Chapter 17
Social Networks and Causal Inference

Tyler J. VanderWeele and Weihua An

Abstract This chapter reviews theoretical developments and empirical studies related to causal inference on social networks from both experimental and observational studies. Discussion is given to the effect of experimental interventions on outcomes and behaviors and how these effects relate to the presence of social ties, the position of individuals within the network, and the underlying structure and properties of the network. The effects of such experimental interventions on changing the network structure itself and potential feedback between behaviors and network changes are also discussed. With observational data, correlations in behavior or outcomes between individuals with network ties may be due to social influence, homophily, or environmental confounding. With cross-sectional data these three sources of correlation cannot be distinguished. Methods employing longitudinal observational data that can help distinguish between social influence, homophily, and environmental confounding are described, along with their limitations. Proposals are made regarding future research directions and methodological developments that would help put causal inference on social networks on a firmer theoretical footing.

Introduction

Although the literature on social networks has grown dramatically in recent years (see Goldenberg et al. 2009; An 2011a, for reviews), formal theory for inferring causation on social networks is arguably still in its infancy. As described elsewhere in this handbook, much of the recent theoretical development within causal inference has been within a "counterfactual-" or "potential outcomes-" based perspective (Rubin 1974; see also Chap. 5 by Mahoney et al., this volume), and this approach typically employs a "no interference" assumption (Cox 1958; Rubin 1980; see also Chap. 16 by Hong and Raudenbush, this volume), that the outcomes of one individual are not influenced by the exposures or treatments of other individuals. Such "no interference" assumptions are implausible in many sociological settings and will generally fail whenever the mechanisms relating a treatment or exposure to an outcome can be influenced by social interaction. Not infrequently, social network data is collected precisely when such interference or social interaction occurs.

T.J. VanderWeele (✉)
Departments of Epidemiology and Biostatistics, Harvard University, Boston, MA, USA
e-mail: tvanderw@hsph.harvard.edu

W. An
Departments of Sociology and Statistics, Indiana University, Bloomington, IN, USA

S.L. Morgan (ed.), *Handbook of Causal Analysis for Social Research*,
Handbooks of Sociology and Social Research, DOI 10.1007/978-94-007-6094-3_17,
© Springer Science+Business Media Dordrecht 2013

Some progress has been made extending this potential outcomes framework to allow for such interference or spillover effects in which the exposure of one individual can affect the outcomes of others (Hong and Raudenbush 2006; Sobel 2006; Rosenbaum 2007; Hudgens and Halloran 2008; Graham 2008; VanderWeele and Tchetgen Tchetgen 2011; Tchetgen Tchetgen and VanderWeele 2012; Manski 2013). However, the literature at present has typically assumed that the persons under study are at least grouped into clusters so that it is the case that the exposure of an individual in one cluster cannot influence the outcome of an individual in a different cluster. Within the context of a group of individuals defined by a social network, often this assumption too will fail. Every individual in the social network may be linked to every other individual by a series of intermediate social ties. Thus, even the recent extensions to the counterfactual-based approach to causal inference essentially fail and are inapplicable in such social network contexts.

In spite of these challenges, several attempts have been made to develop formal approaches to infer causation. When samples from several different social networks are available, randomized trials have been designed that enable causal inferences (Sacerdote 2001; Duflo and Saez 2003; Camargo et al. 2010; An 2011b). When only observational data is available, instrumental variable regressions have been proposed to identify causal network effects, also referred to as causal peer effects (Angrist and Pischke 2008; Bramoullé et al. 2009; O'Malley et al. 2011; An 2011c). More controversially, observational longitudinal data with repeated measures has been used to attempt to draw conclusions about social influence on a network, even when data from only a single social network is available (Christakis and Fowler 2007, 2008, 2013; Snijders 2001, 2005; Steglich et al. 2010).

In this chapter, we will review some of the literature on causal inference for social networks. We will discuss randomized trial designs for social network structures and the types of studies that can be conducted with such designs. We will also review approaches to inferring causation from observational longitudinal social network data that has led to claims that health behaviors and states as diverse as obesity, smoking, happiness, and even loneliness somehow travel through social networks via social influence (Christakis and Fowler 2007, 2008). We will describe why such analyses are seen as controversial, some of the criticisms that have been leveled against the methodology, and some of the responses to these criticisms. We will also describe an alternative approach based on a stochastic actor-oriented model (Snijders 2001, 2005; Steglich et al. 2010) that more explicitly models the selection of ties. Finally, we will discuss developments that may be needed to put causal inference from observational social network data on a firmer theoretical footing.

