Chapter 13
Applications of Internet Methods in Psychology

Lee-Xieng Yang

13.1 Introduction

The birth of the Internet has greatly changed our lives. The search engine on the Internet (e.g., *Google* and *Wikipedia*) is replacing traditional libraries; the digital shopping platform (e.g., Amazon) is replacing traditional department stores; the social media (e.g., Facebook, Twitter, and LINE) is replacing traditional paper media. Most of our needs, which in the past could be fulfilled only via specific channels, now can be done via the Internet. Thus, no one would disagree that the Internet is not only a digital device to us, but also a part of our real live. Psychologists have noticed the potential benefit of the Internet to psychological research in at least three respects: as a data-collecting platform (see Harlow and Oswald 2016), as an online database, and as the research field to uncover the mystery of human mind and behavior. The first is specifically referred to as the Web-based psychological studies. Whether or not, or to what extent, the psychological research is suitably conducted online is the first concern for transferring the laboratory-based study to Web-based study. In Sect. 13.2, I review the studies addressing this issue. In addition to behavioral studies (e.g., experiments and surveys), the database approach study is not rare in psychology. However, in the past, the kinds and numbers of databases were limited to the institutes for maintaining them. Now, with the search engine on the Internet, researchers can gain access to all Web pages of their interest around the world, which provide a bigger than ever database for research. In Sect. 13.3, the studies with the Web search engines are introduced. Since the emergence of blogs, microblogs, and social network sites, people are more and more used to sharing their live events with friends on social media. Thus, the

L.-X. Yang (✉)
Department of Psychology, National Chengchi University, Taipei, Taiwan

S.-H. Chen (ed.), *Big Data in Computational Social Science and Humanities*, Computational Social Sciences, https://doi.org/10.1007/978-3-319-95465-3_13

social media themselves can become a research field for psychologists to observe and understand people. In Sect. 13.4, some pioneer studies in this approach are introduced. Following these three sections is the conclusion of this chapter.

13.2 Online Psychological Studies

As a science, one important principle of psychological research methods is that a study should be replicable. Due to the sampling error and the bigger variance of psychological phenomenon, this is a goal not that easy to achieve for psychology studies. Nonetheless, at least the psychological studies are conventionally asked to be conducted with standardized instruction, stimuli, and procedure and in well-controlled environment. However, the limited size and representativeness of sample (e.g., most participants are college students) are always the barrier to the establishment of ecological validity. When the idea of crowdsourcing emerges, on the Internet, one single task (or experiment) can be assigned to more than hundreds of workers (or participants) all over the world. This seems to be a solution to the low ecological validity of psychological studies. However, the control for the testing environment might not be as standardized as usual. The past studies suggest that the online survey indeed provides data of higher representativeness. For the online experiments, the effect size may be lower than the experiment in laboratory.

13.2.1 Crowdsourcing and Psychological Study

The emergence of the Internet has brought innovations in many aspects of our life. One of them is so called the concept of crowdsourcing first proposed by Jeff Howe (2006).

> *Crowdsourcing is the act of taking a task traditionally performed by a designated agent (such as an employee or a contractor) and outsourcing it by making an open call to an undefined but large group of people.* (Howe 2008, p. 1)

It is not hard to see in this definition that the characteristic of crowdsourcing is an undefined large group of people who work independently to complete a task. The high speed and the security of exchanging information, ideas, and data files on the Internet render the plausibility of crowdsourcing. Although it may look strange to traditional business to have unrelated people independently contribute to a same project, it may be simply nature in psychological studies. Suppose we would like to verify the hypothesis that a new teaching program can effectively enhance students' performance in mathematics. To this end, we might have two groups of students, one as the control group taught with the old way and the other as the experimental group taught with the new way. After being taught for a certain time period, both groups of students will get a test and the comparison between the average scores of these two groups can help us verify our hypothesis.

