

CHAPTER 8

SOCIAL NETWORK EXPERIMENTS

MATTHEW E. BRASHEARS AND ERIC GLADSTONE

WHAT is the most effective way of determining causal processes involving social networks? By definition, social networks are collections of social entities (often individuals) who are linked by one or more relationships, including informal interpersonal bonds (e.g., friendship), formal economic transactions (e.g., employment), obligatory connections (e.g., family), and simple copresence or comembership, to name a few. Such assemblages of individuals or other social actors are often quite large and so are typically studied essentially “in the wild,” imposing significant difficulties in data collection. These challenges include identifying the boundaries of the network (if conducting a sociometric analysis; see Laumann, Marsden, & Prensky, 1992), achieving an adequate sample size (e.g., Hipp et al., 2015; McCarty, Killworth, & Rennell, 2007; Merluzzi & Burt, 2013), selecting the best network instrument (e.g., Brashears & Quintane, 2018; Fischer, 1982), and minimizing priming (e.g., Brashears, 2011; Lee & Bearman, 2017), respondent fatigue (Fischer, 2009), interviewer fatigue (Paik & Sanchagrin, 2013), and other sources of bias. Once the data are collected, the inherent dependence between observations (or between dyads in ego network samples; see Brashears, 2014) forces researchers to use complex replacements for conventional multivariate models (e.g., exponential random graph models; see Lusher, Koskinen, & Robins, 2013). Thus, complex data are met with complex methods in hopes of producing straightforward answers. But there is an alternative.

In this chapter we provide an overview of experimental designs and give several examples of how experimental methods can be employed in social network research. Experiments provide unsurpassed certainty in causal identification by allowing the researcher to control all aspects of the process while in many cases substantially reducing data collection costs. Moreover, with the aid of computers, the internet, and a little creativity, it is often possible to collect highly reliable data (see Kraut et al., 2003; Paolacci, Chandler, & Ipeirotis, 2010; Horton, Rand, & Zeckhauser, 2012) from online samples with the aid of crowdsourcing services, such as Amazon Mechanical Turk. In short, the time is right for network experiments.

AQ1

WHAT IS AN EXPERIMENT?

Most simply put, an experiment is the deliberate creation of a situation for the purpose of measuring the causal linkages between one or more factors (see also Jackson & Cox, 2013). In all experiments, a treatment is given to subjects to determine how their responses change pre- and posttreatment (i.e., a within-subjects design), different groups of subjects are given different treatments to see how their responses vary from each other (i.e., a between-subjects design), or both. Experiments can be further broken down by whether they are true experiments, quasi-experiments, or natural experiments, proceeding in that order from maximum researcher control over conditions to minimum. In a true experiment, the researcher controls all aspects of the experimental setting and randomly assigns participants to conditions. This type of experiment is typically executed in a laboratory where the physical space is entirely under the researcher's control, frequently including interactions between the subject and covert research assistants (i.e., "confederates"). Quasi-experiments involve the random assignment of subjects and administration of treatments in a setting not fully under researcher control (e.g., in a public area). Natural experiments are possible when natural events cause an experiment-like treatment to be applied to a population (e.g., enactment of new legislation or an earthquake), permitting pre- and posttreatment measures to be compared. Crucially, experiments of all types allow us to "rule in" causal mechanisms but never to "rule out" causal mechanisms. When a mechanism is confirmed experimentally, it is accurate to say that the mechanism is sufficient to produce the observed outcome but has not been shown to be *necessary*. That is, there may be other mechanisms that produce the same outcome.

Experimental designs simplify causal identification by synthetically producing matched counterfactual conditions. For example, in a true experiment, researchers design the conditions to differ only on the variable(s) of interest through the use of manipulations (i.e., deliberate changes to the subjects' experiences). They then assign subjects to each condition at random, ensuring that the distribution of unobserved characteristics (e.g., sexual preference, personality traits, etc.) in each condition is equal, within the limits of probability theory. As a result, the only systematic differences between the groups are those produced by the manipulations, and in effect, each condition represents a particular type of counterfactual relative to the others. For example, if we test a drug by giving one randomly selected group the actual medication and another randomly selected group a placebo (e.g., a sugar pill), then the group receiving the placebo is a counterfactual for the group receiving the drug. In one synthetic universe, individuals receive a drug that (hopefully) works, while in the other synthetic universe, they receive a drug that definitely does not work. Differences between groups can therefore be positively attributed to differences in the manipulations as all other characteristics, including the attributes of the subjects, are identical.

Causal identification becomes more challenging as we move from true experiments to quasi-experiments to natural experiments. As the experimental situation is less and less under the control of the researcher, it becomes more likely that outside factors that are unobserved and unmeasured can influence the outcome or moderate any causal relationship. For example, a sudden earthquake can provide the opportunity to examine changes in cortisol levels, but other factors that vary over time (e.g., political events, seasonal weather changes, etc.) are also likely to play a role, complicating efforts to specify a causal link.

