Internet Research in Psychology

Samuel D. Gosling[1] and Winter Mason[2]

[1]Department of Psychology, University of Texas, Austin, Texas 78712; email: samg@mail.utexas.edu

[2]Stevens Institute of Technology, Hoboken, New Jersey 07030; email: m@winteram.com

Annu. Rev. Psychol. 2015. 66:877–902

First published online as a Review in Advance on September 22, 2014

The *Annual Review of Psychology* is online at psych.annualreviews.org

This article's doi: 10.1146/annurev-psych-010814-015321

Keywords

Internet research, online research

Abstract

Today the Internet plays a role in the lives of nearly 40% of the world's population, and it is becoming increasingly entwined in daily life. This growing presence is transforming psychological science in terms of the topics studied and the methods used. We provide an overview of the literature, considering three broad domains of research: translational (implementing traditional methods online; e.g., surveys), phenomenological (topics spawned or mediated by the Internet; e.g., cyberbullying), and novel (new ways to study existing topics; e.g., rumors). We discuss issues (e.g., sampling, ethics) that arise when doing research online and point to emerging opportunities (e.g., smartphone sensing). Psychological research on the Internet comes with new challenges, but the opportunities far outweigh the costs. By integrating the Internet, psychological research has the ability to reach large, diverse samples and collect data on actual behaviors, which will ultimately increase the impact of psychological research on society.

Contents

INTRODUCTION

Since the dawn of the World Wide Web, behavior on the Internet has changed from a relatively solo activity for a minority of technologically savvy people in wealthy Western democratic societies to a vast informational and social web that connects approximately 3 billion people around the globe, or 40% of the world's population (Int. Telecommun. Union 2014; **Figure 1**). This growth has been accompanied by a similar increase in daily usage as the Internet has become more deeply entwined in all of our lives (Wilson et al. 2012). What this means for psychological research—from psychological phenomena resulting from Internet use to novel theoretical and methodological innovations derived from its tools—is the focus of this review.

The first signs of the Internet being used in psychological research appeared in the late 1990s (Gosling & Bonnenburg 1998, Kraut et al. 1998, Young 1998). One early study used the Internet to collect pet owners' ratings of their pets' personalities (Gosling & Bonnenburg 1998). By today's standards, the study was rudimentary, with feedback scores that were generated manually and emailed to participants individually, and the sample was clearly selective (e.g., biased toward technically savvy users and those with Internet access). However, even this early study hinted at some of the promise that the Internet held for psychological research—it collected what in those days was a large sample (i.e., $N = 1{,}640$), cost very little, obviated the need for data entry, and reached participants beyond the typical college undergraduate population.

As more researchers began to use the Internet in their studies, the benefits of doing so became increasingly apparent. These benefits include the improved efficiency and accuracy with which data can be collected, the possibility of instantly checking the validity of protocols and providing participants with immediate feedback, the ability to reach large and diverse samples from around the world, the ability to target populations of participants, and the ease with which various media (e.g., sound, animation, video) can be integrated into studies (Gosling & Johnson 2010, Reis & Gosling 2010). In addition, as the Internet became increasingly integrated into everyday social and professional activities, the distinction between online life and "real life" began to lose its

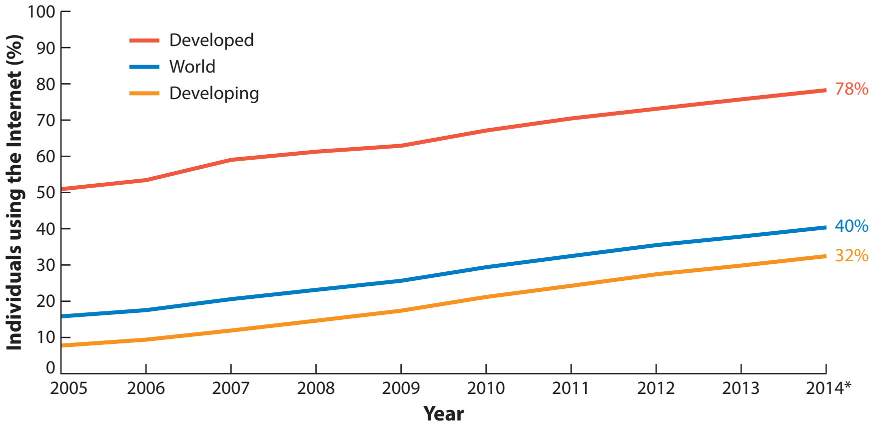

Figure 1

Growth in Internet usage from 2005 to 2014. *Note: 2014 data are estimated. Data taken from International Telecommunications Union World Telecommunications ICT Indicators Database.

usefulness. The prevalence of online interactions brought with it many opportunities to study behavior as it is played out online, narrowing the gap between the way phenomena are studied by psychologists and the way the phenomena occur in everyday life.

Naturally, many researchers were initially cautious about the new method. Much of the early skepticism was based on assumptions about who would be online and how those people would behave. The major concerns were that Internet samples were not demographically diverse and were maladjusted, socially isolated, or depressed (Kraut et al. 1998); that participants would be unmotivated; or that the data would be compromised by the anonymity of the participants and the findings would differ from those obtained with other methods (McKenna & Bargh 2000). However, when those concerns were examined empirically, they turned out to be unfounded (Gosling et al. 2004).

TYPES OF INTERNET RESEARCH

As distrust of the new form of data collection began to dissolve, Internet research began to appear in print. In their 2006 *Annual Review of Psychology* article, Skitka & Sargis identified 22 studies (out of 1,401) published in American Psychological Association journals in 2003/4 that had used the Internet. Skitka & Sargis (2006) divided Internet research into three broad categories, a framework that is still useful today. These categories are translational (implementing traditional methods on the Internet; e.g., personality surveys), phenomenological (studying a topic spawned or mediated by the Internet; e.g., Internet addiction), and novel (providing a new way to study a phenomenon already studied in psychology; e.g., music preferences). Since then, studies using Internet data have proliferated enormously, encompassing a wide range of topics and designs. In fact, studies that use the Internet in one way or another have become so pervasive that reviewing them all would be impossible. More importantly, reviewing all studies would be uninformative because the method covers virtually all areas of psychology. Therefore, we instead use this review to focus on the broad trends in the field, and in the sections below, we offer a brief sampling of research that falls into each category.

Translational Methods

The most straightforward application of Internet methods is simply using Internet technology to deliver surveys, questionnaires, and experiments to participants via the Web. Essentially, these translational methods take studies that would previously have been delivered via paper-and-pencil measures, in-person interviews, and on-site experiments and present them via the Internet. So the change is merely the form of delivery. Nonetheless, the benefits of this change can be significant. Prominent examples of translational studies are the outofservice.com website, which hosts personality questionnaires, and the Project Implicit website (**http://projectimplicit.net/**), which uses the Internet to administer the Implicit Association Test (Greenwald et al. 1998). Both of these portals have reached several million participants.

Delivering standard instruments over the Internet offers a number of advantages over traditional means of data delivery. In particular, these advantages include reducing the use of physical resources (e.g., paper), eliminating the need for data entry, and allowing researchers to take advantage of dynamic features such as automatic checks for item completion or adaptive testing such that the questions presented are contingent on the answers to previous questions. One particularly important advantage of Internet studies has been the ability to provide immediate feedback to participants. This feature proved to be a strong incentive to participants (Reips 2010), especially when using respondent-driven sampling methods (see below), and has allowed the recruitment of samples previously unmatched in size and scope.

Over the past decade or so, delivery of surveys and experiments over the Internet has gone from being quite rare to becoming commonplace. Sometimes the content is delivered no further than to the rooms in a laboratory next door, but as we shall see in the section on recruitment, in many cases data collection can reach well beyond traditional samples and contexts.

Phenomenological Methods

The Internet has transformed society and changed the way people think about and interact with the world and with each other. The Internet has created new forms of behavior (such as online gaming, virtual worlds, and crowdsourcing) and new ways for people to interact with each other (such as online forums, social networking sites, and photo sharing sites)—with both positive and negative effects. In fact, over the past 20 years many different effects of the Internet on people have been studied, including the potential for addiction (Young 1998), the inducement of isolation (Putnam 2000) or loneliness (Kraut et al. 1998), cyberbullying (Tokunaga 2010), the spread of rumors (Friggeri et al. 2014), political polarization (Garrett 2009) and filter bubbles (Pariser 2011), and the Internet's influence on political movements (Tufekci & Wilson 2012). In the present section, we review and summarize the key themes to emerge from these lines of research.

Addiction. Early studies on Internet addiction adapted the definition of dependency from the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) definition for pathological gambling (Young 1998). Other studies adapted different measures of dependency such as the DSM's definition of drug addiction (Brenner 1996, Goldberg 1996), invented novel measures (Morahan-Martin & Schumacher 2003), or combined previous definitions of Internet addiction (Chou & Hsiao 2000). Despite these different approaches, four features are common to most definitions of Internet addiction: excessive use, withdrawal, tolerance, and negative repercussions (Block 2008).

At first, studies found that the majority of Internet addicts were men (Scherer 1997), but later studies found that an equal or larger percentage of addicts were women (Leung 2004). It is unclear whether the change in demographics is due to a change in the recruitment and measurement (cf. Leung 2004), a change in demographics of Internet users (Int. Telecommun. Union 2014), or an actual change in the population of Internet addicts. Treatment for Internet addiction followed patterns similar to the treatment of other addictions, particularly drug use and pathological gambling, such as cognitive behavioral therapy (CBT; Davis 2001). Unfortunately, a relatively recent assessment concluded that Internet addiction is resistant to treatment and has high recidivism (Block 2008).

Loneliness. One of the negative repercussions identified early in psychological research on the Internet was the association between Internet use and loneliness (Kraut et al. 1998). This research was perhaps most popularly captured in the book *Bowling Alone* (Putnam 2000), which held that society in the United States had slowly been losing social capital in part because advances in technology diminished the need to interact with others and increased the ability to enjoy leisure time individually. The Internet, in this case, was seen as just one more step along this path toward isolation. A more recent popular book, *Alone Together*, by Sherry Turkle (2012) has a novel twist on this hypothesis, suggesting that in-person interactions have become lower in quality because they are framed by online interactions that are more superficial, lower risk, and easier to disconnect from. In contrast to Putnam's hypothesis, however, Turkle (2012) suggests that people now have a hard time being alone because they have excessive opportunities to connect with others online.