Before we begin the discussion, it should be noted that within the context of social networks, different forms of causation may be at play. One form is social influence whereby the behavior, states, and characteristics of one individual in a network may influence behaviors, states, and characteristics of others in the network with whom the first individual shares some form of social tie, directly or indirectly. However, within the context of social networks, another form of causation that may be present is that of network formation, whereby the behavior, states, and characteristics of various individuals may exert influence on whether and which social ties are present to begin with or later develop. While our focus in this chapter will be on the first form of causation, namely, social influence (also called peer effects, relational effects, induction, contagion, or network effects), we will also consider network formation (also referred to as friendship formation/selection, homophily, or effects on networks). Indeed the complex interplay between these different forms of causation is in part what makes inferring causation with social network data especially challenging.

More generally if the behavior, states, and characteristics of two individuals with a social tie are found to be correlated, it is possible to envision at least three potential explanations. First, it is possible that the association is due to social influence: One of the persons may have influenced the other, or vice versa, or both. Second, it is possible that the behavior or states are correlated because persons with similar characteristics are more likely to become friends with one another. This phenomenon is sometimes referred to as homophily or selection. Third, it is possible that there is some shared environmental factor that influences the states or behaviors of both individuals so

that they are correlated. With cross-sectional data, it is essentially impossible to distinguish between these three explanations. Manski (1993) referred to this as the "reflection problem." However, as discussed in this chapter, when it is possible to conduct randomized experiments (randomizing either social ties or particular interventions), or when longitudinal data is available, progress can be made in distinguishing influence, homophily, and environmental confounding. We will begin our discussion with randomized experiments on social networks and then turn to observational longitudinal social network data.

Causal Inference and Network Experiments

Experiments are a useful method to identify and estimate causal network effects (e.g., the effect of an intervention for one individual on outcomes of others, magnified or diffused by social ties) or effects on networks (e.g., the effect of an intervention on the structure of the network itself). Roughly speaking, experiments can be divided into controlled experiments (whether in the labs or in the field) and natural experiments. Controlled experiments are conducted purposefully by researchers, while natural experiments exist mostly as unintentional consequences of either natural forces or historical legacy. We will cover both types of experiments in the following review. The effects of interventions can be roughly divided into "network effects" and "effects on networks." Network effects concern how an intervention on one person may affect not only his or her own outcome but also the outcomes of others, and how the effect of the intervention may depend upon the structure of the network. Effects on networks concern how an intervention may alter the structure of the network itself.

There have only been a few network experiments in the literature. Moreover, many of these are not interested in making causal inference about either network effects or effects on networks but are focused on using networks as a medium to facilitate the implementation of interventions (Valente 2005). For example, Campbell et al. (2008) used social network information (more precisely, student nominations of influential peer leaders) to identify opinion leaders in secondary schools and trained them to encourage their classmates not to smoke. Here, the focus is on the intervention effects, and the social network is used just as a media to choose opinion leaders.

Network Effects

We begin our discussion with network effects. Roughly speaking, there are three kinds of network effects: relational effects, positional effects, and structural effects. Relational effects are the micro level of network effects, referring to whether and to what degree possessing certain social relationships affects a focal subject's outcomes, for example, whether the characteristics of one individual will affect the outcomes of his or her social contacts. Most of the research on social capital or relational capital and peer effects falls into this category. Positional effects are the meso level of network effects, which refer to whether and to what degree a person's position in a network affects his or her outcomes. Structural effects are the macro level of network effects, which refer to whether and to what degree the features of the entire network affect the outcomes of the subjects embedded in the network.

The main difficulty in identifying and estimating network effects with observational data is dealing with the endogeneity of networks, namely that the formation of networks may not be a random process but a process driven by selection or unobserved confounding factors that affect both the formation of networks and the outcomes of interest. For example, when studying the monetary returns to relational capital, if measures on human capital are not properly adjusted for, the estimates may be upwardly

biased, as human capital likely determines both how much relational capital a person can possess and how much money a person can earn. In many cases, experiments can help us circumvent such problems.

Relational Effects

Two types of network experiments can be designed to study relational effects (An 2011a). Type I experiments randomly assign social contacts to subjects, the main purpose of which is to eliminate the selection or homophily problem that "birds of a feather flock together." For example, Sacerdote (2001) found that when roommates and dormmates were randomly assigned, the academic achievement of a student had significant impact on the academic performance and social activities of other students in the same room.