In order to prevent any confounding effect from the characteristics of participants (e.g., the control group are all males and the experimental group are all females), we will randomly sample the students from the student pool and randomly assign them to one group or the other. In the view point of crowdsourcing, the experiment is just the project or the task to accomplish. As the students are ideally to be independent to each other[1], the participants in an experiment can be viewed as the undefined group of people in crowdsourcing. Also, it is acknowledged, in order to get the power of statistical analysis large enough, the more participants the better. According to the definition of crowdsourcing, the group size must not be small. Thus, running a psychological experiment is quite similar to doing a project in the way of crowdsourcing, except that the former is traditionally conducted in the laboratories and the latter on the Internet.

In fact, psychologists have begun to collect data online. For instance, on this website (the URL is https://www.socialpsychology.org/expts.htm) maintained by Scott Plous, Wesleyan University, you can find more than 500 online experiments, surveys, and other social psychology studies. By clicking on the title of the study in which you are interested, you will be directed to the instruction Web page of that study and the test will begin after you sign up the informed consent by clicking the button of "I agree." In addition to the website maintained personally, some enterprise also provides similar online platform for the need of crowdsourcing. Amazon Mechanical Turk (MTurk) is a good example, which can be accessed on https://www.mturk.com/mturk/welcome. The slogan on this Web page is as follows.

> *Mechanical Turk is a marketplace to work. We give businesses and developers access to an on-demand, scalable workforce. Workers select from thousands of tasks and work whenever it's convenient.*

It is clear that MTurk is a marketplace where some people as the *requesters* can post their signup sheet/job advertisement to recruit workers to temporarily work for them, while some other people can come and join in the project as the *workers*. On MTurk, *requesters* can list tasks (called "human intelligence task," or HITs) along with a specified compensation. HITs range widely in size and nature, requiring from seconds to hours to complete, and compensation varies accordingly ($0.01 $0.1 per HIT). After seeing the preview of the task, *workers* can choose to accept this HIT and complete the task. Of course, *workers* will get paid according to their performance. This way of recruiting *workers* resembles quite much the way to recruit participants in a psychological study. Apparently, the spirit of crowdsourcing is embodied on MTurk.

[1] Random sampling of participants.

13.2.2 Internet Methods as a Medium for Collecting Human Data?

Since conducting a psychological study is a kind of crowdsourcing, it seems to be proper to do psychological studies on MTurk or similar websites. On the one hand, this idea might be workable in terms of the need to quickly get a large random sample of subjects. On the other hand, this might be a dangerous idea in terms of the concerns about the quality of collected data (e.g., Mezzacappa 2000). After all, psychological surveys and experiments both demand the administration of testing and experimental procedure to be standardized and well controlled, that presumably challenges the Web-based testing. Indeed, there is a debate on whether MTurk or similar websites can be a participant pool in the community of psychology. Let us first see some positive evidence.

Gosling et al. (2004) compared the data on personality measures of Internet samples ($N = 361{,}703$) who are the visitors to the noncommercial advertisement-free website, outofservice.com, and 510 traditional samples published in the *Journal of Personality and Social Psychology* with respect to several domains, such as gender, race, socioeconomic status, geographic region, age, and so on. Their findings suggested that Internet samples are more representative than traditional samples with respect to gender, socioeconomic status, geographic location, and age, and are about as representative as traditional samples with respect to race.

Also, Buhrmester et al. (2011) administered personality questionnaires via MTurk to evaluate the quality of data collected in this way. The main findings include that (1) MTurk participants are significantly more diverse than typical American college samples; (2) although the compensation rate and task length will affect participation, participants can still be recruited rapidly and inexpensively; (3) the compensation rates do not affect data quality; and (4) the data obtained are at least as reliable as those obtained via traditional methods. In addition to the administration of questionnaires, economic experiments can be run on MTurk. Amir et al. (2012) conducted economic game experiments on MTurk (dictator game, ultimatum game, trust game, and public goods game) and found that the results were consistent with previous research conducted in the physical laboratory. Similarly, Mason and Watts (2009) examined the relationship between financial incentives and performance by an MTurk experiment, in which participants were asked to sort the given images in chronological order in one of the conditions crossed by four levels of task difficulty and four levels of compensation. How many sets of images the participants sorted is the measure of quantity and their sorting accuracy is the measure of quality. The results showed that increased financial incentives increase the quantity, but not the quality of participants' performance. Another economic game experiment, public goods game, was found to yield a decreasing trend on subjects' contribution over rounds, which was consistent with the findings of the traditional lab-based experiments (Suri and Watts 2011). Even the Cloze sentence completion task on MTurk showed a high correlation (rho = .75) on word predictability to the lab experiment (Schnoebelen and Kuperman 2010).

