanan, 2000; Mustanski, 2001; Reips, 2000). Duplicate participation can occur intentionally (from participant fraud due to an incentive)
same results for each version of the scale. A preferred method is to use construct validation techniques, which rely on theoretical predictions and empirically confirmed relationships among constructs to determine a scale's integrity (Cron & Brewer, 2002; Krantz & Dela, 2000).

Online measurement reliability and validity concerns are reviewed above, focusing primarily on personality or individual difference inventories from a psychological perspective. No known research has tested individual difference measures for communication constructs or online contexts. Furthermore, the portability of the instrument will require confirmation of their online reliability and validity via comparative and convergent techniques.

The concern for data integrity centers on the quality or length of participation. A review of past studies concluded that more reliable responses for closed-ended items provided open-ended questions in email studies. However, tests of open-ended questions in email studies were not considered in the study. For example, in an experiment in which individuals were randomly assigned to email or mail data collection modes, email participants yielded a non-randomly assigned group of 52%, whereas mail participants yielded a non-randomly assigned group of 36%. The advantage of email data capture is that it is more efficient and convenient for the researcher, whereas mail data capture is more convenient for the participant. Mail participants tend to be more likely to complete the survey, and mail data capture is more likely to result in a higher response rate. However, when data is collected through email, it is necessary to verify that the data is not manipulated by the participant. In one study, participants were asked to enter a code into a survey, which served to verify that the data was not manipulated. This code was then used to determine if the participant had completed the survey. Overall, email data capture is more efficient and convenient for the researcher, whereas mail data capture is more convenient for the participant. It is necessary to verify that the data is not manipulated by the participant.
further research is required to address the causes of, and potential solutions for, item nonresponse in electronic research.

Interdependence of Observations

Data interdependence occurs when observations (or measures) of units of analysis are linked in some way, making them similar to each other compared to nonlinked units (Anderson & Ager, 1978; Kenny, 1995; Sadler & Judd, 2001). Interdependence problems arise when data from linked units are analyzed as if they are not linked, such as when the interdependence among members of a group within an organization who interact regularly are treated the same as members of the same organization who do not belong to the group. The crucial issue is that interdependence be recognized and accounted for, preferably by avoiding mismatches between levels of theory, measurement, and statistical analysis (see Glick & Roberts, 1984; Klein, Danserou, & Hall, 1994).

Many frequently used statistical tests in social scientific research, such as analysis of variance, are based on the assumption that observations of the units of analysis are independent from one another. Thus, data interdependence violates this assumption and can potentially decrease the integrity of data and associated results unless the problem is appropriately addressed. Importantly, data interdependence is based on two factors: (a) knowledge that the units are linked in some way, and (b) similarities or differences between linked and unlinked units, which require empirical validation. Linked units, therefore, do not automatically create a data interdependence problem because tests can determine if the variation among linked units from the same group is greater or smaller than the variation among unlinked units from different groups. Also, interdependence problems can be controlled through statistical procedures. Examples when interdependence could potentially be problematic include dyadic or group interactions, assessments of communication within couples or families, and other situations in which some participants share commonalities and others do not.

The problem of interdependence of observations is not inherent in any particular mode of research. For example, face-to-face dyadic interactions are just as susceptible to data interdependence issues as are computer-mediated dyadic interactions. Potential problems that are unique to online research are twofold. First, online participants may be linked in ways unknown to researchers. For example, data acquired via participant solicitations in online newsgroups may be at risk for data interdependence due to the connections individuals may have with other individuals via the newsgroup. This situation may not be easily assessed by researchers.

The second potential problem involves data collection of nonhuman units of analysis (e.g., online messages, online advertisements, online news articles, and Web-server activity logs). In this case, the concern is that when sampling these sorts of data they may be differentially linked. For example, messages posted to a particular newsgroup may exhibit different levels of dependence on phenomena within that newsgroup than would messages sampled from a different newsgroup. In this manner, data may be treated as independent observations when in fact they are influenced in meaningful ways by membership in specific groups. In such situations, interdependence of data usually can be tested and statistically controlled, if found. Given these concerns, researchers should be aware of potential interdependence problems in research examining phenomena using electronic communication and information technologies and should take appropriate steps to guard against their deleterious effects.