Such random assignment constitutes a powerful approach for assessing peer influence but is subject to a number of limitations. First, it might sometimes be infeasible to randomly assign social contacts to subjects due to ethical or other concerns. For example, to the extent that many student dormitories (and most rooms) are not coed, it may be against university policy to assign a student with a roommate from opposite gender. Analyses would then have to be conducted within strata of gender. Second, even if we can randomly assign social contacts to subjects and even if we indeed find there is a significant correlation in the outcomes of the subjects and their assigned contacts, we cannot necessarily attribute this correlation to relational effects or peer influence, because there still is the possibility that the correlation is driven by some contextual factors that affect both the subjects and their assigned contacts. For example, even if college roommates are randomly assigned and we find that there is a significant correlation in their academic performance, the correlation may be at least partly generated by their common local circumstances (their shared living conditions, e.g., living in a dorm on a quiet versus a noisy street). One potential solution to this problem is to use lagged outcomes of the randomly assigned contacts to predict the focal subject's outcomes. Indeed any characteristic or covariate that is available prior to randomization could be used for this purpose. If a pre-randomization characteristic is used and associations with peer outcomes are assessed, then homophily/selection is eliminated by randomization, and environmental confounding is effectively eliminated by using characteristics that occurred before either individual was in their shared environment. For example, Kremer and Levy (2008) used student drinking behavior in the year before entering college to predict their roommates' drinking behaviors at a large state university where roommates were randomly assigned and found significant peer effects on drinking. A third critique of these experiments is that they implicitly assume dyadic independence and neglect any kind of higher order of network effects. Consequentially, interference between dyads may lead to biased estimates of causal relational effects.

Type II network experiments are particularly useful in situations where random assignment of social contacts is infeasible. In a type II experiment, we do not assign any social contacts to subjects but condition on the existing social relationships among subjects and aim to study how the effects of an external intervention spreads through social contacts. The type II experiments often employ a partial treatment group design, in which only some of the members in the treatment groups will receive an external intervention. See Fig. 17.1 for an illustration, where treatment groups and control groups are usually defined according to natural social boundaries, which can be classrooms, military units, social clubs, etc.

If the treatment status of the individuals in the treatment group is determined randomly, then any difference in the average outcomes between the untreated subjects in the treatment groups and the subjects in the control groups can be attributed to the treatment spillover effects from the treated subjects in the treatment groups. Several kinds of causal peer effects can be inferred from this partial treatment group design. First, if the average outcome of the treated subjects in the treatment groups is significantly different from that of the subjects in the control groups, we can view such a difference

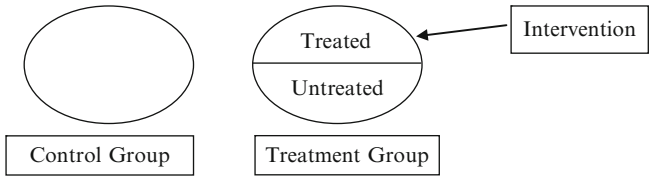

Fig. 17.1 An illustration of the partial treatment group design

as an estimate of the total effect of the intervention for treated subjects or the first-order effects of the intervention. Second, the treatment spillover effects we described above can be viewed as peer effects, indirect effects, or the second-order effects of the intervention. To differentiate such peer effects from the ones we are going to introduce later, we can designate them as peer effects under control (PEC), as the groups being compared do not directly receive any intervention. Third, we could compare the treated subjects in the treated group to the control subjects in the treated group and view this as the direct effect of treatment. Finally, we can compare the treatment groups and the control groups holistically and view any difference in the average outcomes between them as an estimate of the overall effect of the intervention. Such designs and effect definitions have likewise been considered in the context of cluster-randomized vaccine studies (Halloran and Struchiner 1991; Hudgens and Halloran 2008; VanderWeele and Tchetgen Tchetgen 2011) and are clearly applicable to the social network setting as well. As we will see below, information on network structure will allow us to draw conclusions about other types of effects as well.