13.2.3 MTurk as a Platform for Conducting Psychological Experiments

It might still not be sufficient with the above instances to resolve the worries about running psychological experiments online, as most of the psychological experiments demand a high precision (up to milliseconds) in reaction time recording and stimulus presentation, which is not covered by the above instances. However, the current Web browser technology (such as HTML with Java script) does afford millisecond timing functions, which theoretically can fit the need of psychologists. In fact, Crump et al. (2013) endorsed the validity of the cognitive experiments on MTurk via replicating a wide variety of classic cognitive experiments. These authors run on MTurk the Stroop task (MacLeod 1991; Stroop 1935) and found the same response pattern reported in the traditional experiments, that the reaction time of naming a word's color is faster when the word's name is congruent with its color than when the word name and word color are incongruent. Similarly, these authors also replicated on MTurk the cost on switching tasks (see Jersild 1927; Monsell 2003), that is, that people spend a longer reaction time when the to-be-done task is changed to another one. Again, these authors even showed that the Flanker task[2] (Eriksen 1995; Eriksen and Eriksen 1974) and the Simon task[3] (Craft and Simon 1970; Lu and Proctor 1995) were of no problem to run on MTurk.

Crump et al. (2013) further demonstrated that some effects found with rapid stimulus presentation can be replicated on MTurk, such as the visual cuing effect that target is identified more quickly with a correct cue only when the cue-to-target time interval is a small, attentional blink that the detection of the second target will be impaired when the second target appears within 100 500 ms of the first one (Raymond et al. 1992; Shapiro and Raymond 1997), and the marked priming effect that the detection of the probe will be enhanced if the prime is the same as it and vice versa. However, when the task becomes complex, more caution is needed for using MTurk. Crump et al. (2013) tried to replicate the difficulty gradient of learning the six category structures first proposed by Shepard et al. (1961). The stimulus consisted of three dimensions: shape, color, and size as shown in Fig. 13.1.

The Type I problem is the easiest to learn, as the two categories (circles and squares) can be perfectly separated by attending to shape only. The Type II is the second easiest, followed by Type III, Type IV, Type V, and Type VI. The Type VI problem is the hardest, as the categorization rule nonlinear and consists of all three dimensions. This difficulty gradient had been replicated by many studies (e.g., Kruschke 1992; Lewandowsky 2011; Love 2002; Nosofsky et al. 1994). Although in the data of Crump et al. (2013) online experiment, the Type I and Type VI were the easiest and the hardies problems, the error curves for other types (specifically Type

[2]The reaction time of identifying the target alphabet in a line of alphabets becomes longer when the target is different from the others and vice versa.

[3]The reaction time is shorter when the spatial compatibility between stimulus and response key is held and vice versa.