POTENTIAL BENEFITS OF ELECTRONIC COMMUNICATION AND INFORMATION TECHNOLOGIES AS RESEARCH TOOLS

Appropriate sampling techniques, combined with processes that ensure data integrity, are the bases of all sound research. With this foundation in place, electronic communication and information technologies can provide significant advantages for researchers by increasing the ease with which research tasks are performed and through the addition of novel capabilities offered by these tools.

Mundane and Tedious Tasks Made Easier

The research process can be tedious, complex, and slow. To some degree, the use of electronic communication and information technologies can serve to eliminate or minimize monotonous procedures, simplify intricate elements of the research process, and speed up the entire research cycle. For instance, on-line data collection can remove the need for data entry and reduce or eliminate data entry errors because data can be captured electronically in fixed formats (Christians & Chen, 2004; Ilieva et al., 2002; Schmidt, 1997). In addition, online signup for laboratory studies via a website, with automated email reminders, can simplify scheduling and alleviate the need for calling each participant individually (Reips, 2000, 2002a; Stanton & Rogelberg, 2002). Web-mediated focus groups may be more readily available and utilized compared to face-to-face focus groups; further, focus groups of special or hard to acquire populations might be easier to gather online (for summaries of several studies using online focus groups, see Williams & Robson, 2004). Studies themselves also can be largely automated, eliminating the need for individuals to oversee every part of a study. Finally, random assignment of participants to experimental conditions can be automated without participants' awareness (Piper, 1998; see Burton & Walther, 2001 for a Web-based random assignment script) and may be easier and less prone to human error than randomized condition schedules or other procedures, such as flipping a coin.

For analyses of online social interactions (e.g., language use or structure), researchers do not need to spend time on transcription, as they would for face-to-face interactions, because interaction transcripts can be generated automatically with perfect accuracy (Christians & Chen, 2004; Stanton & Rogelberg, 2002; see
Electronic Technologies as Research Tools

level of interactivity can be extended to online interactive games that can be used to assess different communicative phenomena (Ruppertsberg, Givaty, Van Veen, & Buhlhoff, 2001). For example, examinations of group dynamics and leadership emergence could be studied via group-based electronic games or scenarios that are conducted in MUDs. Scholars using diary or log methodologies (e.g., media consumption research) could utilize electronic technologies, such as portable digital assistants (PDAs), to aid participants’ entries, with the novel capability to program PDAs to remind participants to input their entries into the PDA-based digital diary/log on a set schedule. Email or Web-based programs could provide a similar function (see, for example, Haridakis, Rubin, & Rubin, 2003; Rubin, Haridakis, Rubin, & Miraldi, 2002).

Major media research firms are capitalizing on advances in electronic communication and information technologies as well. Nielsen, for example, has begun to implement nationwide unobtrusive television rating using methods, called the Local People Meter (LPM), with New York as one of the most recent markets (Nielsen Media Research, 2004). The LPM is more accurate than traditional paper-and-pencil diary methods and capable of obtaining data quicker from larger and more diverse samples. Relatedly, Nielsen-NetRatings (2004) uses computer monitoring devices to record participants’ Web usage and links these data to participants’ consumer purchasing behaviors, collected via in-home universal product code (UPC) scanners. Finally, HarrisInteractive.com (2004) uses a hybrid telephone-email method, whereby potential participants are first called via telephone and asked for their email addresses. Once they supply their email addresses, participants are emailed to participate in the current study and future research.