Several implicit assumptions should be spelled out clearly here. First, some kinds of relations or connections are assumed to exist between the treated and the untreated subjects in the treatment groups so the treatment effects can possibly spread along social ties. Usually we also assume there is a certain degree of coupling between the social relations and the treatment effects. For example, smoking or drinking behaviors can possibly spread through friendship ties; however, we might expect other outcomes or behaviors such as weight or height to be less subject to social influence (though see below). Second, to speak of the effect of one individual on another, we would ideally like to use only pairs of subjects in the treatment groups which are pairwise independent. If this is not the case, there may be treatment interference between subjects in the treatment groups, which will impose a problem for estimating pairwise causal peer effects (Falk and Ichino 2006). However, in reality, peer effects have an intertwining nature and often are not pairwise independent. Although the partial treatment group design may thus not be able to estimate pairwise causal peer effects (unless all groups have only two individuals and the groups are independent from one another), the design may still serve well for estimating peer effects in more natural settings such as classrooms and clubs. Formal statistical inference in this setting can however be challenging (Hudgens and Halloran 2008; Tchetgen Tchetgen and VanderWeele 2012).

The third concern with this design is that the control groups and the treatment groups should be comparable, particularly regarding the network structures, since relationships are used as channels to diffuse information or behaviors. A two-stage randomized design whereby networks are first randomized to the treatment or control arms and then individuals within the treatment arm are randomized to treatment can help ensure this. However, with studies employing a smaller number of networks, there will likely be chance imbalances. There is no easy way to gauge how similar two networks are, so a large number of treatment and control groups may be needed to mitigate chance imbalance in the network structure. A final concern, which is only applicable in certain particular cases, is that the control groups and the treatment groups need to be reasonably independent from one another. For example, if the above design is implemented in a school and classroom is used as an experimental unit, it is likely some of the treatment effects can spill over to the control classrooms due to cross-classroom friendships. Note that this kind of interference will usually lead to more conservative estimates of peer effects.

There are a couple of examples in the literature using the partial treatment group design to estimate peer effects. For example, Duflo and Saez (2003) found that uninformed employees within departments of a university where some of their colleagues were notified of an information fair about retirement plans were more likely to attend the fair than those employees in the control departments where no one was notified of the fair. Cipollone and Rosolia (2007) used a natural experiment to facilitate identifying peer effects on education. They found that a few cohorts of male students in southern Italy who were granted an exemption from compulsory military service as a result of an earthquake in 1980 not only boosted their own high school graduation rate by over 2 % points but also resulted in the graduation rate of their female counterparts being increased by about 2 % points. As women in Italy were not subject to military draft, the authors argued that the increase in their educational attainment must be a result of the schooling behavior of their male counterparts. The earthquake in 1980 appears to act as an instrument for male students' educational achievement. However, a critique of using this kind of instrumental variable is that natural forces may change the local context that affects the outcomes of both the treated and the supposedly untreated at the same time. For example, it may be the case that the earthquake itself changed the employment patterns in the areas hit by the earthquake or that the additional resources granted for earthquake relief encouraged women to stay in school longer than before. If this is true, then the reported peer effects may have been overestimated.

The partial treatment group design can be adapted in several ways to serve different research purposes and to infer several different kinds of relational effects or peer effects in the social network context. If our goal is to accelerate the diffusion of information or behaviors associated with a certain intervention, as in many public health studies, we can choose subjects who are central in the treatment groups to be the "change agents" (e.g., the individuals towards whom an intervention is directed; see Valente and Davis 1999). We might assume that central subjects have the authority, credibility, or connections to spread any proposed diffusion in a group more widely than random subjects can. If in a trial, groups were randomized to either an arm in which treatment was given randomly or an arm in which treatment was assigned to central subjects, then the difference in the average outcomes between the untreated subjects in the central-subjects treatment arm and the untreated students in the random treatment arm can be viewed as an experimental estimate of peer effects under acceleration (PEA), indicating the added value provided by the central intervention. It should be noted that although a comparison of the average outcomes of both treated and untreated subjects in the random-assignment arm versus the central-subjects arm can be interpreted as a causal effect of the difference in intervention schemes, a comparison of the outcomes of only the untreated subjects in each of the two arms is a combination of the causal effect of treatment and also selection, since the untreated subjects in the two arms may not be comparable. The untreated subjects in the random-assignment arm constitute a random sample, whereas the untreated subjects in the central-subjects arm consists of individuals who are not central (see also An 2011b). To more rigorously infer peer effects in this design would require either regressions or matching to adjust for differences between the untreated subjects in both the central and random intervention arms. Even then, the resulting estimates will be conditional (applying to subjects with certain kinds of characteristics) and sensitive to unobserved differences between the two untreated groups.