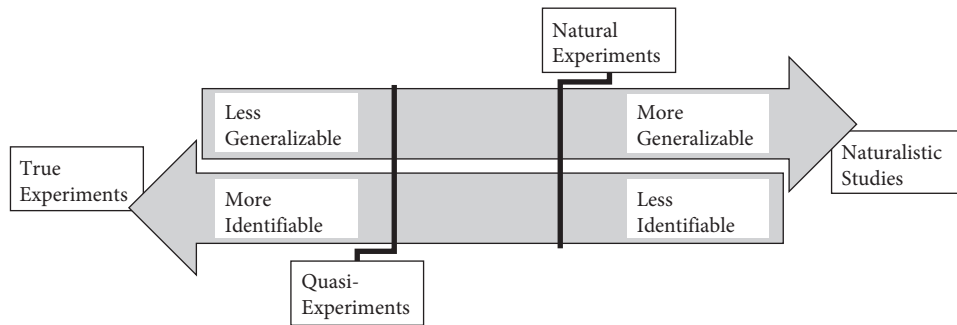


FIGURE 8.1 Ease of causal identification/generalizability by research type.

However, this decrease in ease of causal identification brings with it an improvement in generalizability. True experiments are most difficult to connect to “real life” events because the causal connection is observed only in strictly controlled circumstances. Quasi-experiments and natural experiments take place in less controlled and artificial environments and are therefore more straightforward to relate to “real life” conditions. For example, it may be possible to induce racial prejudice in the laboratory, but this does not demonstrate that such prejudice impacts actual employment decisions. A quasi-experimental design in which job applications, differing only in the race of the applicant, are sent to real employers to measure the frequency of call-backs is more easily generalized but suffers from more opportunities for confounds. The precise type of design selected thus has critical implications for causal identification and generalizability. In comparison, naturalistic studies of social networks (e.g., collecting network data from a selection of managers; see Burt, 1992) maximize generalizability while minimizing the potential for causal identification. As such, traditional studies of social networks are predominantly found at one extreme of the causal identification/generalizability dimension (see Figure 8.1) and are frequently subject to protracted debates over causal identification (e.g., *diffusion of obesity*: Christakis & Fowler, 2007; Cohen-Cole & Fletcher, 2008a, 2008b; Fowler & Christakis, 2008; *diffusion of antibiotic use*: Burt, 1980; Coleman et al., 1966; Van den Bulte & Lilien, 2001). Even when not subject to such debates, naturalistic studies of social networks often suffer from the fact that only a single network is examined. As a result, while many nodes may be contained in the data, there is only one “observation” at the network level. These studies essentially represent the worst of both worlds: too little researcher control to provide easy causal identification, and too few observations to provide generalizability.

CAN SOCIAL NETWORKS BE STUDIED WITH EXPERIMENTS?

Social networks are assemblages of parts, often including dozens to millions of distinct individuals with a potentially very large number of relationships between them. Moreover, most network analysts are of the opinion that networks have effects that are not perfectly reducible to the individual or accounted for with higher-level structural influences. Why



study a network if all of its behavior can be encapsulated in individuals, on the one hand, or large-scale social trends on the other? As a result, it may appear on the surface as though experimental designs are inappropriate for network research, despite their obvious advantages; a network cannot fit into a laboratory, and thus networks cannot be studied with experiments. Nonetheless, these objections can be overcome by doing one of two things: shrinking the process to fit into the laboratory or enlarging the laboratory to encompass the process.

Morris Zelditch (1969) argues for shrinking the process to fit into the laboratory when he asks, “Can you really study an army in the laboratory?” His answer is a resounding “yes” and rests on the point that the goal of most research is not to study a phenomenon *per se* (e.g., a network of managers), but rather to identify and test the underlying causal mechanisms that enable and empower a given phenomenon (e.g., the effect of redundancy on information movement). Given that this is the case, an experimental design is a valid approach so long as the causal mechanism can be brought into the laboratory, even if the context in which it is thought to operate cannot. For example, if status homophily is thought to produce segregated social networks, then an experiment need only induce an arbitrary status order in naïve subjects (using any of a variety of previously validated methods; see, e.g., Simpson & Walker, 2002) and observe their resulting communications behavior in the laboratory. If the mechanism is observed to work in these artificial conditions, then it is logical to assume that it will also operate in larger, more natural groups (e.g., in natural friendship networks; see Brashears, 2008). Mechanisms are smaller than contexts and thus can be made to fit into many laboratory settings quite handily.

Zelditch’s points remain as germane now as they were when he wrote them, but 40 years later, new solutions to the same problem have become available. Not only can we shrink a process to fit into the laboratory, as Zelditch suggests, but also we can enlarge the laboratory to encompass the process. The combination of the internet and ubiquitous connected devices including smartphones and wearables (see Eagle, Pentland, & Lazer, 2009) allows researchers to construct large, but strictly controlled, networks on demand. Connected devices (including but not limited to personal computers, tablets, and cellular telephones) can be used both to administer experimental treatments and as interfaces for subject interaction. The types of interaction can be as constrained or free as the researcher desires, ranging from something as simple as conventional game theory designs where participants divide a pool of points that are tied to their compensation (e.g., the ultimatum game: Rand et al., 2013) to something as complex as full-motion video with audio monitoring. Moreover, as interaction occurs via connected device, researchers can mislead subjects into believing they are interacting with another human when they are in fact communicating with a computer program (i.e., a simulated interactant), allowing highly precise and controlled administration of interactional stimuli. Moreover, thanks to the internet, interaction can take place between individuals drawn from widely separated geographic locations. The laboratory can thus be enlarged enough to enclose the planet, if such is desirable.