Advances for social interaction analyses. Observational research of social interactions also can profit from the use of electronic communication and information technologies. Researchers can use Web-based cameras (i.e., web-cams that capture audio and video and transmit these data across the internet) to record individuals in their natural environments without researchers having to visit and set up equipment in many locations (Stanton & Rogelberg, 2002). Interaction analyses of individuals in homes, organizations, and other contexts can be conducted more easily and less obtrusively via these web-cams and other electronic technologies. “Active badges” or “smart cards” can potentially track the whereabouts of participants (Stanton & Rogelberg, 2002), enabling researchers to analyze communication patterns (although when using this methodology the implicit assumption that communication would be operationalized as colocation must be acknowledged). Based on these data, social network analyses (see Monge & Contractor, 2003; Wasserman & Faust, 1994) could be used to explore the nature and form of connection among actors and behaviors (see Garton, Haythornthwaite, & Wellman, 1999). Similarly, mobile telephones, PDAs, and global positioning system (GPS) devices could be used in this way as well.

Computer monitoring devices (Stanton & Rogelberg, 2002) enable individuals’ computer usage, online navigation of websites, email messages sent and received, and other forms of communication and interaction to be tracked in a

for examples, Rice, 1982; Lemus et al., 2004; Palomares, 2004). Relatedly, although currently not sufficiently accurate, voice recognition software may soon be able to transcribe verbal communications with high fidelity. Finally, electronic communication and information tools enable access to participants from a variety of locations without researchers or participants going anywhere, reducing both time and financial costs.

Electronic technologies also can facilitate mundane and tedious aspects of some content analyses (Krippendorff, 2004). First, Web pages provide an immense data resource for scholars to examine. For example, weblogs (also known as “blogs”) are online journals that can be the subject of content analyses (e.g., Trammell & Gasser, 2004). Researchers analyzing Web-based text also are not required to transcribe content, as text is already in electronic format. Having content in electronic format is especially useful when researchers conduct computer-aided content analyses (for a list and description of computer programs used for content analysis see Neuendorf, 2002). These content analysis programs can increase reliability, reducing errors in data. Computers also can help to facilitate the acquisition of representative samples of specified content, notwithstanding the concerns discussed in the previous section on sampling.

Novel Capabilities

New technologies not only offer assistance for typical research tasks, but also allow innovative and distinctive methods of inquiry. More specifically, electronic communication and information technologies enable the presentation of stimuli, the capture of participants’ responses, and the performance of content analyses to take on novel forms.

Advances for surveys and experiments. Interviews and surveys can be automated via computerized interview robots (or “bots”) that ask individuals questions in seemingly natural conversation (Janetzo, 2002). This tool is extremely useful in reducing the costs of large-scale interviews or surveys and also may be useful for experiments for which control and continuity are high priorities. Further, surveys can be adapted or tailored to participants’ previous responses (Epstein & Klinkenberg, 2001; Hertel et al., 2002) offering, for example, follow-up questions based on particular responses to previous questions. In this fashion, interactive exercises with feedback given to participants are more readily available in online contexts (Hertel et al., 2002). Communication educators interested in pedagogical research may find instructor and course assessments less effortful in online contexts with the ability to query students easily and frequently throughout the duration of a course (e.g., Flanagan, 1999).

Moreover, electronically delivered questionnaires can include much more complex structures and are not limited to text (Best & Krueger, 2002). Experiments with photographic or video stimuli (with varying manipulations or conditions) can be more easily presented via electronic technologies (see, for example, Potter, Mahood, & Yao, 2003; Walther, Slovacek, & Tidwell, 2001). This
variety of settings (e.g., in participants’ homes, organizations, and schools). In addition, a range of variables, such as time spent interacting online or on computer-based work-related activities, are made available through the use of electronic communication and information technologies (Burton & Walther, 2001). Data acquired via mobile telephone technologies similarly could be examined in relation to other variables. These behavioral measures allow for largely unobtrusive observation of participants and could be tested for agreement with attitudinal measures (see, for example, Flanagan & Metzger, 2003b). Researchers also can use these measures as dependent variables for experiments in which independent variables are manipulated and their subsequent effects on these electronic observational measures are determined. These types of data risk a mismatch between conceptual and operational definitions; however, triangulation with other measures, such as self-reports or third-party observations, may provide more robust data. In fact, recent research has indicated that a computer-monitoring program yielded lower levels of internet usage than self-reported diary and survey estimates (Greenberg et al., 2004).