If, however, the average of the untreated persons outcomes in each group is viewed as a property of the group rather than as the average of individual outcomes, then the difference in average untreated outcomes between the random-assignment arm and the central-subjects arm could still be interpreted as a group-level causal effect (without worrying about selection), though not as an average individual causal effect. In the group-level causal effect interpretation, selection is still in some sense present, but it is now being viewed as part of the intervention itself. The group-level interpretation defines the group-level outcome as the average of the untreated subjects, and the effect is unconditional (it applies to all groups in the study); the individual-level interpretation conditions on being untreated, thus introducing selection.

Identifying central subjects in a group is both a theoretical and technical problem. There are roughly two kinds of criteria: instrumental and integrative (Callahan and Robin 1969). The former defines a subject's centrality based on personal attributes, for example, academic performance, gender, and whether playing sports. The latter defines whether a subject is central based on the subject's connections with others. A subject with more friends in a group is more popular and central in that group. In network experiments, the latter is usually the default criterion to choose central subjects, as the main goal of most of these experiments is to change social behaviors such as drinking, smoking, and exercising. But some combination of these criteria may be more effective. For example, we may not want to choose central smokers as the change agents in a smoking prevention program.

Several specific methods that have been proposed in the literature to choose central subjects in a group/network according to the integrative criterion, including the indegree centrality method (Valente and Pumpuang 2007), the eigenvector centrality method (Bonacich 2007), and the key players method (Borgatti 2006). The key computational problem is to choose a fixed number of central subjects who will connect to the largest number of unique members in a group so that the treatment effects can potentially spread via social ties to most subjects. Utilizing the directionality of social ties to help identify central subjects is another key problem. An (2011d) developed a stochastic optimization method that can be used to address both problems.

Another adaptation of the partial treatment group design is that we can assign treatment to an entire set of individuals with social ties instead of assigning treatment to individuals randomly. The purpose is to utilize internal group support or other group dynamics to increase treatment effects. In an appropriately designed trial, we could use the difference in the average outcomes between the subjects who receive the treatment with their friends and the subjects who receive the treatment as randomly selected individuals to measure how much the treatment effects are enlarged via group dynamics. The difference can be viewed as peer effects under treatment (PET), as both of the groups being compared receive the treatment. Note that PET is also a combination of causal and selection effects, as the treated subjects in the group treatment may be different in other aspects from the randomly treated subjects. Regression or matching may be useful in teasing out part of the selection effect. In addition, PET may be underestimated, if the supposedly random subjects assigned to receive the treatment have some connections among one another.

There are a few studies that have used this group-based design to implement interventions. For example, Wing and Jeffery (1999) showed that subjects who were recruited and participated with their friends in a weight loss program were more likely to complete the program and lose more weight than those who were recruited and participated individually. A critical step in conducting the group interventions is to identify peer groups. Valente et al. (2003) showed that group interventions were more effective when natural peer groups were used than when peers were randomly grouped or grouped by teachers. Finding and extracting natural social groups from networks can be challenging even when we know all the social network information of the subjects.

Several difficult questions must be addressed in identifying and defining peer groups: Can a subject belong to several different groups? If so, which group should we assign a subject to? Should the treated groups have different sizes or the same size? When we extract social groups from a social network, should only mutual friends be grouped together so as to maximally utilize group support? Some methods have been developed in the literature to address these questions, including the hierarchical clustering method, the recursive neighborhood mean imputation method (Moody 2001), the edge-betweenness method (Girvan and Newman 2002), and a maximum-likelihood method (Copic et al. 2009). However, more work needs to be done to improve these methods by, for example, utilizing the directionality of social ties and reducing their tendency to generate sparse groups.

One question common to the different designs is whether to train or incentivize the change agents to actively advocate for the proposed diffusion (namely, to utilize active peer effects). For example, in a smoking prevention program, would training of the treated subjects and providing monetary incentives for them to actively advocate their peers not to smoke make the intervention more effective? This

consideration is consequential for the selection of central subjects. For example, to accelerate active peer effects using training, it may be that indegree (the number of received ties) is the more important characteristics, while if the purpose is to accelerate passive peer effects (namely, to utilize peer effects in their natural state, without training or extra incentives), outdegree (the number of outgoing ties) may be more important.

Positional Effects

Positional effects refer to whether and to what degree a person's position in a network affects his or her outcomes. The assessment of positional effects requires that particular measures of position be defined and utilized in analyses. Two of the most popular measures of network positions are structural hole and structural equivalence.