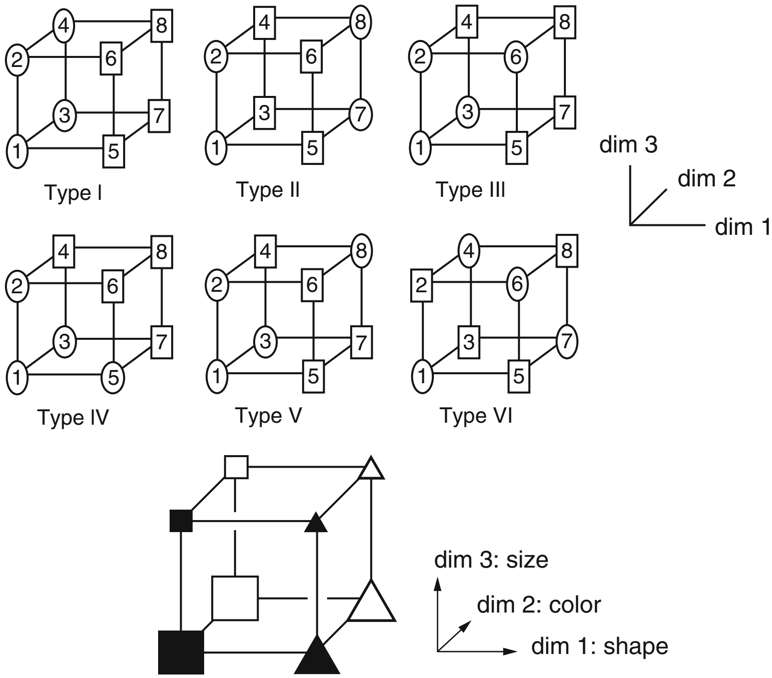

Fig. 13.1 Six types of category learning problems and the stimulus dimensions used in Shepard et al. (1961), quoted from Nosofsk et al. (1994)

II and Type IV) did not match the pattern reported in the precedent studies. These authors initially suspected that the inconsistency came from weak motivation, so they manipulated the incentive level (i.e., raising the payment up to $2.00 and an extra bonus up to $2.50 based on task performance), but the learning patterns for Type II and Type IV were still inconsistent with the precedent studies. As argued by Kurtz et al. (2012) that the Type II advantage can be explained by the extent to which instructions emphasize verbal rules, the observed inconsistency might result from the understanding direction of participants about the instruction. To sum up, when the task is simple (i.e., short in time or simple in experimental design), such as those reaction time or attention and perception tasks, there is no problem to use MTurk; when the task is complex (i.e., complex in experimental design or sensitive to the understanding of instruction), it needs caution to conduct the experiment on MTurk. Nonetheless, the payment does influence the rate of joining in the experiment, but not the quality of performance.

13.2.4 Second Thoughts for MTurk as a Participant Pool

Although the above reviewed studies point up the positive aspect of Web-based studies, some researchers otherwise have warned us of the potential crises in this approach. Among those warnings, sample quality and data quality are the main concerns. For the sample quality, the participants on MTurk often engaged in other affairs (e.g., reading MTurk blogs, listening to music, watching TV, or even chatting online) while doing a HIT (see Chandler et al. 2014). Thus, we do have reasons to doubt that the MTurk experiment is equivalent to those in laboratories. Also, cross talk happens among the participants on MTurk as *workers* can read the MTurk blogs or relevant forums to exchange information about HITs. These discussions are more about the pay rates or the reputation of *requesters*, not too much about the content of a particular HIT. Perhaps the duplication of participants across tasks is the most serious part of sample quality on MTurk. Normally, the researchers assume that the participants are naïve to the research materials because they are from a large participant pool or they have limited exposure to the research. Chandler et al. (2014) showed that some people have been the *workers* on MTurk for several years and are more likely than others to be sampled. Past research has noted that response to psychological measures correlate with proxies of prior participation in similar experiments, such as memory of prior participation (Greenwald and Nosek 2001) and memory of chronological order of studies themselves (Rand et al. 2014). Worse, the non-naïveté of participants can reduce the effect size of experimental treatment (Chandler et al. 2015).

The major concern about data quality resides on the prevalence of data exclusion. Chandler et al. (2014) did a meta-analysis for hundreds of papers published prior to December, 31, 2011, which conducted MTurk experiments and estimated that about one-third of these papers dropped *workers* post hoc for one reason or another. This high ratio is indeed a worry to those who want to conduct an MTurk experiment.

Although the issue of duplication of participants is annoying, it is not impossible to deal with. Chandler et al. (2014) suggested that Amazon Qualifications can be used as a tool for prescreening *workers*. *Requesters* can set up criteria in Qualifications for filtering *workers*, such as female only or the maximum number of HITs before the current task. Accordingly, we can select participants prior to the execution of task. Also, in virtue of a better control for the source of sample provided by Qualifications, it is likely to lower the exclusion rate of data. Thus, whether or not MTurk can be used for conducting psychological studies is not a simple yes/no question.