Finally, virtual reality allows researchers to analyze social interactions and observe communicative behaviors in scenarios that take place in simulated, and potentially highly controlled, environments (Stanton & Rogelberg, 2002). Recent research suggests that computers are viewed as more influential than human partners under some conditions and that participant reaction may vary by the degree of anthropomorphism of the virtual partner (Burgoon et al., 2000). In this fashion, virtual reality enables observations in seemingly natural contexts and facilitates observation of situations otherwise difficult to scrutinize or manipulate. Research indicates that under certain conditions individuals interact with media as if they were human (Reeves & Nass, 1996); as a result, these tools can be accurate reflections of how individuals would interact with real people in real situations.

**Advances for content analysis.** Content analyses also benefit by using electronic technologies. For example, Krippendorff (2004) discussed four types of computer aids for content analyses (i.e., accounts of character strings, text searches, computational content analyses, and interactive-hermeneutic approaches). These four types of computer-assisted content analysis present significant advances on which scholars can capitalize in their research. Content analyses, for example, can be conducted using semantic network analyses to examine the linkages and/or co-occurrences between certain types of content (Rice & Love, 1987). Further, metemetic approaches to content analyses—analyses of how content or texts change and evolve over time—are easier with computerized aids (see, for example, techniques such as centering resonance analysis, Corman, Kuhn, McPhee, & Dooley, 2002). Computer programs also improve content analyses by providing researchers with the ability to manage, organize, manipulate, and examine extremely large amounts of content, providing scholars new perspectives they could not obtain without such technologies.

**Even though the use of electronic communication and information technologies offers substantial opportunities for researchers, it also raises several ethical concerns, which we address next.**

**ETHICAL CONCERNS**

Ethical issues are an important consideration in the conduct of research. The use of electronic communication and information tools prompts particular ethical concerns within the traditional topics of how to obtain informed consent from research subjects, minimize the possibility of participant harm, guard subjects’ privacy, and debrief research participants.

**Informed Consent**

A major ethical concern when conducting research online is how individuals provide informed consent to participate in research studies (Sharf, 1999). Typically, a physical signature is required to signal informed consent, but this is less feasible when researchers and participants do not meet face-to-face. Methods for acquiring informed consent via the internet include using a telnet application, whereby individuals logon to a secure server and indicate their consent to participate (Smith & Leigh, 1997), or issuing instructions on a webpage that inform participants that they give their informed consent by continuing with the research when they click a hyperlink that takes them to the study’s subsequent webpage(s) (Flanagan & Metzger, 2003b; Schmidt, 1997). These methods, however, do not guarantee that participants are at least 18 years of age (unless the frame population ensures this) and are of sound mind, typically requirements for participation in the majority of research studies. In addition, the multilingual nature of the internet can result in potential misunderstandings from participants of different sociolinguistic cultures who may not fully understand the informed consent process. Gaining informed consent from individuals in online groups with fluctuating membership can be difficult, particularly for longitudinal research (Kleinman, 2004). Finally, potential participants can also falsify information (e.g., age) in order to participate in research—particularly if incentives are involved—which can be difficult to detect in online research.

Nonetheless, as Walther (2002) contended, when there is no harm to participants, traditional research conducted face-to-face does not require informed consent, suggesting that online research need not always require informed consent either. That is, regardless of the mode of data collection, if research does not contain significant risks, it should not require participants’ informed consent. Walther also argued that the problems of obtaining proper informed consent (e.g., age falsification and participants’ lack of competence) are not inherent to online research. Assessing a person’s age by inspection, for example, can be difficult in face-to-face
contexts, and assessing a person’s age vocally via the telephone is also problematic. Furthermore, assessing competency or judging a person’s understanding of informed consent material is not inherently better via face-to-face situations either, as it assumes that potential participants relay relevant signals via verbal and nonverbal cues and researchers are able to detect these cues accurately. Walther concludes that the federally implemented ethical guidelines (by which most U.S. research institutions and universities abide) privilege face-to-face research in terms of the informed consent process. Rather than deeming online research a special circumstance where informed consent must always occur, the informed consent process should depend on the specific research regardless of data collection mode. In other words, one should not claim ipso facto that all online research is problematic with regard to informed consent.