The theory that loneliness is a consequence of Internet use had some scientific basis (Kraut et al. 1998), but subsequent empirical studies (Amichai-Hamburger & Ben-Artzi 2003, Morahan-Martin & Schumacher 2003), including work by the original authors (Kraut et al. 2002), found evidence that increased Internet use was more a consequence than a cause of loneliness. For instance, Amichai-Hamburger & Ben-Artzi (2003) found that a model assuming loneliness led to increased Internet usage fit the data significantly better than one that assumed Internet usage caused loneliness.

Cyberbullying. A more recent phenomenon that has received attention is cyberbullying, in which traditional acts of bullying are played out in an online setting. The accepted definition of cyberbullying among researchers is "aggression that is intentionally and repeatedly carried out in an electronic context... against a person who cannot easily defend him- or herself" (Kowalski et al. 2014). Cyberbullying has a lot in common with traditional bullying, particularly aggression and repeated behaviors, but there are also some important differences. In particular, the perceived anonymity of online interactions can lead to deindividuation (Postmes & Spears 1998), encouraging individuals to behave more like they would in an angry mob than they would on their own. For example, one study found that students who would not have otherwise engaged in bullying did so because of the anonymity afforded by the online interaction (Sourander et al. 2010). The anonymity also creates a separation between the bully and the victim, which can reduce empathy and remorse in the bully (Sourander et al. 2010). In addition, the motivations for bullying are potentially changed because the intrapersonal relationship dynamic between bully and victim does not carry through to offline settings when the interaction between bully and victim is masked by anonymity (Kowalski et al. 2014).

Another important distinction between traditional bullying and cyberbullying is the pervasiveness of the behavior. Teens and adolescents spend increasing amounts of time online on an increasingly diverse number of platforms, so the channels through which a person can be targeted are also growing (Tokunaga 2010). Nevertheless, the research also indicates that a large overlap exists between the perpetrators and victims of traditional bullying and cyberbullying (Juvonen & Gross 2008, Smith et al. 2008). This finding suggests that the underlying social dynamics between traditional bullying and cyberbullying are the same, but the reach of cyberbullying is more extensive.

This prevalence has serious consequences. One meta-analysis of research suggested that the outcomes of cyberbullying include depression, low self-esteem, anxiety, stress, drug and alcohol use, loneliness, and suicidal ideation (Kowalski et al. 2014). Another literature review found sudden drops in grades, increased absences, and truancy (Tokunaga 2010). Many of these outcomes are the same as those of traditional bullying (Kowalski & Limber 2013), though the distribution of these outcomes varies. For instance, traditional bullying was more strongly associated with anxiety and health issues, whereas cyberbullying was more strongly associated with self-esteem issues, absences, and poor grades (Kowalski & Limber 2013). Research on cyberbullying, as with the phenomenon itself, is still relatively new, so there is much more to be done. The cultural and contextual factors that are associated with the predisposition toward cyberbullying, as well as ways to intervene and prevent it, are key areas for future research.

Rumor mills. The antecedents, motivations, and consequences of gossip and rumor have been studied for decades (Allport & Postman 1947, DiFonzo & Bordia 2007, Rosnow 1980). Gossip and rumor are conceptually distinct, with gossip focused on structuring and maintaining the social relationships between people, whereas rumor is primarily motivated by sense-making of uncertain information (DiFonzo & Bordia 2007). The Internet provides many channels for sharing information or misinformation, providing opportunities for rumors and gossip to diffuse in ways that are not possible offline. Thus, rumors spreading online likely behave differently than they do offline. The study of rumors on the Internet is a rich area because there is a more permanent record of how information has spread, allowing a finer-grained analysis of rumor transmission than is typically afforded with offline research (Bordia & DiFonzo 2004).

As one example, when a large earthquake hit Chile in 2010, people were thrust into a high-anxiety and high-uncertainty situation, a fertile ground for the creation and transmission of rumors. Researchers used Twitter to understand what kinds of rumors were being created and how they were being criticized (Mendoza et al. 2010). They found that false rumors were tweeted more, on average, but also received more criticisms than true information, suggesting a mechanism for defeating false rumors—although perhaps the mechanism is not fully effective (Friggeri et al. 2014).

Tanaka et al. (2012) looked at the psychological mechanisms behind the sharing of rumors and criticisms surrounding the earthquake that hit off the coast of Japan in 2011. They chose 10 tweets with rumors and 10 tweets with criticisms, and they presented the tweets either rumor-first or criticism-first to participants in Chiba and neighboring prefectures. When people saw the rumor tweets first they were more anxious and perceived the critiquing tweet to be more accurate, which suggests a possible reason why false rumors are tweeted more—they provoke anxiety and thus a desire to share the information. The criticism tweet ameliorates the anxiety, making it seem more important (although perhaps not encouraging sharing in the same way).

Filter bubbles. Another recent phenomenon on the Internet that has attracted a fair amount of media and research attention is the concept of filter bubbles (Pariser 2011), in which the glut of information provided by the Internet becomes problematic because it makes it easy to avoid information that challenges one's beliefs. Part of this phenomenon is attributable to the proliferation of recommender systems that use a person's past behavior to recommend future actions. Some popular examples include movie recommendations on Netflix, product recommendations on Amazon, or the stories that appear in the Facebook newsfeed—all of which are controlled by algorithms that use past history to determine what an individual wants to see. The danger is that these mechanisms diminish the diversity of things to which one is exposed, potentially leading to an unintentional and relatively invisible isolation from new experiences (McNee et al. 2006, Pariser 2011). The consequences are particularly pernicious when these recommendation systems are applied to news, potentially limiting the dialogue that is seen as necessary for political discourse and political change in a community—and this personalization of news content has been increasing (Thurman & Schifferes 2012).

Thus, much research in this area has focused on the role of online news and media consumption in explaining an apparent polarization in political opinions. As early as 2005, a study found that political blogs were highly polarized with respect to the links between them (Adamic & Glance 2005), and this polarization has continued into modern social media such as Twitter (Conover et al. 2011). Within the United States, there is evidence that people are selecting their news sources on the basis of anticipated agreement, potentially leading to increased polarization in opinions (Iyengar & Hahn 2009). However, there is also evidence that this ideological selection of news may be caused more by a desire for opinion-reinforcing news rather than avoidance of opinion-challenging news (Garrett 2009). Moreover, extensive research has found ways to increase the diversity of responses in recommender systems (Ge et al. 2010, Herlocker et al. 2004, McNee et al. 2006), thus avoiding or at least minimizing the filter bubble.

Social movements. In contrast to the relatively negative effects described above, one effect of the Internet has been lauded, namely the opportunity to use it as a public square or as a coordinating mechanism for social change. The most prominent example of this was the role of social media in the Egyptian revolution of 2011, a feature covered heavily in popular media (Cohen 2011, Kirk 2011, Rich 2011). Research on the Tahrir Square protests of 2011 show that social media, particularly Facebook, were instrumental in disseminating information about the protest, facilitating organization through channels not easily controlled by the existing government (Tufekci & Wilson 2012). Some argue that the use of social media was not necessary to the process (Rich 2011, York 2011), but many of the protests were primarily organized through the Internet and social media channels such as Facebook (Lim 2012). Similar uses of Facebook and other social media were identified in efforts to organize youth protests in Chile in 2010 (Valenzuela et al. 2012).

Novel Methods

One key contribution of Internet research has been the new opportunities it has brought to study phenomena that have long been of interest to psychologists. Here we review seven research topics as illustrations of the varied and powerful ways in which Internet studies can enrich psychological science. These topics represent just a small sample of the domains in which Internet studies have been used.

Social networks. As noted in the previous section, one of the key ways online behavior differs from offline behavior is the ease with which information flows from person to person. Moreover, the way in which people are connected is an important factor in how information spreads (Mason et al. 2007). The Internet, particularly with social networking sites such as Facebook and Twitter, has brought this facet to light, both in public awareness and in research (Wilson et al. 2012).

Some recent work has purportedly demonstrated that characteristics (and presumably behaviors) such as obesity (Christakis & Fowler 2007), happiness (Fowler & Christakis 2008), or loneliness (Cacioppo et al. 2009) can spread from person to person. The methodology used in these studies was shown to have flaws that undermined the claims of causality (Lyons 2011), but subsequent work has provided more conclusive evidence of a causal relationship (Coviello et al. 2014) in the spread of moods. There is a significant, albeit small, effect of weather on mood (Hannak et al. 2012); Coviello et al. (2014) were able to detect a decrease in the positive emotion words and an increase in the negative emotion words in Facebook status updates of people in cities when it was raining. Importantly, they also found a significant change in the emotion words for friends of people in the cities with rain, even when the friends' cities were not experiencing rain.

Of course, it is not just mood that passes from person to person online—so does information. There has been extensive research on the way in which different types of information spreads online (Goel et al. 2012), from URLs (Bakshy et al. 2010), to phrases (Leskovec et al. 2009), to online applications (Aral & Walker 2011). The study of the transmission of information is not new to psychology; for instance, Manis et al. (1974) looked at the transmission of attitude-relevant information, and Kashima (2000) looked at the retention of stereotypic and counterstereotypic information in chains of retold narratives. However, the extensive use of online social networking sites and the explicit communication channels in these social networking sites afford illuminating new ways to observe and measure information transmission.

One recent study examined rumors in a way and at a scale not previously possible (Friggeri et al. 2014) by leveraging a website called Snopes (**http://www.snopes.com/**), which is dedicated to documenting and fact-checking rumors and urban legends. By looking for instances where one person posts a photo, and someone else posts a link to a Snopes article in the comments on the link, the authors inferred that the original posted photo referred to a particular rumor or urban legend, and through the Snopes article they could determine the veracity of the rumor. With this method, they were able to identify over 4,000 distinct rumors, which were collectively shared over 62 million times. An example is shown in **Figure 2**; this represents all of the shares of a picture of a receipt from Cabela's that shows a "medical excise tax" on a firearm (**http://www.snopes.com/politics/taxes/medicaldevice.asp**). The rumors could spread quite far, even when many posts of the rumor were commented on with a Snopes link, suggesting that the rumor spread more quickly than the criticism. This is true even though links that received a comment with a link to a Snopes page were more likely to be deleted. In fact, the majority of reshares of a rumor occurred *after* the original post was commented on with a Snopes link, suggesting that the effectiveness of a criticism on the source—if the original poster does not decide to delete it—is limited.