The final technological innovation that has revolutionized the potential for experiments in social network research is the rise of crowdsourcing platforms like Amazon Mechanical Turk.¹ Mechanical Turk, or MTurk, was developed as an online labor market for connecting individuals with time on their hands to groups who are willing to pay money to complete tasks that only humans are capable of (e.g., tagging photos with their contents). MTurk workers peruse a listing of jobs maintained in an online marketplace before selecting the

jobs that are most appealing, completing them, and receiving electronic payment. Both the employer and worker are able to rate each other, leveraging reputation effects to ensure good behavior, and workers can be selected on a variety of criteria, including their average ratings.² Experimentalists have already begun using MTurk as a source of data (e.g., Kraut et al., 2004); workers can be compensated for completing experiments as if they were any other type of job. This approach involves no sampling as usually defined and thus cannot be viewed as statistically representative. But as with laboratory studies, statistical representativeness is not usually a goal. Instead, differences in experimental conditions allow the identification of causal effects, and the more diverse pool of research subjects (i.e., not exclusively college students) reduces the risk that observed results are limited to a subset of the population. Comparisons with laboratory data (Buhrmester, Kwang, & Gosling, 2011; Gosling & Johnson, 2010) indicate that results derived from MTurk are more representative than, and at least as reliable as, traditional college student samples. Results from experiments conducted via MTurk are consistent with those collected in the lab, as well as those collected elsewhere on the internet such as survey websites, message boards, and online forums (Paolacci et al., 2010). Finally, research shows that MTurk samples compare favorably with those drawn using more traditional probability sampling methods (Weinberg, Freese, & McElhattan, 2014). Other options for research crowdsourcing continue to emerge, including Northeastern University's Volunteer Science Initiative³ and the Zooniverse Project,⁴ both of which serve as online sources of participants and human help for scientific research. Therefore, it has become possible to construct online networks and to recruit subjects cheaply and easily from a wider cross-section of society than ever before, somewhat escaping the tradeoff between causal identification and generalizability reflected in Figure 8.1.

In summary, social networks can be studied in the laboratory by focusing on the critical causal mechanisms and capturing them in reduced form. Alternatively, they can be examined experimentally outside the laboratory by using computers to present highly controlled stimuli to subjects and to manage their interactions. Finally, crowdsourcing platforms enable the rapid and inexpensive recruitment of a geographically and demographically diverse (i.e., less WEIRD; see Henrich, Heine, & Norenzayan, 2010) set of respondents. Thus, we have ample reason to believe that social networks can, and indeed *should*, be studied with experiments.

EXPERIMENTAL MANIPULATIONS FOR THE STUDY OF SOCIAL NETWORKS

Experiments of all types derive their power in part from the manipulations used. As the name suggests, manipulations are deliberate changes made by the researcher (or, in the case of natural experiments, the environment) to cast light on a hypothesized causal mechanism. It is effectively impossible to provide a list of the possible manipulations available to researchers as a considerable degree of creativity is frequently involved. However, four broad criteria should generally guide manipulation selection or design: theoretical consistency, fit with scope conditions, availability of validated designs, and manipulation restraint.

First, to provide valuable results, an experimental manipulation must be consistent with the theory it means to test or relies upon. Because experiments work by systematically changing some variable or set of variables thought to have a causal impact on some other variable or variables, they rely on a theoretical model of how those variables are connected. This is most obviously true for formal theory, where the model is specified in exacting (often mathematical; see, e.g., Willer, 1999 or Friedkin, 1998) detail, but it is also true for theories stated as textual rather than mathematical postulates. An appropriate manipulation is one that targets the mechanism specified by the theory and adjusts it in the required manner. However, it is important to keep in mind that even the most simple of data-driven research is, fundamentally, rooted in at least an implicit theory. For example, if researchers want to understand the connections between network structure and emotional well-being, they will at a minimum select some aspect or aspects of the network to connect to emotional outcomes. The specific structural features selected reflect an underlying, likely unstated, theory about how networks shape emotional responses. Thus, virtually any serious network study will be guided by some type of theoretical position. Experiments only differ in the extent to which this is true; while a conventional network study might be purely exploratory, involving minimal theory, a network experiment must have identified a specific set of mechanisms, shaped by theory, that are to be reproduced and manipulated in controlled conditions.

Second, all theories have boundaries that define the set of conditions under which they apply. These boundaries often indicate the types of entities the theory applies to (e.g., humans) as well as the types of situations it is relevant for (e.g., task groups). Such boundaries are known more formally as “scope conditions” or “boundary conditions,” and an experiment that fits within a theory’s scope conditions can serve as a test of that theory, while an experiment that exceeds them cannot. For example, theories of network diffusion cascades generally assume that an item can spread from person to person like a biological contagion. As such, an experiment tracking the movement of an item that cannot be duplicated (e.g., money; see Schaefer, 2007) would violate the scope conditions for network diffusion cascades and not provide an effective test. An unstated but often necessary scope condition for experiments in general is that participants are engaged with and paying attention to the experiment itself. As a result, many experiments use what are called attention checks or manipulation checks (Keppel, 1991) to confirm that the subjects are engaged and that the manipulation is working (though not necessarily that the manipulation is having the anticipated effect). Scope conditions can be relaxed by subsequent theoretical development and empirical research, but at any given moment, a manipulation based in, or intended to test, a given theory must be consistent with its scope conditions.