Participant Harm

Another issue is the potential for participant harm when using electronic communication and information technologies as research tools. The ease and speed with which participants can obtain feedback in a study, coupled with the lack of personal debriefing, is one concern. Online research makes it quite simple to provide participants specific results for a study (indeed, this feature may be used as a recruitment technique, as previously discussed). Some scholars warn, however, that information should not be given to participants via the internet if harm may result from such knowledge (Buchanan, 2001, 2002; Kleinman, 2002). For example, in a study examining intelligence or depression levels, the revelation of low intelligence scores or high depression scores can cause psychological harm to participants. Of course, the appropriateness of providing specific feedback is not unique to online research methods, but the potential for participant harm can be magnified given the ease of feedback and the need to supply a qualified individual who is capable of interpreting results and providing necessary counseling when harmful or threatening information is made available.

Another concern regarding participant risk is potential negative effects of interaction with nonhuman agents. Individuals sometimes interact with, and react to, computers as if they had human characteristics (Reeves & Nass, 1996); as a result, researchers should be cautious about the “personalities” imparted to electronic technologies used for research, since participants could be affected by computers’ personalities in a negative way. Relatedly, researchers’ use of virtual reality and other computer-simulated interactions may also prompt negative events and behaviors, even if participants are told that the simulated “social” interactions are not real. Again, although the interactions themselves may not take place with humans, the experiences of these interactions (e.g., negative responses or unwanted arousals) may affect participants as if the social interaction were real. Harm to participants is a serious concern to which researchers should pay appropriate attention.

Privacy

Electronic communication and information technologies can create new privacy issues for researchers. Some scholars contend that despite the internet’s public nature, public data obtained online differ, for example, from observing individuals on public streets, because the internet is often part of people’s homes (Elgesem, 2002). Moreover, some note that privacy may be invaded through intrusive email messages sent to potential research participants (Cho & LaRose, 1999). Further, many electronic research devices noted earlier (e.g., active badges, computer monitoring devices, and web-cams) warrant special consideration with regard to potential privacy violations because of their unobtrusive and ubiquitous capabilities (Stanton & Rogelberg, 2002). These technologies enable less obtrusive observation of individuals in many contexts; as a result, some scholars feel that participants may forget about these monitoring devices and reveal information or behave in ways they would otherwise wish to remain private (Stanton & Rogelberg, 2002). There also is the potential with these devices (e.g., web-cams) to collect data from individuals who have not provided their consent to participate in research but are in close proximity to individuals who have provided their consent. Assuming, however, that participants consent to such unobtrusive observations and have the option to withdraw or remove certain information from the researcher’s database, ethical violations regarding privacy are not likely. Without participants’ informed consent, of course, the use of these observational techniques would most likely result in privacy violations if the situations are not public or the risks to participants are great.

Use of online public data not acquired through participant-researcher interaction (such as content analyses of online message postings, online advertisements, websites, and other types of nonhuman participant data) is somewhat controversial with regard to privacy concerns. Messages posted online are usually submitted to newsgroups or other online communities that are available to any member of the public with internet access. Some scholars feel that use of these sorts of data can be unethical if participants’ consent is not formally given because individuals do not intend for their messages to be used for research and have a right to be able to keep certain information private (Bellotti, 1997; Kleinman, 2002, 2004; Sharf, 1999). On the other hand, other scholars feel that information posted on the internet, including online postings in newsgroups, online communities, and other online discussions, is in the public domain and fair use regulations regarding copyright dictate that use of these data for research purposes should not be regulated by guidelines that govern the use of human participants (Bruckman, 2002; Walther, 2002). Use of online communications also does not require researcher-participant interaction (e.g., participants are not surveyed for the information), and thus using these data should be exempt from the regulations regarding human research participation. Bruckman even suggested that online messages should be linked to particular individuals, because information posted on the internet
is semipublished information and, if no risk is involved, then authors should be given credit for their postings.