Keeping secrets. Secrets are a topic of great psychological interest and clinical significance (Lane & Wegner 1995, Pennebaker & Sussman 1988) but are hard to study because they tend to focus on things that people do not want to reveal to others. Another obstacle to studying secrets is that there are typically several parties involved, consisting of the secret keeper, the target (from whom the secret is being kept), the confidant, and uninvolved others, and it is difficult to study their behavior without alerting them to the existence of the secret. As a result, naturalistic studies of all the parties involved in a secret in a real-world setting have been virtually impossible to undertake, with most past research relying on lab experiments, surveys, and interviews. However, the advent of pervasive email communication has provided new opportunities for examining the complex interactions associated with keeping secrets.

Tausczik et al. (2014) used email correspondences furnished by the holders of significant real-world secrets. To find out what happens to social relationships at the onset of secrets, the researchers collected email correspondences between the secret keeper and the target, the confidant, and uninvolved others. Tausczik et al. (2014) were particularly interested in testing whether secrets lead to withdrawal between the secret keeper and the target (because this could be a strategy for maintaining the secret) or whether participants become hypervigilant as they step up the degree to which they monitor their social relationships. The results generally supported the hypervigilance hypothesis. One particularly interesting finding was that the target of the secret increased the

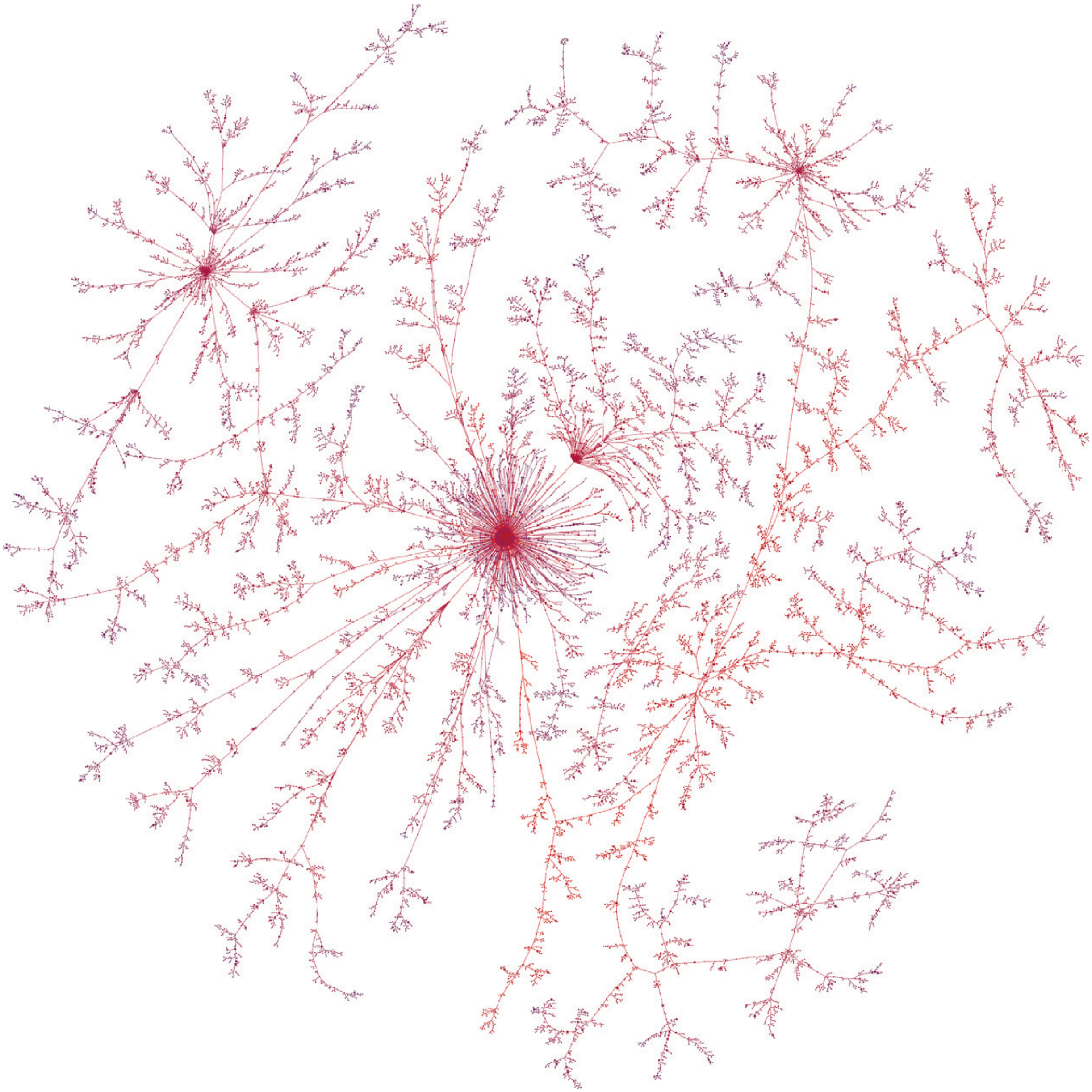

Figure 2

The spread of a rumor on Facebook. Each point is a person, and an edge (connecting line) indicates a share of the rumor. Earlier shares of the rumor are darker red.

number of questions he or she posed to the secret keeper even while remaining unaware of the secret. It seemed as though, at the onset of the secret, the target of the secret was somehow alerted to the fact that something was amiss. In addition, the secret keeper and confidants forged closer social bonds during the period of secret keeping, with the keeper sending the confidant many more emails than before the secret was held.

Testing Heider's balance theory. In some cases, the Internet provides a means to test a classic theory with much greater precision than was previously possible. Early in the history of relationship research, Heider (1946) noticed an intuitive relationship, captured in the idiom, "the enemy of my enemy is my friend," and "the enemy of my friend is my enemy." Extending these notions, Heider, and later Cartwright & Harary (1956), developed the theory of structural balance that posits certain kinds of dynamics in the sentiments between people. Specifically, whenever there is a set of three people that have relationships with each other, there will be a psychological motivation to ensure structural balance, such that there are no conflicting ties in the triad.

One assumption in balance theory is that the relationships between two people are mutual: If you are my friend, then I am your friend, and if you are my enemy, than I am your enemy. In other words, balance theory assumes reciprocal relationships. When one considers directed relationships (e.g., you think I am your friend, but I secretly despise you), many more permutations of positive and negative relationships become possible, only a small subset of which balance theory makes specific predictions about.

An alternative theory, from a very different line of research, does make predictions about directed relationships. This is known as status theory (Guha et al. 2004, Leskovec et al. 2010a), and it suggests that negative relationships flow downstream and positive relationships flow upstream, so anyone higher in a hierarchy looks down on subordinates, whereas people lower in a hierarchy have positive relationships with superiors. Considering these different theories carefully, one can see when and what predictions they make about the probability of different types of triangular relationships between people (e.g., A likes B and C, B likes A and C, C dislikes A and B).

These theoretical predictions were tested by gathering data from three different sites that make explicit or implicit signed relationships between people (Leskovec et al. 2010b): Epinions (a product review site in which users can indicate trust or distrust in other users' opinions), Slashdot (a technology news website that allows users to tag each other as "friend" or "foe"), and Wikipedia (where users cast positive or negative votes for the promotion of one user to the role of administrator by another user). At the time of testing, the Epinions dataset had nearly 120,000 users with 841,000 relationships, the Slashdot dataset had over 82,000 users with approximately 550,000 relationships, and the Wikipedia dataset had over 7,000 users and approximately 104,000 relationships. Clearly, without the Internet it would be impossible to study this phenomenon at this scale.

Leskovec et al. (2010b) found that when you look at reciprocal relationships, balance theory holds true, but when you look at directed relationships, status theory holds true. Essentially, each theory works when you look at the level in which it was formulated. In addition, because these investigators decomposed the pattern of relationships into all different possible triangles, they extended the understanding of signed relationships beyond the domain on which either theory makes predictions.

Impression formation. With so many real-world interactions now mediated by the Internet, questions formerly asked about impression formation in offline contexts (e.g., based on physical appearance) have been redirected to online contexts. For example, a number of studies have examined the impressions elicited on the basis of Facebook profiles (Back et al. 2010) and Twitter posts (Qiu et al. 2012). People have great control over what they post on Facebook profiles and what they tweet, so it is tempting to think that people may use these mediums to portray themselves in an unrealistically positive light (e.g., via selective use of wall posts, comments, and photo tags). However, studies directly testing this question in the context of Facebook have shown that impressions conveyed by profiles tend to converge more strongly with measures of what the profile holders are really like (derived from self- and informant reports) than with measures of what the profile owners would ideally like to be like; these findings suggest that social media profiles convey fairly accurate personality impressions of profile owners (Back et al. 2010, Waggoner et al. 2009, Weisbuch et al. 2009). However, in other contexts where there is less expectation of meeting the target in person (e.g., online gaming profiles; Graham & Gosling 2013) or where there is an incentive to portray oneself in a positive light (e.g., dating websites; Ellison et al. 2012), the accuracy of impressions diminishes.

Analyses of language and social behavior. Interpersonal communication is a key feature of many aspects of social behavior. Communication through spoken language leaves no trace of the interaction, but communication through written language on the Internet does leave a trace, which enterprising psychologists can harvest and study (Pennebaker et al. 2003). For example, in the aftermath of the September 11, 2001, terrorist attacks on the World Trade Center, Back et al. (2010) analyzed the messages sent to text pagers before, during, and after the attack. The researchers were able to access the messages from the WikiLeaks website where they had been posted. Automated text analyses of the messages sent from 85,000 pagers were used to examine the dynamic development of sadness, anxiety, and anger; the analyses assessed emotions by capturing words related to the emotions (e.g., sadness was assessed via words such as "crying" and "grief"). In response to the attacks, sadness words showed only a small increase, anxiety words rose in the immediate aftermath but quickly returned to baseline levels, and anger words steadily increased. By mining this publicly available dataset, the researchers were able to glimpse an in-the-moment window onto psychological processes as they unfolded in the wake of a significant social event, something that would have been virtually impossible before the Internet first mediated (via messages) and then made available (via WikiLeaks) the real-world indicators of social behavior. However, as we note later, their analyses also revealed some pitfalls that can arise in such automated analyses (Back et al. 2011, Pury 2011).

Blogs are another domain of language that have proved particularly fruitful for psychological research, in part because blogs permit researchers to retroactively obtain critical baseline measures that would otherwise be difficult to access. Again, in the context of responses to the September 11 terrorist attacks, Cohn et al. (2004) analyzed more than 1,000 online journal entries spanning the period two months before and two months after the attacks. Soon after the attacks, the bloggers' language conveyed increases in negative emotions, cognitive and social engagement, and psychological distance. Within a two-week period, the bloggers' moods and social referencing had returned to preattack levels, and their levels of psychological distancing were reduced. Their levels of cognitive processing dropped sharply from their peak immediately after the attack and continued to decline beyond baseline levels. Together, the patterns point to the conclusion that the cognitive effects of traumatic events may outlast the emotional effects.