Structural hole measures the degree of disconnection among the contacts of a focal subject (Burt 1995). The proposition is sometimes put forward that the more holes in a person's local network (i.e., the higher degree of disconnection among a person's contacts), the more benefit this person can derive from occupying such a network position.

Structural equivalence measures how similar the patterns of connections of two subjects are. Two subjects are structurally equivalent if they are connected to the same contacts in the same manner (Wasserman and Faust 1994). The proposition is sometimes put forward that people are more likely to be affected by others who are structurally equivalent to them, because they may share similar information and risk exposure, may use each other as a reference, or may compete for status.

Although there have been observational studies that evaluate these propositions (Burt 1987, 2004; Cornwell 2009; Podolny 2001), experimental research is rare. One possible experimental design to evaluate these propositions would be to randomly assign network positions to subjects and see how individuals' outcomes are affected by these positions.

Structural Effects

The above review thus far focused on micro or meso levels of network effects. In contrast, the holistic features of a network may have profound impact as well. For example, in a recent study, Centola (2010) showed that networks with many local clusters were more effective for behavioral diffusion than random networks. The rationale behind this is somewhat similar to what we find in group interventions: Individuals are much more likely to adopt a health behavior when they receive social reinforcement from multiple contacts. But the focus here is on comparing the efficacy of different kinds of networks.

There is not much experimental research to date on macro network effects. One possible experimental design to study macro network effects is to randomly assign different kinds of networks to subjects and then compare the resulting outcomes, as has been done in Centola (2010). For example, we can reconfigure organizational structures or processes to change the communication or collaboration networks within an organization and examine whether the changes in the networks affect the productivity, innovativeness, and performance of the employees. Alternatively, we could also randomly assign students to different kinds of schools and study how different school networks affect student learning. For example, in many school integration programs in which selected students from urban areas are bussed into suburban schools, we could study how the networks of the hosting schools affect the achievement of the incoming students. But a problem with this example is that the enrollment of students itself may change the existing school networks. Moreover even if this were not the case we would not know whether network structure in fact gave rise to different effects or whether

it were instead some other feature associated with network structure. To make valid causal inference on macro network effects, the effects such as endogenous network changes need to be removed or adjusted for.

From a theoretical perspective, several steps would need to be taken to advance research in this area. First, metrics to measure the macro features of networks need to be defined. These can be roughly broken into three categories: (1) cohesion, (2) hierarchy, and (3) clustering. Cohesion can include measures on network density, transitivity, fragmentation (the proportion of subjects that cannot reach one another), heterogeneity, or embeddedness (the proportion of multiplex ties that mix instrumental relationships with integrative relationships). Hierarchy can include measures on degree centralization or network balance (the proportion of mutual ties). Clustering can include measures about the localness or small-world-ness of the networks. Many such measures have been developed in the literature (see, e.g., Wasserman and Faust 1994).

We also need to clarify what kinds of outcomes macro network studies may be relevant to. For example, macro network features can be used to assess how and to what extent network cohesion is related to trust, reciprocity, engagement, social control, and social support in an organization and community. Morgan and Todd (2009) found that intergenerational closure (as measured by the density of parental networks) might increase student achievement in Catholic schools but not in public schools, possibly due to different school norms. On the other hand, Portes (1998) has pointed out four possible facets of the dark side of social capital: exclusion of outsiders, excessive claims of obligations on group members, restrictions on individual freedom, and downward leveling norms. An interesting set of questions would arise from the study of how network features are related to any of these dark sides of social capital and to contrast these effects with the positive features.

Another possible area of research would be to contrast the performance of networks with predominantly bridging ties (i.e., ties connecting distant individuals, groups, or communities) and networks with predominately bonding ties (i.e., ties concentrated within a small group or community). There are a few nonexperimental studies in this area. According to a study by Morgan and Sørensen (1999), public schools featured by horizon-expanding patterns of social relations (i.e., students' teachers and parents do not know each other but predominantly connect to adults outside of the school community) and Catholic schools characterized by norm-enforcing patterns of social relations (i.e., students' teachers and parents know one another well) are more effective in fostering student learning than their counterparts with other kinds of networks. Uzzi and Spiro (2005) found that the effect of the small-world network on the creativity of artists who produced Broadway musicals was parabolic: Too many bonding or bridging ties were both detrimental to the creativity of the artistic teams. Borgatti (2005) studied the relations between network features and individual innovativeness and contended that large diffuse personal networks (i.e., those with a lot of bridging ties) were more conducive for incremental innovations, while radical innovations usually came from people who worked alone or possessed only sparse networks.