Third, while substantial creativity is often involved in the development and deployment of experimental manipulations, it is generally desirable to use manipulations already developed by other researchers, when possible. As with survey items, the behavior of previously used manipulations is more likely to be known and thus the results that they produce can be more easily understood. Likewise, potential pitfalls in the design (e.g., demand characteristics) are more likely to be known, and thus possible to adjust for. To date (as we discuss later), experiments have presented networks to participants in textual form (e.g., vignettes), have shown them sociograms, and have depicted interactions between computer-generated avatars, among other approaches. These methods each have their own advantages, but there

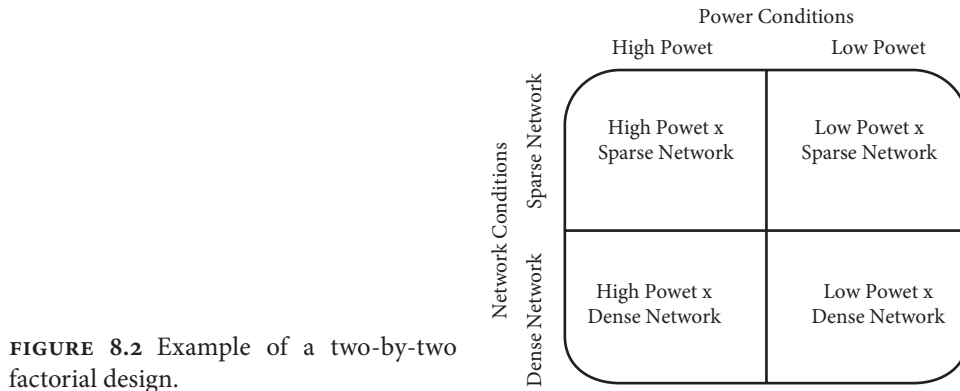


FIGURE 8.2 Example of a two-by-two factorial design.

is little consensus on which is best for a given purpose or what the potential biases may be. Thus, considerable work remains to be done.

Finally, manipulations must be applied with a degree of restraint. This doesn't mean that manipulations must be designed to be small in effect, but rather that they should be applied cautiously to maximize causal identification. For example, an experiment with two manipulations with two levels each (e.g., high vs. low power priming and dense vs. sparse network) can be performed with a simple two-by-two crossing.

This results in four conditions, each reflecting a combination of manipulations (e.g., high/dense, high/sparse, low/dense, low/sparse). Critically, however, the experimental conditions form pairs, between which only one factor has changed (e.g., high/dense vs. low/dense), allowing the impact of that specific manipulation to be measured under otherwise constant conditions (see Figure 8.2). If manipulations are applied such that these pairs don't form and conditions always differ in more than one way at a time, it becomes impossible to identify causal relationships with certainty. The goal is thus not simply to apply manipulations, but to do so to make it straightforward to identify and measure causal links if they indeed exist.

EXAMPLES OF NETWORK EXPERIMENTS

While experimental designs do not make up a majority of social networks research, there have been a number of examples, some quite recent, in the literature. Here we review several of these to provide some idea of the types of questions that can be addressed using experimental approaches. These are not chosen because they are necessarily the most important experiments that have been performed, but instead because they provide good examples of true experiments (e.g., Brashears & Gladstone, 2016; O'Connor & Gladstone, 2015) and quasi-experiments (e.g., Centola, 2011; Salganik, Dodds, & Watts, 2006). We do not provide examples of natural experiments for reasons of brevity and because these of necessity are relatively more unique than other types of experiments (but for an example see Godechot, 2016).



Homophily and the Spread of Health Behaviors

Damon Centola (2011) used a quasi-experiment to investigate how homophily (the tendency for individuals to associate with those like themselves or for birds of a feather to flock together) influences the likelihood of adopting a new health-related behavior. Centola argued that individuals are more likely to adopt a new behavior from others who are similar to themselves along a relevant dimension. For example, a healthier person is more likely to adopt a health-related behavior from another healthier person rather than from someone who is less healthy. An experimental design was particularly useful in this context because it isolates network structural topology from frequently associated confounds such as homophily, geographic proximity, and interpersonal affect (Blau & Schwartz, 1984; McPherson, Smith-Lovin, & Cook, 2001).

Centola examined the spread of a health behavior through an internet-based social networking environment, created for the experiment, populated by 710 participants recruited from an online fitness website. Individual traits including age, gender, and body mass index (BMI) were used to produce two types of networks: networks generated by allocating individuals at random and networks generated to produce homogeneity on the individual-level traits. These conditions allowed Centola to compare diffusion of behavior in homophilous conditions to diffusion of behavior without homophily, and individuals were assigned at random to one or the other condition.

Participants generated an anonymous social profile that displayed their gender, BMI, fitness level, diet, and favorite exercise. Subjects were then matched with other participants in the community (i.e., network ties were imposed). An online dashboard informed both parties about each other and their respective adoption of healthy behaviors. Participants ultimately made decisions regarding the adoption of an internet-based diet diary. Importantly, only those in the experiment knew of or could access the diet diary, ensuring that the only way to learn about or adopt the behavior was to receive a dashboard notification from a community partner. The experiment began when the public dashboard of a healthy seed—a member of the experimental community with above-average exercise minutes and fitness and a below-average BMI—displayed the adoption of the health diary. If a neighbor observed the notification via the online dashboard and adopted the behavior as well, then his or her neighbors were alerted in turn. After seven weeks of observation, the experiment concluded and Centola examined the rates of diet diary adoption across conditions.