Another creative use of blogs analyzed the links between personality and language use in over 500 bloggers (Yarkoni 2010). The analyses provided fascinating portraits of what bloggers high and low on the Big Five personality dimensions do and do not talk about. For example, the words used by people high (versus low) on Neuroticism essentially reflect their focus on what is bad in the world and what could go wrong, by using words such as awful, worse, depressing, terrible, stressful, horrible, and annoying. In contrast, the blogs of people high (versus low) on Agreeableness focus on the positive and connections with others, as reflected by high use of words such as wonderful, together, beautiful, joy, visiting, and hug.

Other domains in which researchers have studied language include Twitter (e.g., De Choudhury et al. 2013b), Facebook wall posts (e.g., Bazarova et al. 2013), and online support groups (e.g., Vambheim et al. 2013). For example, De Choudhury and colleagues (2013b) developed a statistical model to estimate the risk of depression from an individual's tweets over a year preceding the onset of the depressive episode; predictors of subsequent depression included linguistic signs of decreases in social activity, increases in negative affect, and greater religious involvement. Patterns of depression derived from these Twitter-based analyses were generally consistent with the seasonal, diurnal, gender-based, and geographic patterns published by the Centers for Disease Control and Prevention (De Choudhury et al. 2013a), providing support for the validity of the method.

The emergence of linguistic conventions on the Internet. Another study that capitalized on the fact that the Internet is a voluminous record of written language examined the emergence of a linguistic convention (Kooti et al. 2012). The authors focused on a particular linguistic convention on the Twitter platform, namely the way to indicate that a tweet is a repost of another's tweet and attributing the source.

The very first convention used was "via," which was borrowed from natural language and therefore was very intuitive to use. Shortly thereafter the variation "HT" (short for hat tip) was introduced, this time borrowed from a convention among bloggers. However, by the end of data collection three years later, the most dominant variation was "RT" (short for retweet). This is interesting because it is not as intuitive as via and is the same length as HT, yet it appeared later. In between HT and RT, people started to use "retweet" and "retweeting," a variant that is costly in the context of Twitter, which caps messages at 140 characters and therefore puts a premium on extreme brevity. Nevertheless, these variants took off, perhaps because they served as a signal of belonging to the Twitter community (Berger & Heath 2008). Presumably the cost (in terms of length) was not sustainable, because the first use of RT was in a tweet of exactly 140 characters by someone who had previously used both retweet and retweeting. After its first use, RT took off. Still, during the early days of RT, the recycle symbol emerged as an even more cost-effective variant, serving the role with a single character. Moreover, the variant was introduced by a hub in the network, one of the founders of Twitter. Despite these advantages, the recycle symbol was never widely used, possibly because of the difficulty of replicating the symbol and possibly because RT was already on its way to popularity.

Kooti and colleagues (2012) also discovered that for any variant, the original users were much more likely to describe themselves with tech-related words (e.g., "developer," "geek," or "entrepreneur"), had a generally higher in-degree, were more central in the network, and were more likely to use different features in the Twitter platform (e.g., have a written bio or have a list of Twitter users) than were typical users. All of this supports Rogers's (1962) theory of the diffusion of innovations that posits the first to adopt a new technology will be people predisposed to exploration and innovation.

Another interesting finding in this study is that, unlike the diffusion of other information such as URLs (Bakshy et al. 2012), the adoption of a particular variant typically required multiple exposures, making it a form of "complex contagion" (Centola & Macy 2007). In other words, the choice to adopt a variant on a linguistic convention seems to be less a result of "informational influence" and more a form of "normative influence" (Cialdini & Goldstein 2004, Deutsch & Gerard 1955). This distinction affects the form of the diffusion network, which for informational influence is less clustered and more tree-like, whereas for normative influence it has many cross-cutting branches and many triangles, indicating the use and reuse of a particular variant as it circulates in a community and gains acceptance.

Massive experiment on voting behavior. On the day of the US congressional election in 2010, researchers ran an experiment using a message that appeared in the newsfeed of 61 million Facebook users encouraging them to vote (Bond et al. 2012). Roughly 60 million users received a "social information" message that included a button that said "I voted," a link to information about local polling stations, and crucially, an array of up to six profile pictures of their friends who had clicked on the "I voted" button. In addition, approximately 500,000 people were in the informational condition, which had all of the above content except the array of friends' profile pictures, and 500,000 people were in the control condition and were shown no information.

The social information condition led to more clicks of the "I voted" button than did the purely informational condition. The social information also led to more clicks on the link to information on polling locations. The most amazing part of this study, however, is that the authors were able to match a sizeable number of the participants to actual voting records, which were released to the public domain; these data revealed that the social information condition led to significantly more actual votes than those in the control condition. There was no difference between the nonsocial information condition and the control condition in actual voting behavior.

Prior work had established that, at least within households, voting information can have contagious effects on voting behavior, so Bond et al. (2012) also looked at the effect of the message on the Facebook users' friends. They found a small but significant effect of the message on the likelihood of close friends actually going out and voting. In estimated real numbers, the direct effect of the message led to an additional 60,000 votes, and the indirect effect (through friends) led to an additional 280,000 votes, equivalent to an increase in turnout of the eligible voting population of about 0.14%, a substantial proportion of the actual increase in turnout from 2006 to 2010 of 0.6%.

Cautionary Note: Studying a Moving Target

Research on the Internet is complicated by the rapid and considerable changes in the Internet since it became an object of study. These changes generally occur much more quickly than changes found in typical offline contexts. For example, in the domain of Internet addiction, the definition of the Internet itself changed from 1998, when the term was used "to denote all types of on-line activity," including information protocols, WWW, email, news groups, MUDs, and chat rooms (Young 1998), to 2009, when the term also captured kinds of Web content (e.g., blogs, social networking sites, collaboratively authored encyclopedias) that did not exist a decade earlier (Byun et al. 2009).

These rapid changes in content are exacerbated by changes in the ways that Web content is consumed. The most visible change over the past decade has been the skyrocketing use of mobile devices. Indeed, it is unlikely that the social movements that found fertile ground in social media would have blossomed as they did if it weren't for the ability to access the Internet via mobile phones (Tufekci & Wilson 2012) or if the social media sites such as Twitter and Facebook had not existed in the first place.

Psychological research on the Internet, whether using novel techniques to study classic problems or studying novel phenomena specific to the Internet itself, must be conducted and consumed cautiously. Research must be designed to account for the (online) environment in which it is being conducted, and importantly, should be documented sufficiently to provide context about the environment. For example, the linguistic convention on Twitter described previously (Kooti et al. 2012) became obsolete when Twitter introduced the retweet button, and if the study were conducted on data after the retweet button was introduced, the outcome may have been quite different. So it is vital for researchers and readers to keep such characteristically rapid changes in mind.

Summary. The past decade of research has yielded many attempts to harness the benefits of Internet research to address longstanding questions in psychological science. Many of these studies have brought added statistical power, greater generalizability, and improved ecological validity to research. In a sense, the advent of the Internet has provided a viable domain for psychological researchers to leave the artificial confines of lab studies, surveys, and questionnaires and study real behavior as it unfolds in the real world.

METHODOLOGICAL ISSUES

We next turn to some enduring methodological issues in the domain of Internet research.

Recruitment

Since the 1960s, repeated calls have been made to extend the empirical base of psychology beyond the college undergraduates that have become the go-to population for research in psychology, especially in the United States (Sears 1986). For decades these calls went unheeded. However, as researchers began gathering Internet samples, it quickly became apparent that this new medium could easily reach well beyond the traditional samples of college students (Gosling et al. 2004). Below, we discuss four broad recruitment strategies.

Delivery to standard populations. In its most basic form, the Internet can be used as a new method for delivering questionnaires and experiments to participants. For example, rather than distributing paper-and-pencil measures to participants, a researcher might simply direct participants to a website where the questionnaire or experiment is hosted. Doing so can take advantage of the various benefits of Internet studies (e.g., data checks, adaptive testing, providing feedback), thereby increasing the validity and efficiency with which data can be collected (Kraut et al. 2004). One can also use online panels—groups of people who have volunteered to be contacted for online research—which allows the research to have a more reliable pool of participants and facilitates the collection of longitudinal data that otherwise might be expensive or difficult (Göritz 2004).

Respondent-driven sampling. The use of respondent-driven sampling methods has become particularly popular for recruiting participants (Goel & Salganik 2010). Such methods recruit participants through word of mouth, often via social media and other informal channels. One element that has been essential for making such strategies effective is generating interesting feedback for users. The outofservice.com website has accumulated samples with millions of participants by offering no more than automated personalized feedback. The integration of feedback with social media allows word of mouth to spread rapidly as participants post their results on Facebook or tweet about them. An increasing number of surveys and experiments are delivered as apps within social networks (e.g., MyType, myPersonality). For example, the myPersonality Facebook app provides participants with feedback on their personality ratings and permits these data to be combined with other surveys and other information made available to the researchers, with the users' permission, via Facebook (Stillwell & Kosinski 2012). It is important to design these studies with the analysis of the results in mind because traditional methods for estimating an effect are likely to mask variability in the estimates (Goel & Salganik 2010).

Delivery to targeted populations. Many interesting questions require tapping populations that are hard to reach by conventional means. With so many specialist groups online, many populations can now be targeted via online forums and interest groups. For example, for a study of motivations driving online game-playing behavior, researchers recruited participants via an advertisement posted on wowinsider.com, a popular news website for players of World of Warcraft, a massively multiplayer role-playing game (Graham & Gosling 2013).

Many populations in which psychologists might be interested are hard to reach because they are engaging in behavior that is socially undesirable or embarrassing. An example of the former is the research done on hate groups, in which investigators posed as newbies on hate-group chat rooms to interview white supremacists about their attitudes regarding advocating violence toward blacks (Glaser et al. 2002); without the anonymity afforded to the researchers and the participants by the chat-room context, it would have been difficult to obtain access to these individuals.

Researchers can also use the Internet to target participants (e.g., bloggers, tweeters) even if they have not self-selected into specialist groups. For example, a researcher who wants to contact people who are interested in a topic or who are from a specific location could search for users posting on that topic or in that location using tags (e.g., in Instagram) or using general web search engines. Another option for reaching desired populations is advertising via such services as Facebook, where participants can be targeted on such features as demographics, geographic location, and preferences (e.g., via likes).