Effects on Networks

Formation of social relationships (or social networks) is another area that demands more experimental research. Theoretically, formation of social relationships, for example, friendships, is based on three broad groups of factors: contextual factors, network factors, and individual factors. Contextual factors refer to physical opportunities which determine how likely and frequently subjects can potentially meet or interact, including such factors as neighborhood features, population composition, organizational activities, or institutional rules. Network factors refer to social opportunities to meet or interact with others that arise due to subjects' network positions and the features of the networks they are embedded in. These factors can range from homophily (people tend to associate with others who

are similar to them, as shown by McPherson et al. 2001), reciprocity (people are friends of those who think they are their friends too), transitivity (a friend's friend is also friend), to preferential attachment (popular subjects tend to accumulate even more connections over time). The last group of factors is individual factors including subjects' personal characteristics, family background, or preference for particular types of connections. Dividing these factors into three groups is principally for conceptual purposes. In reality, these three groups may overlap with one another, and some factors may belong to multiple groups at the same time. In theory, we could envision experiments and interventions focused on modifying any of these various contextual, network, or individual factors and defining and assessing causal effects of such interventions on network formation.

Various types of experimental designs could be developed to study network formation. First, as indicated above, one might randomly intervene on specific contextual, network, or individual factors. One could for instance use a design similar to the random-assignment design of social contacts (described above) but now aimed at studying how spatial proximity and social factors like race and gender affect subjects' formation of relationships. For example, Camargo et al. (2010) reported that randomly assigned roommates were likely to become friends regardless of their roommates' race and that white students were likely to have a higher percentage of black friends, if they were assigned with black roommates than with white roommates.

Second, one could also consider the partial treatment group design as described above to study the effect of the treatment intervention (e.g., a smoking prevention program) not only on primary outcomes (e.g., smoking behavior) but also on network structure and the number of ties treatment and control participants in each group have. The intervention may change not only primary outcomes but the structure of the network itself, and both changes may potentially have feedback effects on each other later on. We give an example of this in the following section.

Finally, other types of designs could be used that mix subjects (e.g., mingling subjects with different backgrounds together) to examine how connections are formed among them over time and what factors seem to explain relationship formation. This design is good at assessing homophily or "task complementarity" (i.e., that people tend to form ties with others who have different skills to complete different aspects of a common task) in social networks. Vissa (2011) studied a panel of Indian entrepreneurs who first met each other at a conference and found that they intended to form ties with others who were from the same caste, who spoke the same language, and who had greater task complementarity.

An Integrative Example

In An (2011b), the author used a partial treatment group design to study peer effects on smoking among adolescents in 76 classes of six middle schools. From each grade of the six middle schools, four classes were selected and randomly assigned into four conditions: (1) control condition, in which no students would receive a smoking prevention intervention; (2) random treatment condition, in which a number of randomly selected students were chosen to receive the smoking prevention intervention; (3) central treatment condition, in which a number of central students were selected to receive the intervention; and (4) group treatment condition, in which students and their close friends were chosen to receive the intervention. In all of the treated classes, only 25 % of the students were chosen to receive the intervention.

The results of the experiment indicated that there were peer effects on some (but not all) outcomes of interest. For example, there was evidence of peer effects regarding information exchange or behavior change. Specifically, compared with the students in the control classes, the untreated students in the random treatment condition were more likely to have exchanged information about the intervention with other students and to remain nonsmokers or change to nonsmokers over time.

The evidence for other kinds of network effects comparing the random treatment condition, the central treatment condition, and the group treatment condition was not as clear cut. Compared with the untreated students in the random treatment condition, the untreated students in both the central treatment condition and the group treatment condition had lower intention to smoke next month. But the estimated "peer effects" on three other outcomes of interest, including information exchange, knowledge accumulation (change in knowledge regarding the negative effects of smoking), or behavior change, were indistinguishable between the random intervention and the social network-based interventions, though, as noted above, these comparisons are a combination of causation and selection because the untreated subjects may not be comparable across arms. Even when the effects on both treated and untreated subjects (i.e., the overall intervention effects) are of concern, the evidence on the benefits of social network-based interventions is somewhat ambiguous. The major benefit of the social network-based interventions, over the random treatment intervention, appeared to be reinforcing student intention not to smoke rather than on disseminating information, facilitating knowledge accumulation, or promoting behavior change. For example, compared with the students in the random intervention, fewer students in both the central intervention and the group intervention intended to smoke next month, but they were roughly equally likely or even less likely to remain nonsmokers or change to be nonsmokers.