Walther (2002), however, argued that messages posted online need not be linked to individuals at all because most social scientific research would merely quantify online messages and report general trends. On the other hand, much qualitative research, such as conversation analysis, would use specific exemplars that, even if pseudonyms are employed, could have the potential to be traced to the corresponding authors through the archival and search capabilities on the internet. Researchers using specific online messages as examples should therefore be aware of the potential to link these messages to participants, especially if links could create a risk for the authors of messages. Observations of online messages that cannot be linked to individuals, and that do not use identifiable examples, on the other hand, should not be considered human participant research.9

A final privacy concern in online research is maintaining participant confidentiality. Given the public nature of the internet, keeping information confidential, secure, and private is important when conducting online research. Buchanan (2002) suggested that information transferred between participants’ computers and researchers’ Web servers should be encrypted to maintain confidentiality, especially when participants’ responses are associated with identification information. Participants in online research also should be informed that email messages they send are usually saved on their computers and that Web browsers typically cache webpages they have visited, so that if participants share their computers with others or if they use public computers, their responses could be read by others (Walther, 2002; see also Sharf, 1999). An additional recommendation is that researchers should opt for Web-based research over email-based studies because email-based studies contain participants’ email addresses, which could decrease the likelihood of participant confidentiality (Cho & LaRose, 1999; Im & Chee, 2002) and may spawn email viruses. Another method to keep participants’ responses confidential is to separate informed consent webpages, which require participants’ identification information, from the webpages containing participants’ responses, by not linking them in databases (Cho & LaRose, 1999). Additionally, organizations participating in research efforts can randomize user identifications and link them to survey data, so that researchers are unaware of individuals’ identities (see, for example, Rice, Hughes, & Love, 1989). A final recommendation to ensure participants’ confidentiality is to have the researcher’s websites undergo third-party certification, such as that provided by TRUSTe.com. These methods are important to ensure the confidentiality of participants’ responses in online research.

Debriefing

A final ethical concern is debriefing participants upon the completion of a research study. Debriefing is particularly important when deception is used, in order for researchers to reveal the true purpose of the study to participants. Debriefing also can inform participants of hypothesized relationships, experimental manipulations, and data regarding participants’ individual performance measures. Typically, debriefing does not pose major problems. However, online research participants can leave studies more freely before debriefing takes place (Piper, 1998). Debriefing online typically takes the form of a text-based statement via either a webpage or email; participants (intentionally or unintentionally), however, can easily ignore these statements online. Suggestions for increasing the likelihood that debriefing occurs successfully in Web-based research (for participants who complete the study and for those who drop out early) include emailing a debriefing statement to participants after they leave the study, making debriefing information available via a hyperlink from a button saying, “click here to leave the study,” and automatically providing debriefing information to participants once they leave the study at any stage through a computer script (Nosek, Banaji, & Greenwald, 2002). Of course, ensuring that participants are given and actually read the debriefing information is not a problem limited to online research (Walther, 2002); participants, for example, can choose to ignore debriefing information in online research just as easily as in face-to-face research.

Even though the use of electronic technologies in the conduct of research presents particular ethical considerations, scholars should treat all forms of research equally with regard to ethical guidelines, whether data are collected face-to-face or electronically. Thus, above all else, researchers should strive to do no harm regardless of the tools involved. “No harm” research is particularly vital as future electronic tools reveal unforeseen ethical complexities.

THE FUTURE OF COMMUNICATION RESEARCH AIDED BY ELECTRONIC COMMUNICATION AND INFORMATION TECHNOLOGIES

As scholars continue to apply electronic communication and information technologies to research endeavors, methodological issues covered in this essay remain important and additional concerns become relevant. Indeed, several of these concerns are critical to consider for the future of communication research, including (a) the evolving role of advanced electronic technologies in research activities; (b) an informed awareness of the methodological issues associated with the use of electronic technologies when conducting research; (c) communication-specific issues implied by the use of electronic tools in research; and (d) future applications of electronic technologies, based on current uses and perspectives.