Centola found that the rate of behavioral adoption was significantly higher in the homophilous networks compared to the nonhomophilous networks. The results suggest that the most effective health interventions will be mediated by social influence from similar individuals, which has obvious public policy implications.

This research benefited substantially from a quasi-experimental design. While all aspects of the subjects' environment were not under researcher control (e.g., they were logging in from unknown physical locations), the conditions were identical with the exception of the experimental manipulation. In other words, the interfaces were the same, starting conditions were the same, and community features were identical, aside from the creation of homogeneous or heterogeneous networks. Causal identification was substantially improved by the design as it eliminated selection effects; all individuals

in the experiment were identical in selecting into a health-related website and no unobserved variables influenced assignment to a condition. Finally, because the experiment was conducted online, it was relatively more generalizable than either a lab study or a naturalistic study of a single bounded context. That is, the online participant pool is more representative of the general population than is a behavioral laboratory containing, largely, undergraduates. Likewise, the results, generated from multiple artificially induced networks, are more generalizable than a naturalistic study of a single social network.

Networks and the Matthew Effect

Salganik et al. (2006) investigated how social influence impacts the likelihood that a song is viewed favorably. In particular, the authors were interested in whether or not perceptions of successful bands would become a self-fulfilling prophecy (Merton, 1948). In short, a self-fulfilling prophecy means that an initially false state of events evokes a new state of events that makes the originally false situation come true. They explored this question by artificially manipulating the initial popularity of various songs in an online music market populated by 14,341 participants.

The authors recruited participants from among those who had been involved in prior studies using email solicitations. The emails contained a link to a custom-designed website, as well as some information regarding the purpose of the study. As was typical, the emails did not reveal the full goals of the study so as to not introduce a confound. The link led participants to a menu containing 48 songs, which they could listen to, rate (from 1 to 5), and download for free, if desired. Participants could only rate songs if they had listened to them first, and could only download them if they had rated them. These three behaviors formed the basis of the authors' dependent variable.

Participants were randomly assigned to one of two conditions: an independent condition and a social influence condition. In the independent condition, songs were randomly presented in a list of 48, and no information about prior downloads by others was provided. In the social influence condition, participants were randomly assigned to one of eight initially identical "worlds," allowing eight replications of the same process from the same starting conditions. In the social influence condition, each song was accompanied by a number indicating how often songs were downloaded by prior participants. Salganik et al. also generated an additional condition wherein they manipulated the popularity of a song. Songs that were previously the most popular were now displayed as least popular, and vice versa for the least popular songs. Ultimately, Salganik et al. found that many songs experienced self-fulfilling prophecies in which initially false perceptions of popularity became real over time.

In this example, the authors used two manipulations: the presence or absence of social influence and the deliberate modification of popularity scores. Both allowed the authors to show that judgments of song quality depend significantly, though not entirely, on perceptions of the views of others. Further, the authors demonstrated that web-based experiments can yield powerful results in the study of macro-sociological processes that otherwise would be quite difficult to study in brick-and-mortar settings.

Error and Error Correction Process in Network Diffusion

Brashears and Gladstone (2016) used experimental methods to study (1) how rapidly errors accumulate during social networks contagions, (2) whether some message formats are more likely to accumulate error when compared to others, and (3) what human efforts at correcting errors do to the contagion.

The authors designed an experiment wherein participants read, remembered, and retransmitted a series of 10 sentences (13 to 16 words in length) drawn from popular press books, ensuring a standard and relatively undemanding reading level. Reading and retransmitting a series of messages acted as an analog for the more general process of receiving and sending other types of network contagions, all of which involve the movement of information.

While the study was administered via computer, as with the prior examples, all participants completed the experiment in a physical laboratory, allowing the authors total control over environmental conditions. A given participant was presented with a sentence for five seconds and was then forced to watch a blank screen for five seconds, before being given unlimited time to enter a new sentence into the computer. The time limits were used to capture real-world time and cognitive restraints, without being excessively demanding. This process was repeated until the participant read and retransmitted all 10 sentences. These rekeyed sentences then acted as the new stimulus sentences for the next participant.⁵

Brashears and Gladstone introduced two manipulations: a format manipulation and an error correction manipulation. First, the authors hypothesized that the redundancy of the message format (see Shannon, 1950) would influence the accuracy of retransmitted messages. In information theory, redundancy, or entropy, refers to the extent to which a given signal can be identified based on a subset of the elements. For example, at one extreme, an endless sequence of one character (e.g., “1111111...”) is maximally redundant (minimal entropy) because the entire sequence is known once the first character has been identified. At the opposite extreme, an endless sequence of random numbers is minimally redundant (maximum entropy) because each character identified provides no improvement in the ability to identify the full sequence. The stimulus sentences were thus presented either in high-redundancy standard English (e.g., the format of this chapter) or in low-redundancy text messaging pidgin (e.g., “I see you” becomes “I c u”). Second, the authors argued that humans would attempt to correct errors that they detected, and that these corrections would impact message fidelity. In the No Correction condition, participants were instructed to read and retransmit the message verbatim. In the Correction condition, participants were asked to read and reproduce what they believed to be the intended meaning of the message (i.e., to paraphrase). These manipulations were crossed to produce a two (message format: standard English vs. texting pidgin) by two (attempts to correct messages: Correction vs. No Correction) design, yielding all possible combinations of the manipulations. Participants were randomly assigned to a condition, thus ensuring that individual variation did not affect the results. The results indicated that high-redundancy formats are more robust against error, that error correction efforts are generally effective, and that failed error corrections produce more widely varying descendant messages. Thus, language with more built-in fault tolerance can weather error storms to a greater degree. Further, and somewhat counterintuitively, the (failed) act of correcting errors gives rise to substantially more distinct forms of