The Web-based (or MTurk) research indeed has some advantages that are better than the traditional studies, such as that it makes inexpensive and fast the recruitment of participants as well as it provides a relatively diverse sample. However, there is no such thing as a free lunch. It is worth noting that the validity of Web-based experiments might be reduced, due to the duplication of participants or/and inappropriate trimming of data post hoc. Nonetheless, as long as we consider the nature of our task and do necessary prevention in advance (e.g., setting up Qualifications), this kind of research is still a viable choice for psychologists.

13.3 Psychological Studies with Online Search Engine

In addition to collecting data from real human beings, psychological studies can also be done with the established databases. For instance, the psycholinguistic study often relies on the corpus, which provides information about vocabulary, such as word frequency and word types. However, maintaining a database costs a lot of time and money. Also, it is not guaranteed to have a database matching our research interest. Thus, the feasibility of database approach research is constrained. Since the Internet is composed of billions of Web pages and each of which can be treated as a source of data, a straightforward idea is why not treat the whole Internet as the gigantic database for research. There are two types of studies embodying this idea. First, the search engine is used to provide data from Web pages. Second, the database of the search engine is the target for research. Both types of studies unveil a new landmark of database approach research in psychology.

13.3.1 Search Engine on the Internet as Research Tool

In social sciences, database research has long been a typical way to provide a large-scale description for the topic of interest. Different databases contain data for different topics, such that the database of the World Bank contains data about economic indices of nations all over the world or that WordNet contains English nouns, verbs, adjectives, and adverbs which are structurally grouped as a useful tool for computational linguistics and natural language processing. A database is normally established and maintained by governments, academic institutions, or enterprises for public use. Most of the databases are established for a specific topic. For instance, WordNet is specifically established as a corpus of English. If we want to get the statistics of the number of vocabularies an English speaker would have, WordNet is a suitable choice. However, if we want to know GDPs of all nations in the world, we definitely will not go with WordNet. In addition to the specificity, sometimes a database can be accessed only by those who have been approved. It is noted that for the consideration of information security, some databases can be accessed only by the approved users.

In addition to the databases established for specific purposes, the search engine on the Internet, such as *Google* and Wikipedia, might also be a research tool, just like the database. However, with a number of essential differences, the search engine is not simply a database. First, the data contained in a regular database are structuralized, but the data in the Web pages, which can be accessed by the search engine, are not. For instance, the Web pages found by *Google* for our needs are often metadata which cannot be used directly until being sorted out by suitable processes. Although the Wikipedia may be more structuralized with clear catalogs and indices, what the users can get on it basically are the Web pages and the contents of those Web pages are the source of data, not just the data.

Second, the search engine, especially *Google*, can provide a far wider variety of data than the regular database. Unlike the regular database which is established by a group of specialists following the coding instructions, the Web pages can be created by anyone all over the world. Also, there is no constraint on the format (e.g., spread sheet or a plain text), type (e.g., text or non-text), or length of the content of a Web page. Thus, the data transferred from the Web pages can be far more divergent than those structuralized data in a normal database. Third, the search engines can provide us not only the Web pages meeting our criteria but also the information about other people's search for those Web pages. For instance, the Web pages on the first page of the search results with *Google* are more likely searched by other people also. The normal database cannot provide such information. Some researchers in psychology and social sciences have started to apply *Google* and Wikipedia to research and gained quite inspiring results.