The Evolving Role of Technology in Research Endeavors

To date, electronic communication and information technologies have influenced scholarly research in at least three ways, each of which has implications for future applications of these tools. First, researchers have transported traditional methods...
and procedures into electronic arenas. The use of technologies in this manner is readily apparent and typically straightforward. Relevant examples include Web-based experiments, online focus groups, email questionnaires, and computerized assessments of face-to-face behaviors. In these instances, traditional methods are transferred to an electronic format, with minimal methodological modifications and typically without embellishment. In many cases, the conversion of methods to electronic means can be achieved without substantial pitfalls, although appropriate methodological vigilance is, of course, critical.

Next, electronic technologies can be used to alter, augment, and expand traditional methods. Researchers can use electronic technologies to generate new or more complex means to assess research questions in novel ways. In such cases, technological advances have the potential for more accurate research methods (e.g., computer-aided content analyses, Web-based questionnaire administration, or online network data collection), new ways to assess phenomena (e.g., adapted or tailored experimental manipulations or computer-captured Web log data), and the exploration of new phenomena that may result (e.g., the use of active badges might prompt the assessment of novel network phenomena). Importantly, such extensions are not new: For example, conversation analytic methodologies did not gain momentum until the advent of the tape recorder, which allowed researchers to capture conversation and make accurate transcripts that served as the basis for systematic analysis (cf. Psathas, 1995). Electronic technologies are similarly capable of ushering in domains of research that have not yet been fully explored because limitations of time, accuracy, expense, and effort.

Third, electronic tools have enabled scholars to direct their research efforts toward understanding electronic technologies themselves. For example, researchers can examine new technology usage via inspection of Web log files and cellular telephone usage. Similarly, research efforts can focus on investigating the methodological issues posed herein, such as the causes of online participant dropout, the convergent validity of a common paper-and-pencil measure administered online, or comparisons of self-report to system-collected usage data.

**Informed Awareness**

A vital mechanism to maintain the rigor and quality of research using electronic tools is an informed awareness and thorough consideration of the issues outlined in this essay. Anyone conducting electronic-based research must weigh the costs and benefits in order to make informed and thoughtful decisions. Those reading, critiquing, and/or reviewing such research also must pay attention to the issues we have discussed because these issues can influence their appraisal of research. Being cognizant of these issues will help to ensure that future research is sufficiently rigorous to maintain trust in research findings and knowledge claims. Indeed, among our goals in this essay is to prompt discussions among scholars interested in electronic technologies about the conceptual, methodological, and empirical issues surrounding the use of these tools in research endeavors.

**Communication-Specific Issues Implied by the Use of Electronic Research Tools**

The methodological issues presented in this review are many respects particularly relevant for communication scholars, given that interindividual interactions are at the core of what communication scientists examine (Chaffee & Berger, 1987). Indeed, many of the methodological issues addressed in this review are fundamentally communicative in nature. For example, reactions to different modes of questionnaire delivery, interaction among individuals online versus in person, and problems with social desirability biases are intrinsically issues of communication delivery, style, intent, and mode. Thus, communication scholars are well equipped to play a major role in assessing the methodological dilemmas we presented. As a consequence, communication scientists should find in this review ideas on research procedures as well as research opportunities.

**Future Applications of Electronic Tools in Research**

The future presents many possibilities for the use of electronic communication and information technologies as research tools. Many scholars likely will rely increasingly on electronic communication and information technologies in the conduct of their research. Specifically, given the benefits and novel capabilities of using electronic communication and information technologies outlined herein, studies using traditional methodologies will be transferred into electronic environments and existing concepts and measures will be transported into online contexts. Replications of previous findings also will be tested in electronic settings. In addition, novel methodologies will emerge that capitalize on unforeseen features in yet-to-be-realized technologies. Scholars will introduce new concepts and accompanying variables in order to test original ideas and phenomena generated by electronic technologies. Most importantly, however, the future holds an explicit discussion and empirical investigation of relevant issues addressed throughout this essay. Active research directly conceptualizing and testing the methodological issues associated with using electronic tools for research is necessary to inform the rigor and quality of our research endeavors.