the original message than do the original errors. Thus, diversity emerges not in spite of attempts to conform, but because of it, producing “innovation from imitation.”

The use of a true experimental design in this study allowed the precise identification of causal processes; differences between conditions could be attributed conclusively to message format, or the presence of error correction. Moreover, unlike previous similar research tracking error in online network contagions (e.g., Simmons, Adamic, & Adar, 2011), Brashears and Gladstone were able to monitor the network at all times, ensuring that all predecessor and descendant messages were identified and properly linked. Finally, this study shows how a mechanism (i.e., network contagion) can be reduced to fit into a laboratory even though the context in which it normally unfolds (i.e., an actual network) is not.

Network Recall and Social Exclusion

O'Connor and Gladstone (2015) use true experimental methods to explore the ways in which people perceive and recall novel network structures. The authors hypothesized that social exclusion would systematically and negatively impact an individual's ability to decode the network structures embedded in unfamiliar social situations. More specifically, the authors argued that compared to those who were not excluded, excluded participants would view new social groups as relatively denser than (1) they actually were or (2) those who were not excluded.

The authors conducted three separate studies, only two of which we review here. To determine whether exclusion systematically impacts perceptions of novel networks, they manipulated the description of the network itself (Figure 8.3). Using Amazon Mechanical Turk, O'Connor and Gladstone recruited roughly 140 participants from the United States, Canada, the United Kingdom, and Australia. After providing their informed consent to participate in the experiment, participants were asked to rate the extent to which they felt (1) popular (reverse-scored), (2) part of a social group (reverse-scored), (3) included (reverse-scored), and (4) ostracized. These variables were then merged to form a composite index of overall feelings of social exclusion.⁶

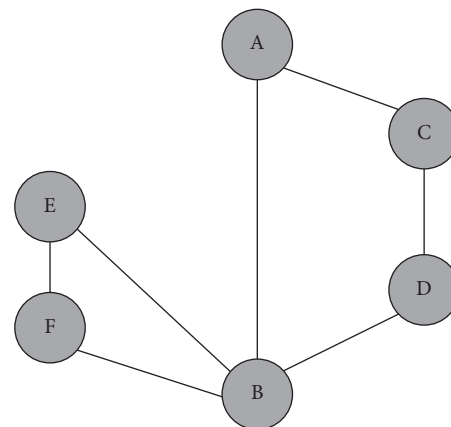


FIGURE 8.3 Stylized social network used in Studies 2 and 3.

After providing their perceptions of social exclusion, participants observed a network sociogram containing seven ties (Figure 8.3) for 30 seconds before the software auto-advanced to a new screen containing questions regarding network perceptions. While all participants saw the same network structure, half received a network graphic described as a social group, while the other half received a network graphic described as a road and town map. The dependent variable used by the authors was the number of relationships/roads participants recalled after seeing the networks.

The authors found that as participants' self-reported feelings of social exclusion increased, so too did their tendency to overestimate the number of social relations in the social network graphic (i.e., the social group condition), but not in the road and town network graphic (i.e., the road map condition). The results suggested that social exclusion has a significant impact on the recall of social relations, but not relations in general.

In Study 2, O'Connor and Gladstone sought to establish a stronger causal relationship between ostracism and perceptions of novel social networks. In Study 1, a correlate of feeling socially excluded (e.g., taxed cognitive abilities) may have driven the observed effects. To remedy this problem, O'Connor and Gladstone relied on psychological priming, a manipulation used commonly in psychology, social psychology, and economics (e.g., Brashears, Hoagland, & Quintane, 2016; Stevens, 1951). Priming is often used to study a variable of interest without introducing that variable's correlates into the experiment, for example, to induce a feeling of being powerful without also bringing along high education, wealth, or formal authority. There are many established ways to prime an individual (for the seminal work on priming, see Meyer & Schvaneveldt, 1971), but the general aim remains the same: to induce a state of the variable of interest without that variable's correlates being present.

After providing their informed consent, participants were directed to a new screen that depicted a game called Cyberball (Williams, 2006). Cyberball is a three-person computer-based ball toss game designed to prime and manipulate participants' experiences and feelings of social exclusion. Participants in the lab believed, falsely, that they were playing this game with other individuals in the room.

The authors created two conditions: inclusion and exclusion. In the exclusion conditions, participants were thrown the ball one-fifth of the time, whereas in the inclusion condition, participants were thrown the ball four-fifths of the time. After playing the game for four minutes, participants completed a brief inventory that served as the author's manipulation check.