In addition to surveying or recruiting participants, targeted strategies can be useful for identifying relevant existing data. For example, to examine patterns of behavior, feelings, thoughts, and social connections expressed by people suffering from eating disorders, one study analyzed the language of blogs focused on eating disorders; analyses compared the content of pro-eating-disorder blogs, recovery blogs, and control blogs (Wolf et al. 2013). Another study examined the role of thinking styles and social connections as possible explanations for why Christians tend to be happier than atheists (Ritter et al. 2014); by examining approximately two million tweets from over 16,000 Twitter users who followed either Christian or atheist leaders, the researchers showed that intuitive (versus analytic) thinking style and frequency of words related to social relationships partially mediated the higher happiness of Christians relative to atheists.

Crowdsourcing. One recent trend has been the use of crowdsourcing websites, most prominently Amazon's Mechanical Turk (MTurk; Buhrmester et al. 2011, Mason & Suri 2012, Paolacci et al. 2010). Crowdsourcing websites such as MTurk, CrowdFlower, and Clickworker are online marketplaces designed to match individuals who have tasks [known as HITs (human intelligence tasks)] they need completed (requesters) with people willing to complete those tasks (workers). For psychological researchers, the tasks are surveys and experiments. The system provides a mechanism for creating the tasks, matching requesters with workers, compensating the workers, and rating the quality of the requests and worker performance. Requesters can create and post virtually any task that can be done at a computer (i.e., surveys, experiments, writing, etc.) using simple templates or technical scripts or linking workers to external online survey tools (e.g., SurveyMonkey). Workers can browse available tasks and are paid upon successful completion of each task. Requesters can refuse payment for subpar work.

These sources have proved to be exceptionally popular because they typically allow data to be collected rapidly and inexpensively and reach populations that are more diverse than typical student samples, without compromising data quality (e.g., Weinberg et al. 2014). For example, in a study of data quality obtained using a series of classic experiments on judgment and decision making, Paolacci et al. (2010) found that MTurk participants had better completion rates than did participants selected via Internet discussion boards, with the effects replicating across recruitment methods. Another series of studies replicated the effects from key experiments in cognitive psychology using MTurk participants and did so in a fraction of the time taken to run studies in the lab (Crump et al. 2013). Moreover, a relatively straightforward series of steps (e.g., inclusion of comprehension checks) could bring the data quality up to the standards of lab studies (Crump et al. 2013).

As more and more psychologists have turned to MTurk for their samples, some issues have arisen as a result of its popularity. For example, with greater numbers of psychologists using MTurk, it is becoming increasingly difficult to find naïve participants who have not taken part in psychology studies before. A nonnegligible number of MTurk workers participate in multiple studies, with some of the most prolific workers (superturkers) accounting for a disproportionately high number of HITs (Chandler et al. 2014); Chandler et al. analyzed the participants in 132 studies and found that the top 1% of the most prolific workers completed 11% of the total HITs. The very prolific workers can be especially problematic in cases where there is some selection criterion (e.g., being left-handed, a first born, or a smoker) because the MTurk workers can share information about selection criteria, manipulation checks, and frequently used designs and measures on the community message boards. However, once researchers are aware of such issues, they can be dealt with by using careful prescreening techniques (see Chandler et al. 2014), and these superturkers can actually be helpful for some designs (e.g., building longitudinal panels).

Sample Diversity

One persistent critique of Internet studies has been that the samples are selective and are not diverse. Ideally, sample populations would be fully representative; however, the appropriate point of comparison is with conventional and viable alternatives. To get a sense of the characteristics of viable alternative samples, one study examined a year's worth of conventional samples published in the *Journal of Personality and Social Psychology* (Gosling et al. 2004). The characteristics of these conventional samples were then compared to the characteristics of a large Internet sample. The comparison revealed that the Internet sample was more diverse and more representative of the general US population with respect to gender, socioeconomic status, geographic region, and age than were the conventional samples.

One particularly important concern about sample diversity stems from the so-called digital divide, which refers to the differential access to the Internet based on socioeconomic class and other demographic variables. In Western countries, there are concerns that individuals from lower classes and some minority groups are severely underrepresented in Internet samples (e.g., Van Dijk & Hacker 2003). One analysis looking at the representativeness of a large Internet sample with respect to race and social class at the level of US states revealed that the sample was reasonably representative (Rentfrow et al. 2008). The analysis was done by correlating the percentages of certain demographic groups from each state as determined by the Internet sample and by the US Census Bureau's figures. The correlations for African Americans, Asians, Latinos, whites, and "other" ethnicities were 0.88, 0.96, 0.96, 0.93, and 0.74, respectively. With regard to social class, the correlations for working-, lower-middle-, middle-, upper-middle-, and upper-class participants were 0.52, 0.64, 0.41, 0.66, and 0.43, respectively. So, overall, the Internet-based sample was generally representative of the population at large with respect to ethnicity, but it somewhat underrepresented individuals from lower and upper classes. Internet samples are also known to overrepresent younger participants, and the gender composition of samples can vary widely depending on the sampling method (Gosling et al. 2004). Regardless, researchers should report the characteristics of their samples in detail so readers can evaluate claims regarding generalizability accordingly.

Internet samples are biased in favor of those who are sufficiently educated and wealthy to access the Internet. However, this selectivity should diminish as Internet connectivity reaches saturation in Western societies and continues to grow elsewhere. Already, the Internet can be used to reach samples that only a decade ago would have been hard to access. This feature was well illustrated in a study examining differences in the trajectories of age-related personality changes across different cultures (Bleidorn et al. 2013). Unsurprisingly, the US subsample was much larger than those from other countries. Nonetheless, there were sufficient numbers of participants from other countries meeting their inclusion criteria (i.e., at least 100 participants in each of five age bands spanning 16–20, 21–25, 26–30, 31–35, and 36–40) to examine 62 different countries, including such countries as El Salvador, Serbia, Iran, and Singapore.

In fact, Internet samples can be used to address concerns that psychological science overly represents participants from societies that are Western, educated, industrialized, rich, and democratic (WEIRD; Henrich et al. 2010). One analysis revealed that Internet samples, although still far from perfect, are substantially more diverse than conventional samples; specifically, one Internet sample that collected personality data using respondent-driven sampling in English, German, Dutch, and Spanish found that 19% of participants were not from advanced economies, 20% of participants were from non-Western societies, 35% of the Western society participants were not from the United States, and 66% of the US participants were not in the 18–22 (college) age group (Gosling et al. 2010). These numbers may be far from ideal, but they are better than most viable alternatives, and given that Internet samples can be very large, even small percentages represent large absolute numbers; for example, the 19% of participants not from advanced economies consisted of 104,928 participants.

Anonymity

One potential concern with Internet samples is that participants can often engage in the research with complete anonymity. This anonymity could lower the accountability of the participants, potentially increasing the chances that they would engage in sloppy responding, provide false answers, and participate in other behaviors detrimental to collecting valid data. On the other hand, under some circumstances, the anonymity may actually constitute an advantage, especially when the studies involve topics that are embarrassing or illegal (Rains 2014). For example, one study of sexsomniacs (a medical condition in which individuals engage in sexual activity during their sleep) took advantage of the anonymity afforded by the Internet to survey individuals who might be too embarrassed to report on their condition in other circumstances (Mangan & Reips 2007); the reach and anonymity afforded by the Internet allowed the researchers to gather data on five times more sexsomnia sufferers than had been reached in all previous studies combined from the previous two decades of research.

Techniques for Improving the Quality of Data Collected Via the Internet

Data gathered via the Internet are frequently of equal or even higher quality than those gathered via traditional means (Dodou & de Winter 2014, Gosling et al. 2004, Luce et al. 2007). This effect is likely a result of relying on participants who are intrinsically motivated to take part rather than the ubiquitous college samples that tend to be doing the studies reluctantly. Nonetheless, key threats to the validity of data collected via the Internet remain. These threats largely stem from the fact that researchers cannot easily supervise participants to make sure they are whom they say they are and that they are responding truthfully, and researchers cannot monitor the participants' alertness and attentiveness (Johnson 2005). A variety of methods have been developed to combat some of the concerns.

Methods have been developed to detect the degree to which participants are attending to the experimental materials and following instructions properly (Johnson 2005). One tool—the Instructional Manipulation Check (IMC)—detects participants who are not paying close attention or following instructions (Oppenheimer et al. 2009); it does so by including a question that looks similar to the standard questions in length and response format but instead requires that participants do something unexpected (e.g., clicking a small blue circle rather than the standard response options). By removing error from participant responses, the IMC increases statistical power and improves the reliability of a dataset.

Many additional strategies can be implemented to improve data quality. These strategies include (*a*) using automated procedures designed to monitor data quality (e.g., flagging participants who take unusually short or long times to complete the study or who engage in unlikely response patterns; Johnson 2010); (*b*) using automated procedures to alert participants to errors (e.g., missed items; Johnson 2010); (*c*) employing statistical checks of the data (e.g., evaluating reliability and factor structure; comparing with previous research); (*d*) implementing techniques to reduce incentives to provide false information (e.g., in a survey that offers feedback as an incentive, researchers can provide a link to all possible forms of feedback to stop participants from submitting multiple responses to see what other kinds of feedback they could get); (*e*) providing participants with a means for indicating which responses are valid (e.g., providing a check box where participants can indicate if they have taken the survey before); (*f*) increasing accountability by recruiting participants where faking one's identity is not trivial (e.g., the personality test administered by myPersonality is done via Facebook, so multiple Facebook accounts would be needed to take the test more than once; Stillwell & Kosinski 2012); (*g*) offering appropriate incentives (Göritz 2010); and (*h*) formatting the materials to optimize clarity and usability (Reips 2010).

Ethics

Some of the features that raise methodological challenges to doing Internet research also pose ethical issues that must be considered by anyone doing research online. Specifically, the facts that researchers have less control over and knowledge of the research environment and cannot monitor the experience of the participants, or indeed their true identities, raise a number of ethical issues (Buchanan & Williams 2010). For example, it is quite possible that participants could drop out of an experiment or a technical issue could interrupt the experiment before the participants have been properly debriefed. In addition, it would be possible for people to misrepresent their identity in ethically relevant ways (e.g., minors representing themselves as adults).

On the other hand, in some instances Internet research can provide a favorable ethical context in comparison with traditional contexts for research (Buchanan & Williams 2010). For example, participants taking a study online can simply close the browser if they wish to discontinue a study, making participation less coercive than in a context where an experimenter is present. In addition, Internet research facilitates a greater degree of anonymity than can easily be guaranteed in traditional in-person studies.