13.3.2 Research with Google and Wikipedia

Stewart et al. (2005) proposed a simple model, the decision-by-sampling (DbS) model, to account for why the descriptive psychoeconomic functions take the form that they do. The idea of the DbS model is that people's decision for an attribute value depends on the relative rank of it in a random sample from their memory. These authors assumed that the contents of memory reflect the structure of the world. They also showed that the distribution of time in our memory can be gained by googling different strings which represent different time periods (e.g., 1 day, 2 days, ..., 1 year) and accumulating the counts of articles searched by *Google*. This distribution as other utility functions in Economics follows a power function. Similarly, Olivola and Sagara (2009) asked participants to estimate the frequencies of natural and industrial disasters given different number of deaths. The aggregated data show a function following a power function. These authors also searched with *Google News Archives* for the given numbers of death tolls (e.g., "10 people died"). The results showed that the distribution of fatalities from natural and industrial disasters made from *Google* search is quite similar to the one of real records and the participants' estimation followed a similar power function. These studies supported the idea that the result of *Google* search can reflect the structure of the real world and represent the contents of our memory.

Psychologists have long acknowledged that human beings are actually not as rational as assumed in economic theories on making their decision and instead we rely on many types of heuristic (Tversky and Kahneman 1974) to help us make decisions. These heuristics quite often lead us to fallacy in making our decisions. One of the famous fallacies is called the conjunction fallacy (Tversky and

Kahneman 1983). The typical problem is: "Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations." The decision-makers are then asked whether she is more likely to be a bank teller or a feminist bank teller. Since the event of "bank teller" covers the event of "bank teller *AND* feminist," the probability for Linda being a bank teller should be higher than being a banker teller and feminist. However, people tend to choose the latter as their answer. Bhatia (2015) developed a heuristic judgment model which is a neural network model learning approximately 3.2 million articles on Wikipedia to recover the co-occurrence structure of 300,000 words. If two words appear in one article, then the frequency of co-occurrence for them is 1. Based on the co-occurrence structure implicitly retained in the articles on Wikipedia, this model predicts 67% of this conjunction fallacy and this result is positively correlated with another study with human participants ($r = .29$, Shafir et al. 1990).

More surprisingly, the search engine data can be used as a predictor of human action. Preis et al. (2013) found that words were classified as more financially relevant were more strongly related to subsequent stock market moves. Moat et al. (2013) even showed that increases in views of Wikipedia pages relating to companies listed in the Dow Jones Industrial Average or to more general economic concepts, tended to be followed by stock market drops.

13.4 Psychological Studies with Social Media

The emergence of social media extends our real (or offline) life to the virtual (or online) networks. Although there is not too much research so far, researchers have started to investigate the influence on people brought from social media. One interesting issue is whether the algorithm of social media to order the posts would manipulate people's attitude toward particular issue. The current evidence seems to support that the polarization of people's attitude does not result from the algorithm but people themselves. However, people's mood is indeed easily influenced by their friends even via their posts on social media. In addition to social influence, how people's footprints on social media can reflect their psychological construct is another interesting topic. Some research has started this quest and gained interesting findings. Although these cases are relevant to social and personality psychology, it can be reasonably expected that more studies with social media will be proposed in many other areas of psychology.

13.4.1 Social Influence on Individuals in Social Media

It is acknowledged that we are influenced by important others in our social circles or the media and it becomes more salient on social media, such as Facebook. Some researchers suspected that Facebook's social endorsement algorithm has something to do with the polarization of opinions and selective exposure on the media (Messing and Westwood 2012). However, some others defensed for the Facebook algorithm and showed that individuals' choices play a bigger role than the algorithm on selecting news (Bakshy et al. 2015). Nonetheless, it is of no doubt that we are living in a world without any influence from others, no matter whether the social endorsement algorithm exacerbates the fragmentation of the citizenry.

Another interesting case about the influence through the social media is the study of Coviello et al. (2014) on emotional contagion. These authors identified the emotional expression of millions of posts of Facebook users in terms of the words used in the posts. The regression results showed that the individual-specific factor (e.g., some people are always happier than others), the exogenous factor (e.g., rainfall), and the influence from the user's Facebook friends are valid predictors to the emotional expression of post. Generally speaking, when it rains, we feel sad. What these authors found gives us a surprising story, that even though there is no rain in my town, I could express negative emotion if my friend feels sad due to the rain in his/her town. The emotion contagion through social media is one instance of how psychologists can investigate individuals in a social context via social media. Of course, other topics in social psychology (e.g., group thinking) should be able to study via social media.