Participants then moved on to the second component of the study, where they were presented with a custom-designed video animation showing figures (known as avatars) interacting in a number of social settings. The video was eight minutes long and depicted the avatars talking and dancing across three separate scenes: a beer garden, walking into a club, and dancing to music. The authors used a video rather than a stylized network graph to increase realism and ensure that the findings in Study 1 were consistent across different methods of presentation. Importantly, the underlying network structure of the social interactions depicted in the video remained identical to that of Study 1 (Figure 8.3). Again, the authors found that experiences of ostracism systematically altered the way in which social networks were recalled such that ostracized individuals perceive other networks as denser than they are.

Using true experimental procedures, O'Connor and Gladstone were able to link psychological processes to behaviors that generate macro-level network structures. This experiment demonstrates, first and foremost, how multiple studies can be used to improve causal

identification. This is simple and fast in an experimental framework, but substantially more difficult in a naturalistic setting. Second, this experiment showcases how modern experiments can present material in multiple formats (e.g., videos of avatars, sociograms, etc.) and can mislead participants into believing they are interacting with others when they are only interacting with a computer program. Finally, this experiment combined a quasi-experimental design (Study 1) with a true experimental design (Study 2), thereby improving both causal identification and generalizability.

CONCLUSION

In this chapter we have argued that experimental designs are not just feasible for the study of social networks, but in many respects optimal. They provide maximum causal identification and, with the use of modern crowdsourcing services, often require only minimal sacrifices to generalizability. Thus, by either shrinking a process to fit into the lab or expanding the lab to encompass a process, we are able to more precisely study social network mechanisms with experiments than nearly any other method. We additionally provided several examples of social network experiments from the existing literature, though we were by no means exhaustive in our listing (e.g., Muchnik, Aral, & Taylor, 2013). These give an idea of the range of manipulations that are available, the topics that can be studied, and the insights that can be achieved by using experimental designs. As such, we hope that the reader agrees that the time is right for network experiments.

While we are enthusiastic about the capabilities of network experiments, we do not mean to suggest that all research questions should be addressed with an experiment. Naturalistic designs must remain a significant part of the repertoire of social network analysts, and we do not mean to imply otherwise. Nevertheless, it is a mistake to assume that the ideal context in which to study a process is one that duplicates the natural environment as precisely as possible. A globe the size of the earth is of no use. The natural environment is complex, messy, and prone to generating incorrect causal models. The laboratory, in contrast, provides a calm and controlled environment in which the hidden workings of the social world may be uncovered. Armed with this detailed knowledge, it is then possible to examine the “real world” for signs of these mechanisms in action and to use deviations from those expectations to identify missing variables. In short, experiments have a critical role to play alongside natural studies, but neither of the two can replace the other.

NOTES

1. <https://www.mturk.com/mturk/welcome>.
2. For sociological research on the community of MTurk workers and the future of crowdsourcing, see Kittur et al. (2013).
3. <https://volunteerscience.com/volunteerscience.html>
4. <https://www.zooniverse.org/>
5. The similarity to the children’s game of “telephone” (also known by too many other names to mention here) is intentional.

6. A composite index of at least three items is generally preferable when assessing psychological and social psychological processes and should yield Cronbach's alpha values of at least 0.7 (Cronbach & Shavelson, 2004).

REFERENCES

- Blau, P. M., & Schwartz, J. E. (1984). *Crosscutting social circles*. New Brunswick, NJ: Transaction Publishers.
- Brashears, M. E. (2008). Sex, society and association: A cross-national examination of status construction theory. *Social Psychology Quarterly*, 71, 72–85.
- Brashears, M. E. (2011). Small networks and high isolation? A reexamination of American discussion networks. *Social Networks*, 33, 331–341.
- Brashears, M. E. (2014). “Trivial topics” and rich ties: The relationship between discussion topic, alter role, and resource availability using the “important matters” name generator. *Sociological Science*, 1, 493–511.
- Brashears, M. E., & Gladstone, E. (2016). Error correction mechanisms in social networks can reduce accuracy and encourage innovation. *Social Networks*, 44, 22–35.
- Brashears, M. E., Hoagland, E., & Quintane, E. (2016). Sex and network recall accuracy. *Social Networks*, 44, 74–84.
- Brashears, M. E., & Quintane, E. (2018). The weakness of tie strength. *Social Networks*, 55, 104–115.
- Buhrmester, M. D., 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.
- Burt, R. S. (1980). Innovation as a structural interest: Rethinking the impact of network position on innovation adoption. *Social Networks*, 2, 327–355.
- Burt, R. S. (1992). *Structural holes: The social structure of competition*. Cambridge, MA: Harvard University Press.
- Centola, D. (2011). An experimental study of homophily in the adoption of health behavior. *Science*, 334, 1269–1272.
- Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine*, 357, 370–379.
- Cohen-Cole, E., & Fletcher, J. M. (2008a). Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic. *Journal of Health Economics*, 27, 1382–1387.
- Cohen-Cole, E., & Fletcher, J. M. (2008b). Detecting implausible social network effects in acne, height, and headaches: Longitudinal analysis. *British Medical Journal*, 337, a2533.
- Coleman, J. S., Katz, E., Menzel, H., Columbia University, & Bureau of Applied Social Research. (1966). *Medical innovation: A diffusion study*. Indianapolis, IN: Bobbs-Merrill Co.
- Cronbach, L. J., & Shavelson, R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement*, 64(3), 391–418.
- Eagle, N., Pentland, A. S., & Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences, USA*, 106, 15274–15278.
- Fischer, C. S. (1982). *To dwell among friends: Personal networks in town and city*. Chicago, IL: University of Chicago Press.
- Fischer, C. S. (2009). The 2004 GSS finding of shrunken social networks: An artifact? *American Sociological Review*, 74, 657–669.