A number of ethical issues with Internet research arise, in part, because of the need to apply principles that were generally developed before the existence of the Internet and also because of the rapid developments in technology. One key issue that has continued to elicit discussion is the question of what counts as public behavior (Buchanan & Williams 2010). Should blogs and entries on discussion boards be considered public behavior? Some researchers have argued that virtually anything online can be considered public behavior and is therefore subject to the same ethical guidelines as other public behaviors. Others have argued that in some contexts, such as support groups, members may reasonably expect their behaviors not to be subject to research. As new platforms and technologies continue to be developed, raising new ethical issues, researchers can refer to a list of 16 systematic questions developed by Buchanan & Williams (2010) for framing their ethical implications; these questions concern such issues as traceability, dropout rates, inducements to participate, and data security. In addition, researchers can refer to the guidelines developed by the Association of Internet Researchers (Markham & Buchanan 2012).

EMERGING DIRECTIONS

By its nature, the field of Internet research changes rapidly, with new research possibilities emerging monthly. Here we highlight two emerging domains that seem set to be particularly influential in the near future.

Big Data

One major new avenue of research made viable by the Internet is the exploitation of very large datasets—sometimes known as big data—to address psychological questions. The sizes of these datasets often have a number of desirable features, including high statistical power, a broad reach of participants, and the ability to tap real-world behaviors. These benefits were nicely demonstrated in the massive 2012 Facebook-based experiment on voting behavior noted above, in which social information was used to encourage people to vote (Bond et al. 2012).

Increasingly, these large datasets are collected specifically to address psychological questions. For example, the myPersonality dataset collected via a Facebook application has collected millions of personality questionnaires (Stillwell & Kosinski 2012); when users sign up for the app, they also provide consent for the app to collect information on their demographic details, likes (a thumbs-up sign to express positive regard for online content), photos, wall posts, and other information, providing a rich trove of information on a major medium of social behavior. As a result, the myPersonality dataset has already yielded a large number of studies, many of which could only be undertaken with the vast quantities of data to which the researchers have access. For example, in one paper, the research team used Facebook likes to statistically predict—with great accuracy—a wide range of traits and attributes, such as sexual orientation, ethnicity, gender, age, religious and political views, and personality traits (Kosinski et al. 2013).

Other very large datasets can be found or generated by enterprising researchers even if they were not originally collected for the purpose of testing psychological questions. Examples include the studies of Twitter noted above, in which researchers used massive samples of tweets to examine the spread of social conventions (Kooti et al. 2012).

Another data-mining approach, psychoinformatics (Yarkoni 2012), aims to apply tools and techniques from the computer and information sciences to improve the acquisition, organization, and synthesis of psychological data. One particularly compelling example is the development of a program called Neurosynth (Yarkoni et al. 2011), which searched the research literature for articles on neuroimaging topics (e.g., pain, emotion) and then automatically extracted activation coordinates from all tables reported in these studies. These coordinates were then aggregated and mapped to automatically extracted psychological terms. In their demonstration of the method, Yarkoni et al. (2011) automatically synthesized the results of nearly 3,500 neuroimaging studies.

Of course, such datasets can be subject to a range of concerns not typically found with conventional datasets. When enormous samples are analyzed, the details of individual cases are easily lost, with the result that some large-scale errors can occur. For example, in the dataset of pager messages described previously and analyzed by Back et al. (2010), the researchers coded words signaling anger. However, subsequent analyses revealed that 35.9% of the anger words were an automatically generated message sent with great and increasing frequency to a single pager (Pury 2011). This message included the word "critical," which was used to indicate the message was urgent, but it influenced the analyses because "critical" was one of the words used to mark anger in the Back et al. analyses. This error initially went unnoticed, underscoring the need for careful control mechanisms to be implemented in both the preparation and the analysis of large datasets (Back et al. 2011).

Smartphone Sensing

Smartphones have the potential to be a powerful tool for psychological research. They are almost constantly near the person and have the capability to detect features of the environment (e.g., noise, light, the presence of others), record online and offline behaviors (e.g., physical activity, sleep, conversations, texting, Web surfing), and provide timely automated interventions. As a result, smartphones are poised to have a big impact on how data are collected in psychological science (Miller 2012), realizing the promise of ambulatory assessment (Mehl & Conner 2012).

Many of the sensing functions served by a smartphone could technically be performed without Internet connectivity, but the value of these functions is magnified by their connection to the Internet. In particular, Internet connectivity allows real-time monitoring, the possibility of instantaneously uploading and processing data in the cloud, the ability to provide participants with real-time feedback and interventions, and the ability to monitor the aspects of social and psychological behavior that are mediated via the Internet, such as texting, emailing, information searching, Web surfing, and consuming media.

A number of challenges must be faced before smartphone sensing becomes commonplace. For example, smartphones are still relatively expensive, resulting in problems of sample selectivity unless the devices are provided to the participants at the cost of the researcher. Additionally, the sensors are so sensitive that issues of privacy are raised regarding both the participants and others in their vicinity. However, the biggest practical obstacles to the large-scale adoption of smartphone sensing methods in psychological research presently are the technical challenges of creating and maintaining the apps. These tasks are typically beyond the capability of most psychologists and require collaborations between computer scientists and psychologists.

A number of conferences and workshops promoting such collaborations have been held (e.g., Campbell & Lane 2013, Mascolo & Rentfrow 2011, Rentfrow & Gosling 2012), and computer scientists have been proactive in developing and demonstrating apps and initiating collaborations with psychologists. For example, as early as 2007, Andrew Campbell and collaborators demonstrated CenceMe, which used sensors in a smartphone to track the amount of time engaged in activities such as sitting, standing, walking, running, and chatting (Miluzzo et al. 2007). Researchers at Cambridge University and elsewhere have also taken a lead in developing smartphone apps for use in psychological research (e.g., Lathia et al. 2013, Rachuri et al. 2010, Wrzus et al. 2012). Thus, studies using smartphone sensing are still quite rare but are becoming increasingly prevalent as the practical challenges are solved. Moreover, with the widespread development and adoption of wearable sensors (e.g., to measure activity, heart rate, posture), the capabilities of smartphone sensing are continuing to increase dramatically.

CONCLUDING POINTS

The rapid and global expansion of the Internet has transformed it from an oddity to a daily presence in an incredibly large and diverse population. As such, the Internet is becoming a fundamental part of psychological research. In this review we have discussed how traditional methods have been adapted to the online setting; how new phenomena that owe their existence to the Internet, such as Internet addiction and cyberbullying, have become topics of psychological research; and how novel research paradigms have been developed using the new capabilities of the Internet. We also discussed new methodological issues that arise when doing research, such as the anonymity the Internet provides, as well as new methodological opportunities, such as data mining big data.

Research on the Internet comes with new challenges, but the widespread adoption of Internet methods in the field suggests that the opportunities far outweigh the costs. With the ability to reach large, diverse samples and collect data on actual behaviors, psychologists have little excuse for relying exclusively on college samples and contrived lab studies. Although small-scale laboratory research will always be a necessary approach to answer some questions, the Internet can provide a broader set of participants and methods that will ultimately increase the impact of psychological research on society.

DISCLOSURE STATEMENT

Winter Mason is currently employed by Facebook, Inc., which is mentioned several times in the article. However, at the time of writing he was employed by Stevens Institute of Technology.

ACKNOWLEDGMENTS

We are grateful to Friggeri and colleagues (2014) for preparing **Figure 2**. Preparation of this article was supported, in part, by a Brotherton Fellowship from the University of Melbourne.

LITERATURE CITED

Adamic LA, Glance N. 2005. The political blogosphere and the 2004 U.S. election: Divided they blog. *Proc. WWW2005 Workshop Weblog. Ecosyst., Jpn.*, pp. 36–43

Allport GW, Postman LJ. 1947. *The Psychology of Rumour*. New York: Holt, Rinehart & Winston

Amichai-Hamburger Y, Ben-Artzi E. 2003. Loneliness and Internet use. *Comput. Hum. Behav.* 19(1):71–80

Aral S, Walker D. 2011. Creating social contagion through viral product design: a randomized trial of peer influence in networks. *Manag. Sci.* 57(9):1623–39

Back MD, Küfner ACP, Egloff B. 2010. The emotional timeline of September 11, 2001. *Psychol. Sci.* 21:1417–19

Back MD, Küfner ACP, Egloff B. 2011. "Automatic or the people?" Anger on September 11, 2001, and lessons learned for the analysis of large digital data sets. *Psychol. Sci.* 22(6):837–38

Back MD, Stopfer JM, Vazire S, Gaddis S, Schmukle SC, et al. 2010. Facebook profiles reflect actual personality, not self-idealization. *Psychol. Sci.* 21:372–74

Bakshy E, Hofman JM, Mason WA, Watts DJ. 2010. Identifying "influencers" on Twitter. *ACM Int. Conf. Web Search Data Min., 4th, Hong Kong*, pp. 1–10

Bazarova NN, Taft JG, Choi Y, Cosley D. 2013. Managing impressions and relationships on Facebook: Self-presentational and relational concerns revealed through the analysis of language style. *J. Lang. Soc. Psychol.* 32(2):121–41

Berger J, Heath C. 2008. Who drives divergence? Identity signaling, outgroup dissimilarity, and the abandonment of cultural tastes. *J. Personal. Soc. Psychol.* 95(3):593–607

Bleidorn W, Klimstra TA, Denissen JJA, Rentfrow PJ, Potter J, Gosling SD. 2013. Personality maturation around the world: a cross-cultural examination of social investment theory. *Psychol. Sci.* 24:2530–40

Block J. 2008. Issues for DSM-V: Internet addiction. *Am. J. Psychiatry* 165(3):306–7

Bond RM, Fariss CJ, Jones JJ, Kramer ADI, Marlow C, et al. 2012. A 61-million-person experiment in social influence and political mobilization. *Nature* 489:295–98

Bordia P, DiFonzo N. 2004. Problem solving in social interactions on the Internet: rumor as social cognition. *Soc. Psychol. Q.* 67(1):33–49

Brenner V. 1996. An initial report on the online assessment of Internet addiction: the first 30 days of the Internet Usage Survey. *Psychol. Rep.* 70:179–210

Buchanan T, Williams JE. 2010. Ethical issues in psychological research on the Internet. See Gosling & Johnson 2010, pp. 255–71

Buhrmester M, Kwang T, Gosling SD. 2011. Amazon's Mechanical Turk: a new source of inexpensive, yet high-quality, data? *Perspect. Psychol. Sci.* 6:3–5