**CONCLUSION**

Rigorous research methods are the foundation of all sound research, regardless of the specific tools invoked. The goal of this chapter has been to explore how the application of electronic tools to the research process provides new possibilities while alerting researchers to specific problems that may arise from this application. On one hand, electronic technologies enable researchers to perform tasks more efficiently while facilitating new forms of data collection and analysis. On the other hand, with these new functions come new obstacles. Only by heeding proper research methods can researchers capitalize on the promise of electronic technologies as research tools while avoiding the perils associated with their use.
Indeed, the promise of these tools is great. Among the most valuable aspects of electronic technologies for research endeavors are their application to make research tasks easier, reach a variety of populations, gather data from a diversity of sources, capture several forms of social interaction, and provide novel ways to conduct experiments, network analyses, surveys, content analyses, and other types of research. Core concerns prompted by these applications are proper sampling, including attention to participant recruitment, response rates, generalizability, and the viability of internet-based samples, and data integrity concerns such as threats to validity, measurement reliability, item response issues, and independence among observations. In addition, ethical concerns specific to the use of electronic technologies are crucial in the appropriate use of these tools.

As the use of electronic tools for various research tasks evolves, more research on the tools themselves is required to properly assess their strengths, weaknesses, and appropriateness in research pursuits. Importantly, such examinations should be cognizant of confounds between media and their features and attempt to disentangle the two (for related arguments, see Eveland, 2003; Griffith & Northcraft, 1994; Nass & Mason, 1990; Rice, 1992). For example, a study testing whether social desirability occurs more often in paper-and-pencil surveys versus online surveys should also investigate other feature-based variables, such as anonymity and confidentiality, by fully crossing all variables where possible to avoid confounds. Moreover, conceptual motivations should drive research, as opposed to variable analytic impulses. Doing so will enable researchers to draw conclusions about features and technologies that may signal important distinctions when various tools are employed in future research efforts.

Even though a great deal of research remains to be done to properly assess the utility of electronic technologies as research tools, a first step is an informed awareness of the core issues implied by the use of electronic communication and information technologies in the research process and an attendant understanding of the major methodological issues implied by their use. Our hope is that this essay serves as a preliminary step in outlining these issues, establishing the framework of a research agenda, alerting scholars to the promise and perils associated with the use of electronic technologies as research tools, and sparking an informed and productive discussion on these issues.

NOTES
1. Of course, units of analysis besides individuals, such as television programs, also can be sampled. The focus of the current section, however, is on sampling individuals because that is the primary focus of most research. A discussion of sampling other units of analysis besides individuals, nevertheless, is provided at the end of this section.
2. One potential problem generated is participant self-selection bias; however, we directly address this issue later, when discussing data integrity.
3. Some argue that a representative sample is not necessary with experimental designs, by virtue of the fact that many phenomena under scrutiny do not vary significantly across populations (see Birnbaum, 2001). Consequentially, as long as random assignment to experimental conditions is maintained, results adhere unless and until researchers wish to generalize experimental results to wider populations.
4. For additional information, including websites about content analyses, bibliographies on content analyses, message archives, a program that can calculate reliability, information on content analyses programs, and other resources, visit The Content Analysis Guidebook Online at http://academic.csuohio.edu/kneuendorf/content/.
5. Researchers should consult their specific institutional regulations before conducting this type of online research because current regulations employed by human participant committees (or internal review boards) at universities and other institutions vary. For a discussion of obtaining approval from internal review boards, see Johns, Hall, and Crowell (2004).

REFERENCES


Werner, A. (2002). Contact measurement in the WWW. In B. Batinic, U.-D. Reips, & M. Bosnjak (Eds.), Online social sciences (pp. 251-262). Toronto: Hogrefe & Huber.


Williams, M., & Robson, K. (2004). Reengineering focus group methodology for the online environment. In M. D. John, S. L. P. Chen, & C. G. Hall (Eds.), Online social research: Methods, issues, and ethics (pp. 25-45). New York: Peter Lang.