13.4.2 Private Traits are Predictable from Digital Records

Psychologists always seek the valid criteria of the psychological traits or constructs. These criteria can be the items in a personality test, which can be a sentence or an adjective for people to judge to what extent the description matches their situation. For instance, a sentence that I like to go out with friend might be a criterion for extraversion.

Kosinski et al. (2013) investigated whether the digital records of human behavior can be used to estimate the personal attributes. These authors used the participants' Facebook likes as digital record of their behavior on Facebook. Also, these participants' scores on the famous Big-5 personality test were collected, using the questionnaire in the International Personality Item Pool (IPIP). These authors first constructed a matrix to represent the users and their likes. In order reduce the dimensionality of this matrix, the linear algebra technique SVD (Singular Value Decomposition) was applied to transfer the user-like matrix to user-component matrix. Subsequently, with this user-component matrix as a predictive variable, logistic regression model or linear multiple regression was established to predict the users' psychodemographic profiles. It was shown that their model can correctly discriminate between homosexual and heterosexual men, African Americans and Caucasian Americans, and between Democrats and Republicans. For the personality trait "Openness," the prediction accuracy is close to the test–retest accuracy of a standard personality test. Therefore, it is supported that the digital records on social media can predict some personal traits that people would typically assume to be private.

13.5 Conclusion

There are basically three approaches psychologist would take to apply the Internet methods to their study. The first is using the Web environment as a new platform for recruiting participants and conducting studies. The advantage of this approach is that we can normally get the participants quickly and inexpensively. However, duplication of participants and high exclusion rate of data post hoc would be the cost. Whether or not a study is suitable for the Web environment depends on the nature of it (the type of dependent variable, prior knowledge interference, etc.). Instead of behavioral studies, the second approach goes for the search engine in the Internet (e.g., *Google* and Wikipedia) and can be regarded as a database study. With the gigantic amount of materials on the Internet as the representation of our knowledge about the world, it becomes possible to examine the mental representation and processing (memory, categorization, decision-making, etc.) at the population level. Different from the traditional experimental design, this approach requires some Internet techniques for dealing with the texts on Web pages (html coding, Web crawler, text mining, latent semantic analysis, etc.). Finally, the third approach provides a possibility to examine individual actions in social network. Comparing with the first two, this approach is more ambitious and requires more considerations on with respect to the procedure of data collection and the analysis and interpretation of data. Specifically, the related ethical issue about collecting the data of users on social media needs thorough considerations. Nonetheless, the stand of this chapter is that the Internet methods, just like any other research techniques (e.g., EEG or fMRI), can contribute to our studies of human beings, as long as the research topic, target behavior, and testing conditions are appropriate.

References

Amir, O., Rand, D. G., & Gal, Y. (2012). Economic games on the Internet: The effect of $1 stakes. *PLoS ONE, 7*(2), e31461. https://doi.org/10.1371/journal.pone.0031461.

Bakshy, E., Messing, S., & Adamic, L. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 1–5.

Bhatia, S. (2015). The power of the representativeness heuristic. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 232–237). Austin, TX: Cognitive Science Society.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*, 3–5.

Chandler, J., Mueller, P., & Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior Research Methods, 46*, 112–130. https://doi.org/10.3758/s13428-013-0365-7.

Chandler, J., Paolacci, G., Peer, E., Mueller, P., & Ratliff, K. (2015). Non-naïve participants can reduce effect sizes. In K. Diehl & C. Yoon (Eds.), *NA—advances in consumer research* (Vol. 43, pp. 18–22). Duluth, MN: Association for Consumer Research.

Coviello, L., Sohn, Y., Kramer, A. D. I., Marlow, C., Franceschetti, M., Christakis, N. A., et al. (2014). Detecting emotional contagion in massive social networks. *PLOS ONE, 9*, e90315. https://doi.org/10.1371/journal.pone.0090315.