Byun S, Ruffini C, Mills JE, Douglas AC, Niang M, et al. 2009. Internet addiction: metasynthesis of 1996–2006 quantitative research. *CyberPsychol. Behav.* 12(2):203–7

Cacioppo JT, Fowler JH, Christakis NA. 2009. Alone in the crowd: the structure and spread of loneliness in a large social network. *J. Personal. Soc. Psychol.* 97(6):977–91

Campbell AT, Lane ND. 2013. *Smartphone sensing: a game changer for behavioral science*. Workshop held at Univ. Calif., Davis

Cartwright D, Harary F. 1956. Structural balance: a generalization of Heider's theory. *Psychol. Rev.* 63(5):277–93

Centola D, Macy M. 2007. Complex contagions and the weakness of long ties. *Am. J. Sociol.* 113(3):702–34

Chandler J, Mueller P, Paolacci G. 2014. Nonnaïveté among Amazon Mechanical Turk workers: consequences and solutions for behavioral researchers. *Behav. Res. Methods* 46:112–30

Chou C, Hsiao M-C. 2000. Internet addiction, usage, gratification, and pleasure experience: the Taiwan college students' case. *Comput. Educ.* 35(1):65–80

Christakis NA, Fowler JH. 2007. The spread of obesity in a large social network over 32 years. *N. Engl. J. Med.* 357:370–79

Cialdini RB, Goldstein NJ. 2004. Social influence: compliance and conformity. *Annu. Rev. Psychol.* 55(1):591–621

Cohen R. 2011. Facebook and Arab dignity. *New York Times*, Jan. 24. **http://www.nytimes.com/2011/01/25/opinion/25iht-edcohen25.html**

Cohn MA, Mehl MR, Pennebaker JW. 2004. Linguistic markers of psychological change surrounding September 11, 2001. *Psychol. Sci.* 15:687–93

Conover M, Ratkiewicz J, Francisco M, Gonçalves B, Menczer F, Flammini A. 2011. Political polarization on Twitter. *Proc. 5th Int. Conf. Weblogs Soc. Media*, pp. 89–96

Coviello L, Sohn Y, Kramer ADI, Marlow C, Franceschetti M, et al. 2014. Detecting emotional contagion in massive social networks. *PLOS ONE* 9(3):e90315

Crump MJC, McDonnell JV, Gureckis TM. 2013. Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLOS ONE* 8:e57410

Davis RA. 2001. A cognitive-behavioral model of pathological Internet use. *Comput. Hum. Behav.* 17(2):187–95

De Choudhury M, Counts S, Horvitz E. 2013a. Social media as a measurement tool of depression in populations. *Proc. 5th ACM Int. Conf. Web Sci., Paris, Fr.*, May 2–May 4

De Choudhury M, Gamon M, Counts S, Horvitz E. 2013b. Predicting depression via social media. *Proc. 7th Int. AAAI Conf. Weblogs Soc. Media, Boston, MA*, July 8–July 10

Deutsch M, Gerard HB. 1955. A study of normative and informational social influences upon individual judgment. *J. Abnorm. Soc. Psychol.* 51(3):629–36

DiFonzo N, Bordia P. 2007. *Rumor Psychology: Social and Organizational Approaches*. Washington, DC: Am. Psychol. Assoc.

Dodou D, de Winter JCF. 2014. Social desirability is the same in offline, online, and paper surveys: a meta-analysis. *Comput. Hum. Behav.* 36:487–95

Ellison NB, Hancock JT, Toma CL. 2012. Profile as promise: a framework for conceptualizing veracity in online dating self-presentations. *New Media Soc.* 14(1):45–62

Fowler JH, Christakis NA. 2008. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *BMJ* 337:23–27

Friggeri A, Adamic LA, Eckles D, Cheng J. 2014. Rumor cascades in social networks. *Proc. 8th Int. AAAI Conf. Weblogs Soc. Media (ICWSM)*

Garrett RK. 2009. Echo chambers online? Politically motivated selective exposure among Internet news users. *J. Comput.-Mediat. Commun.* 14(2):265–85

Ge M, Delgado-Battenfeld C, Jannach D. 2010. Beyond accuracy: evaluating recommender systems by coverage and serendipity. *Proc. 4th ACM Conf. Recomm. Syst.*, pp. 257–60. New York: ACM

Glaser J, Dixit J, Green DP. 2002. Studying hate crime with the Internet: What makes racists advocate racial violence? *J. Soc. Issues* 58:177–93

Goel S, Salganik MJ. 2010. Assessing respondent-driven sampling. *Proc. Natl. Acad. Sci. USA* 107(15):6743–47

Goel S, Watts DJ, Goldstein DG. 2012. The structure of online diffusion networks. *Proc. 13th ACM Conf. Electron. Commerce*, pp. 623–38. New York: ACM

Goldberg I. 1996. *Internet addiction disorder*. **http://www.urz.uni-heidelberg.de/Netzdienste/anleitung/wwwtips/8/addict.html**

Göritz AS. 2004. Recruitment for online access panels. *Int. J. Mark. Res.* 46(4):411–25
Göritz AS. 2010. Using lotteries, loyalty points, and other incentives to increase participant response and completion. See Gosling & Johnson 2010, pp. 219–33
Gosling SD, Bonnenburg AV. 1998. An integrative approach to personality research in anthrozoology: ratings of six species of pets and their owners. *Anthrozoös* 11:148–56
Gosling SD, Johnson JA, eds. 2010. *Advanced Methods for Conducting Online Behavioral Research*. Washington, DC: Am. Psychol. Assoc.
Gosling SD, Sandy CJ, John OP, Potter J. 2010. Wired but not WEIRD: the promise of the Internet in reaching more diverse samples. *Behav. Brain Sci.* 33:94–95
Gosling SD, Vazire S, Srivastava S, John OP. 2004. Should we trust Web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. *Am. Psychol.* 59:93–104
Graham LT, Gosling SD. 2013. Personality profiles associated with different motivations for playing World of Warcraft. *Cyberpsychol. Behav. Soc. Netw.* 16:189–93
Greenwald AG, McGhee DE, Schwartz JLK. 1998. Measuring individual differences in implicit cognition: the Implicit Association Test. *J. Personal. Soc. Psychol.* 74:1464–80
Guha R, Kumar R, Raghavan P, Tomkins A. 2004. Propagation of trust and distrust. *Proc. 13th Int. Conf. World Wide Web*, pp. 403–12. New York: ACM
Hannak A, Anderson E, Barrett LF, Lehmann S, Mislove A, Riedewald M. 2012. Tweetin' in the rain: exploring societal-scale effects of weather on mood. *Proc. 6th Int. AAAI Conf. Weblogs Soc. Med.*, pp. 479–82
Heider F. 1946. Attitudes and cognitive organization. *J. Psychol.* 21(1):107–12
Henrich J, Heine SJ, Norenzayan A. 2010. The weirdest people in the world? *Behav. Brain Sci.* 33:61–135
Herlocker JL, Konstan JA, Terveen LG, Riedl JT. 2004. Evaluating collaborative filtering recommender systems. *ACM Trans. Inform. Syst.* (*TOIS*) 22(1):5–53
Int. Telecommun. Union. 2014. *The World in 2014: ICT Facts and Figures.* Geneva, Switz.: Int. Telecommun. Union
Iyengar S, Hahn KS. 2009. Red media, blue media: evidence of ideological selectivity in media use. *J. Commun.* 59(1):19–39
Johnson JA. 2005. Ascertaining the validity of Web-based personality inventories. *J. Res. Personal.* 39:103–29
Johnson JA. 2010. Web-based self-report personality scales. See Gosling & Johnson 2010, pp. 149–66
Juvonen J, Gross EF. 2008. Extending the school grounds? Bullying experiences in cyberspace. *J. Sch. Health* 78(9):496–505
Kashima Y. 2000. Maintaining cultural stereotypes in the serial reproduction of narratives. *Personal. Soc. Psychol. Bull.* 26(5):594–604
Kirk M. (Writer/Producer) Feb. 22, 2011. Revolution in Cairo. *Frontline.* Boston, MA: WGBH Educ. Found. [telev. ser. episode]. **http://www.pbs.org/wgbh/pages/frontline/revolution-in-cairo/**
Kooti F, Yang H, Cha M, Gummadi KP, Mason WA. 2012. The emergence of conventions in online social networks. *Proc. 6th Int. AAAI Conf. Weblogs Soc. Media*, pp. 194–201
Kosinski M, Stillwell DJ, Graepel T. 2013. Private traits and attributes are predictable from digital records of human behavior. *Proc. Natl. Acad. Sci. USA* 110:5802–5
Kowalski RM, Giumetti GW, Schroeder AN, Lattanner MR. 2014. Bullying in the digital age: a critical review and meta-analysis of cyberbullying research among youth. *Psychol. Bull.* 140:1073–137
Kowalski RM, Limber SP. 2013. Psychological, physical, and academic correlates of cyberbullying and traditional bullying. *J. Adolesc. Health* 53(1):S13–20
Kraut R, Kiesler S, Boneva B, Cummings J, Helgeson V, Crawford A. 2002. Internet paradox revisited. *J. Soc. Issues* 58(1):49–74
Kraut R, Olson J, Banaji M, Bruckman A, Cohen J, Couper M. 2004. Psychological research online: report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. *Am. Psychol.* 59(2):105
Kraut R, Patterson M, Lundmark V, Kiesler S, Mukopadhyay T, Scherlis W. 1998. Internet paradox. A social technology that reduces social involvement and psychological well-being? *Am. Psychol.* 53:1017–31
Lane DJ, Wegner DM. 1995. The cognitive consequences of secrecy. *J. Personal. Soc. Psychol.* 69:237–53
Lathia N, Pejovic V, Rachuri KK, Mascolo C, Musolesi M, Rentfrow PJ. 2013. Smartphones for large-scale behavior change interventions. *IEEE Pervasive Comput.: Mobile Ubiquitous Syst.* 12:66–73

Leskovec J, Backstrom L, Kleinberg J. 2009. Meme-tracking and the dynamics of the news cycle. *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 497–506

Leskovec J, Huttenlocher D, Kleinberg J. 2010a. Predicting positive and negative links in online social networks. *Proc. 19th Int. Conf. World Wide Web*, pp. 641–50. New York: ACM

Leskovec J, Huttenlocher DP, Kleinberg JM. 2010b. Governance in social media: a case study of the Wikipedia promotion process. *Int. AAAI Conf. Weblogs Soc. Media* (*ICWSM*)