- Fowler, J. H., & Christakis, N. A. (2008). Estimating peer effects on health in social networks: A response to Cohen-Cole and Fletcher; and Trogdon, Nonnemaker, and Pais. *Journal of Health Economics*, 27, 1400–1405.
- Friedkin, N. E. (1998). *A structural theory of social influence*. New York, NY: Cambridge University Press.
- Godechot, O. (2016). The chance of influence: A natural experiment on the role of social capital in faculty recruitment. *Social Networks*, 46, 60–75.
- Gosling, S. D., & Johnson, J. A. (Eds.). (2010). *Advanced methods for conducting online behavioral research*. Washington, DC: American Psychological Association.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33, 1–75.
- Hipp, J. R., Wang, C., Butts, C. T., Jose, R., & Lakon, C. M. (2015). Research note: The consequences of different methods for handling missing network data in stochastic actor based models. *Social Networks*, 41, 56–71.
- Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2012). The online laboratory. *Experimental Economics*.
- Jackson, M., & Cox, D. R. (2013). The principles of experimental design and their application in sociology. *Annual Review of Sociology*, 39, 27–49.
- Keppel, G. (1991). *Design and analysis: A researcher's handbook*. Upper Saddle River, New Jersey: Prentice-Hall.
- Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological research online: Opportunities and challenges. *American Psychologist*, 59, 105–117.
- Laumann, E. O., Marsden, P. V., & Prensky, D. (1992). The boundary specification problem in network analysis. In L. C. Freeman, D. R. White, & A. K. Romney (Eds.), *Research methods in social network analysis* (pp. 61–79). New Brunswick, NJ: Transaction Publishers.
- Lee, B., & Bearman, P. (2017). Important matters in political context. *Sociological Science*, 4, 1–30.
- Lusher, D., Koskinen, J., & Robins, G. (2013). *Exponential random graph models for social networks: Theory, methods, and applications*. New York, NY: Cambridge University Press.
- McCarty, C., Killworth, P. D., & Rennell, J. (2007). Impact of methods for reducing respondent burden on personal network structural measures. *Social Networks*, 29, 300–315.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27, 415–444.
- Merluzzi, J., & Burt, R. S. (2013). How many names are enough? Identifying network effects with the least set of listed contacts. *Social Networks*, 35, 331–337.
- Merton, R. K. (1948). The self-fulfilling prophecy. *The antioch review*, 8(2), 193–210.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90(2), 227.
- Muchnik, L., Aral, S., & Taylor, S. J. (2013). Social influence bias: A randomized experiment. *Science*, 341, 647–651.
- O'Connor, K. M., & Gladstone, E. (2015). How social exclusion distorts social network perceptions. *Social Networks*, 40, 123–128.
- Paik, A., & Sanchagrin, K. (2013). Social isolation in America: An artifact. *American Sociological Review*, 78, 339–360.
- Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5, 411–419.



- Rand, D. G., et al. (2013). Evolution of fairness in the one-shot anonymous Ultimatum Game. *Proceedings of the National Academy of Sciences, USA*, 110(7), 2581–2586.
- Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311, 854–856.
- Schaefer, D. R. (2007). Votes, favors, toys, and ideas: The effect of resource characteristics on power in exchange networks. *Sociological Focus*, 40, 138–160.
- Shannon, C. E. (1950). Memory requirements in a telephone exchange. *The Bell System Technical Journal*, 29(3), 343–349.
- Simmons, M. P., Adamic, L. A., & Adar, E. (2011). Memes online: Extracted, subtracted, injected, and recollected. *International Conference on Web & Social Media*, 2011.
- Simpson, B., & Walker, H. A. (2002). Status characteristics and performance expectations: A reformulation. *Sociological Theory*, 20(1), 24–40.
- Stevens, S. S. (1951). *Handbook of experimental psychology*. Hoboken, New Jersey: Wiley Publishing.
- Van den Bulte, C., & Lilien, G. L. (2001). Medical innovation revisited: Social contagion versus marketing effort. *American Journal of Sociology*, 106, 1409–1435.
- Weinberg, J. D., Freese, J., & McElhatten, D. (2014). Comparing data characteristics and results of an online factorial survey between a population-based and a crowdsourcing-recruited sample. *Sociological Science*, 1, 292–310.
- Willer, D. (1999). *Network exchange theory*. Westport, CT: Praeger.
- Williams, K. D., & Jarvis, B. (2006). Cyberball: A program for use in research on interpersonal ostracism and acceptance. *Behavior research methods*, 38(1), 174–180.
- Zelditch, M. (1969). Can you really study an army in the laboratory? In A. Etzioni (Ed.), *Complex organizations* (2nd ed., pp. 528–539). New York, NY: Holt, Rinehart and Winston.

Author Query

AU1: Year in references is 2004; please check.