Craft, J. L., & Simon, J. R. (1970). Processing symbolic information from a visual display: Interference from an irrelevant directional cue. *Journal of Experimental Psychology, 83*, 415–420.

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE, 8*(3), e57410. https://doi.org/10.1371/journal.pone.0057410.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*, 143–149.

Eriksen, C. W. (1995). The flankers task and response competition: A useful tool for investigating a variety of cognitive problems. *Visual Cognition, 2*, 101–118.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. *American Psychologist, 59*, 93–104.

Greenwald, A. G., & Nosek, B. A. (2001). Health of the implicit association test at age 3. *Zeitschrift für Experimentelle Psychologie, 48*(2), 85–93.

Harlow, L. L., & Oswald, F. L. (2016). Big data in psychology: Introduction to the special issue. *Psychological Methods, 21*, 447–457.

Howe, J. (2006). The rise of crowdsourcing. *Wired Magazine, 14*(06), 1–5.

Howe, J. (2008). *Crowdsourcing: Why the power of crowd is driving the future of business*. New York: Crown Publishing Group.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*, *14*, (Whole No. 89).

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *PNAS, 110*, 5802–5805.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99*, 22–44.

Kurtz, K. J., Levering, K. R., Stanton, R. D., Romero, J., & Morris, S. N. (2012). Human learning of elemental category structures: Revising the classic result of Shepard, Hovland, and Jenkins (1961). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Online first publication. doi:https://doi.org/10.1037/a0029178

Lewandowsky, S. (2011). Working memory capacity and categorization: Individual differences and modeling. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 720–738.

Love, B. C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic Bulletin & Review, 9*, 829–835.

Lu, C., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychological Bulletin & Review, 2*, 174–207.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*, 163–203.

Mason, W., & Watts, D. J. (2009). Financial incentives and the performance of crowds. In *Proceedings of the ACM SIGKDD Workshop on Human Computation* (pp. 77–85). New York: ACM.

Messing, S., & Westwood, S. J. (2012). Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. *Communication Research, 41*, 1042–1063. https://doi.org/10.1177/0093650212466406.

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

Moat, H. S., Curme, C., Avakian, A., Kenett, D. Y., Stanley, H. E., & Preis, T. (2013). Quantifying Wikipedia usage patterns before stock market moves. *Scientific Reports, 3*, 1801.

Monsell, S. (2003). Task switching. *Cognitive Science, 7*, 134–140.

Nosofsk, Y. R. M., Gluck, M. A., Palmeri, T. J., McKinley, S. C., & Glauthier, P. (1994). Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961). *Memory & Cognition, 22*, 352–369.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review, 101*, 53–79.

Olivola, C. Y., & Sagara, N. (2009). Distributions of observed death tolls govern sensitivity to human fatalities. *Proceedings of the National Academy of Sciences, 106*, 22151–22156.

Preis, T., Moat, H. S., & Stanley, H. E. (2013). Quantifying trading behavior in financial markets using Google Trends. *Scientific Reports, 3*, 1684.

Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., et al. (2014). Social Heuristics shape intuitive cooperation. *Nature Communications*.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink. *Journal of Experimental Psychology: Human Perception & Performance, 18*, 849–860.

Schnoebelen, T., & Kuperman, V. (2010). Using Amazon Mechanical Turk for linguistic research. *Psihologija, 43*, 441–464.

Shafir, E. B., Smith, E. E., & Osherson, D. N. (1990). Typicality and reasoning fallacies. *Memory & Cognition, 18*, 229–239.

Shapiro, K. L., & Raymond, J. E. (1997). The attentional blink. *Trends in Cognitive Science, 1*, 291–296.

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs, 75*(13), 1–42 (Whole No. 517).

Stewart, N., Chater, N., & Brown, G. D. A. (2005). Decision by sampling. *Cognitive Psychology, 53*, 1–26.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Suri, S., & Watts, D. J. (2011). Cooperation and contagion in web-based, networked public goods experiments. *PLoS ONE, 6*(3), e16836. https://doi.org/10.1371/journal.pone.0016836.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 141–162.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*, 293–315.