Leung L. 2004. Net-generation attributes and seductive properties of the Internet as predictors of online activities and Internet addiction. *CyberPsychol. Behav.* 7(3):333–48

Lim M. 2012. Clicks, cabs, and coffee houses: social media and oppositional movements in Egypt, 2004–2011. *J. Commun.* 62(2):231–48

Luce KH, Winzelberg AJ, Das S, Osborne MI, Bryson SW, Taylor CB. 2007. Reliability of self-report: paper versus online administration. *Comput. Hum. Behav.* 23:1384–89

Lyons R. 2011. The spread of evidence-poor medicine via flawed social-network analysis. *Stat. Polit. Policy* 2(1):Article 2

Mangan M, Reips UD. 2007. Sleep, sex, and the Web: surveying the difficult-to-reach clinical population suffering from sexsomnia. *Behav. Res. Methods* 39:233–36

Manis M, Cornell SD, Moore JC. 1974. Transmission of attitude relevant information through a communication chain. *J. Personal. Soc. Psychol.* 30(1):81–94

Markham A, Buchanan E. 2012. *Ethical Decision-Making and Internet Research: Recommendations from the AOIR Ethics Work. Comm. Tech. Rep.* **http://aoir.org/reports/ethics2.pdf**

Mascolo C, Rentfrow PJ. 2011. *Social Sensing: Mobile Sensing Meets Social Science*. Workshop held at Cambridge Univ., Cambridge, UK

Mason W, Suri S. 2012. Conducting behavioral research on Amazon's Mechanical Turk. *Behav. Res. Methods* 44:1–23

Mason WA, Conrey FR, Smith ER. 2007. Situating social influence processes: dynamic, multidirectional flows of influence within social networks. *Personal. Soc. Psychol. Rev.* 11(3):279–300

McKenna KY, Bargh JA. 2000. Plan 9 from cyberspace: the implications of the Internet for personality and social psychology. *Personal. Soc. Psychol. Rev.* 4(1):57–75

McNee SM, Riedl J, Konstan JA. 2006. Being accurate is not enough: how accuracy metrics have hurt recommender systems. In *CHI'06 Ext. Abstr. Hum. Factors Comput. Syst.*, pp. 1097–101. New York: ACM

Mehl MR, Conner TS, eds. 2012. *Handbook of Research Methods for Studying Daily Life*. New York: Guilford

Mendoza M, Poblete B, Castillo C. 2010. Twitter under crisis: Can we trust what we RT? *Proc. 1st Workshop Soc. Media Anal. (SOMA '10)*, pp. 71–79. New York: ACM

Miller GF. 2012. The smartphone psychology manifesto. *Perspect. Psychol. Sci.* 7:221–37

Miluzzo E, Lane ND, Eisenman SB, Campbell AT. 2007. CenceMe—injecting sensing presence into social networking applications. *Proc. 2nd Eur. Conf. Smart Sensing Context* (*EuroSSC 2007*). Lake District, UK

Morahan-Martin J, Schumacher P. 2003. Loneliness and social uses of the Internet. *Comput. Hum. Behav.* 19(6):659–71

Oppenheimer DM, Meyvis T, Davidenko N. 2009. Instructional manipulation checks: detecting satisficing to increase statistical power. *J. Exp. Soc. Psychol.* 45:867–72

Paolacci G, Chandler J, Ipeirotis PG. 2010. Running experiments on Amazon Mechanical Turk. *Judgm. Decis. Mak.* 5:411–19

Pariser E. 2011. *The Filter Bubble: What the Internet is Hiding From You*. New York: Penguin

Pennebaker JW, Mehl MR, Niederhoffer KG. 2003. Psychological aspects of natural language use: our words, our selves. *Annu. Rev. Psychol.* 54: 547–77

Pennebaker JW, Sussman JR. 1988. Disclosures of traumas and psychosomatic processes. *Soc. Sci. Med.* 26:327–32

Postmes T, Spears R. 1998. Deindividuation and antinormative behavior: a meta-analysis. *Psychol. Bull.* 123(3):238–59

Pury CLS. 2011. Automation can lead to confounds in text analysis: Back, Kufner, and Egloff (2010) and the not-so-angry Americans. *Psychol. Sci.* 22:835–36

Putnam RD. 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster

Qiu L, Lin H, Ramsay J, Yang F. 2012. You are what you tweet: personality expression and perception on Twitter. *J. Res. Personal.* 46:710–18

Rachuri KK, Musolesi M, Mascolo C, Rentfrow PJ, Longworth C, Aucinas A. 2010. EmotionSense: a mobile phones based adaptive platform for experimental social psychology research. *Proc. 12th ACM Int. Conf. Ubiquitous Comput.* (*UbiComp'10*), *Copenhagen*, *Den.*, pp. 281–90

Rains SA. 2014. The implications of stigma and anonymity for self-disclosure in health blogs. *Health Commun.* 29:23–31

Reips U-D. 2010. Design and formatting in Internet-based research. See Gosling & Johnson 2010, pp. 29–43

Reis HT, Gosling SD. 2010. Social psychological methods outside the laboratory. In *Handbook of Social Psychology*, ed. ST Fiske, DT Gilbert, G Lindzey, 1:82–114. New York: Wiley. 5th ed.

Rentfrow PJ, Gosling SD. 2012. Using smartphones as mobile sensing devices: a practical guide for psychologists to current and potential capabilities. *Preconf. Annu. Meet. Soc. Personal. Soc. Psychol.*, *San Diego*, *CA*

Rentfrow PJ, Gosling SD, Potter J. 2008. A theory of the emergence, persistence, and expression of geographic variation in psychological characteristics. *Perspect. Psychol. Sci.* 3:339–69

Rich F. 2011. Wallflowers at the revolution. *New York Times*, Feb. 5. **http://www.nytimes.com/2011/02/06/opinion/06rich.html**

Ritter RS, Preston JL, Hernandez I. 2014. Happy Tweets: Christians are happier, more socially connected, and less analytical than atheists on Twitter. *Soc. Psychol. Personal. Sci.* 5:243–49

Rogers EM. 1962. *Diffusion of Innovations*. New York: Free Press

Rosnow RL. 1980. Psychology of rumor reconsidered. *Psychol. Bull.* 87(3):578–91

Scherer K. 1997. College life on-line. Healthy and unhealthy Internet use. *J. Coll. Stud. Dev.* 38:655–65

Sears DO. 1986. College sophomores in the lab: influences of a narrow data base on social psychology's view of human nature. *J. Personal. Soc. Psychol.* 51:515–30

Skitka LJ, Sargis EG. 2006. The Internet as psychological laboratory. *Annu. Rev. Psychol.* 57:529–55

Smith PK, Mahdavi J, Carvalho M, Fisher S, Russell S, Tippett N. 2008. Cyberbullying: its nature and impact in secondary school pupils. *J. Child Psychol. Psychiatry* 49(4):376–85

Sourander A, Klomek AB, Ikonen M, Lindroos J, Luntamo T, et al. 2010. Psychosocial risk factors associated with cyberbullying among adolescents: a population-based study. *Arch. Gen. Psychiatry* 67(7):720–28

Stillwell DJ, Kosinski M. 2012. myPersonality project: use of online social networks for large-scale social research. *1st ACM Workshop Mobile Syst. Comput. Soc. Sci.* (*MobiSys*), *Univ. Cambridge*, *UK*

Tanaka Y, Sakamoto Y, Matsuka T. 2012. Transmission of rumor and criticism in Twitter after the Great Japan Earthquake. *Annu. Meet. Cogn. Sci. Soc.*, pp. 2387–92

Tausczik YR, Chung CK, Pennebaker JW. 2014. Tracking secret-keeping in emails. Manuscript under review

Thurman N, Schifferes S. 2012. The future of personalization at news websites. *J. Stud.* 13(5–6):775–90

Tokunaga RS. 2010. Following you home from school: a critical review and synthesis of research on cyberbullying victimization. *Comput. Hum. Behav.* 26(3):277–87

Tufekci Z, Wilson C. 2012. Social media and the decision to participate in political protest: observations from Tahrir Square. *J. Commun.* 62(2):363–79

Turkle S. 2012. *Alone Together: Why We Expect More from Technology and Less from Each Other*. New York: Basic Books

Valenzuela S, Arriagada A, Scherman A. 2012. The social media basis of youth protest behavior: the case of Chile. *J. Commun.* 62(2):299–314

Vambheim SM, Wangberg SC, Johnsen JK, Wynn R. 2013. Language use in an Internet support group for smoking cessation: development of sense of community. *Inform. Health Soc. Care* 38(1):67–78

Van Dijk J, Hacker K. 2013. The digital divide as a complex and dynamic phenomenon. *Inf. Soc.* 19(4):315–26

Waggoner AS, Smith ER, Collins EC. 2009. Person perception by active versus passive perceivers. *J. Exp. Soc. Psychol.* 45:1028–31

Weinberg JD, Freese J, McElhattan D. 2014. Comparing data characteristics and results of an online factorial survey between a population-based and a crowdsource-recruited sample. *Sociol. Sci.* 1:292–310

Weisbuch M, Ivcevic Z, Ambady N. 2009. On being liked on the web and in the "real world": consistency in first impressions across personal webpages and spontaneous behavior. *J. Exp. Soc. Psychol.* 45:573–76

Wilson RE, Gosling SD, Graham LT. 2012. A review of Facebook research in the social sciences. *Perspect. Psychol. Sci.* 7:203–20

Wolf M, Theis F, Kordy H. 2013. Language use in eating disorder blogs: psychological implications of social online activity. *J. Lang. Soc. Psychol.* 32:212–26

Wrzus C, Brandmaier AM, von Oertzen T, Müller V, Wagner GG, Riediger M. 2012. A new approach for assessing sleep duration and postures from ambulatory accelerometry. *PLOS ONE* 7:e48089

Yarkoni T. 2010. Personality in 100,000 words: a large-scale analysis of personality and word use among bloggers. *J. Res. Personal.* 44:363–73

Yarkoni T. 2012. Psychoinformatics: new horizons at the interface of the psychological and computing sciences. *Curr. Dir. Psychol. Sci.* 21:391–97

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. 2011. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* 8:665–70

York JC. 2011. *Not Twitter, not WikiLeaks: a human revolution* [Web log post]. **http://jilliancyork.com/2011/01/14/not-twitter-not-wikileaks-a-human-revolution/**

Young KS. 1998. Internet addiction: the emergence of a new clinical disorder. *CyberPsychol. Behav.* 1(3):237–44