An unintended but important result associated with the social network-based interventions, however, was that smokers' popularity was decreased to a larger degree in the social network-based interventions than in the random condition – the smokers in the social network-based interventions lost over two friend nominations more than their counterparts in the random condition. This result suggests that the intervention had impact of changing not only students' smoking status but also their networks. Conceptually, we can conjecture that the changed network structure could further affect student smoking behavior and the changed smoking behavior could in turn consolidate the changes in networks, and so forth.

Causal Inference from Observational Social Network Data

Having considered causal inference from randomized experiments on social networks, we now turn to the more controversial area of causal inference from observational social network data. As noted above, with cross-sectional social network data, when outcomes of peers are associated, it is essentially impossible to distinguish whether this is due to social influence, homophily, or environmental confounding. However, an approach based on longitudinal social network data has more recently emerged that is able to at least partially circumvent such problems. As discussed below, even the analyses based on longitudinal data are subject to various limitations and critiques, but such analyses certainly constitute an advance over the use of cross-sectional data. Most prominent among these are a series of analyses conducted by Christakis and Fowler and colleagues (e.g., Christakis and Fowler 2007, 2008; Fowler and Christakis 2008; Cacioppo et al. 2009; Christakis and Fowler 2013) claiming that social influence plays an important role in the spread of a variety health-related attributes, behaviors and psychological states including obesity, smoking, happiness, depression, drug use, and even loneliness. In what follows, we will present the methodology employed by Christakis and Fowler in their analyses, and we will discuss some of the critiques of their approach and responses to these critiques. We will furthermore describe an alternative stochastic actor-oriented approach which also employs longitudinal social network data and more explicitly models changes in network structure.

Longitudinal Analyses of Social Network Data

The longitudinal social network analysis approach of Christakis and Fowler (2007, 2008; Fowler and Christakis 2008; Cacioppo et al. 2009) essentially consists of regressing one individual's (the ego's) state (e.g., obesity) on another's (the alter's) state, along with the alter's lagged state, the ego's lagged state, and other covariates for the ego. Significant association between the ego's state and the alter's state when also controlling for the ego's and alter's lagged state and other variables is then taken as evidence for a contagion effect.

More formally, suppose individual i names individual k as a friend. Let $Y_i(t)$ and $Y_i(t+1)$ denote the ego's outcome at times t and $t+1$, respectively. Let $Y_k(t)$ and $Y_k(t+1)$ denote the alter's outcome at times t and $t+1$, respectively. Let $Z_i(t+1)$ denote the ego's covariates at time t. Christakis and Fowler regress $Y_i(t+1)$ on $Y_i(t)$, $Y_k(t)$, $Y_k(t+1)$, and $X_i(t+1)$ using either repeated measures logistic regression for binary outcomes or repeated measures linear regression for continuous outcomes. The coefficient for $Y_k(t+1)$ in the regression model for $Y_i(t+1)$, which we will call here β, is taken as the contagion effect (i.e., a measure of social influence). Robust standard errors are computed using generalized estimating equations.

Christakis and Fowler argue that with obesity as the outcome, adjusting for the alter's lagged obesity status helps to control for homophily (Christakis and Fowler 2007; Carrington et al. 2005). The reasoning is that the latent factor giving rise to homophily would have to explain both the ego's obesity and the alter's obesity via pathways other than through the alter's lagged obesity for such a factor to generate an association in the absence of genuine social influence. As discussed below, this approach to control for homophily, although partially effective, has come under critique.

Even if we grant adequate control for homophily, interpreting associations, even with longitudinal social network data, as evidence for contagion effects is potentially problematic because of the possibility that a shared environmental factor might in fact affect both the ego's and the alter's state. Christakis and Fowler argue against this as an explanation by noting that the effect estimates for ego-nominated friends are larger than those for alter-nominated friends, which would not occur if the associations were purely due to environmental confounding. This argument, which can be traced back to Duncan et al. (1968) and was formulated in Anagnostopoulos et al. (2008) as the edge-reversal test, may not hold in situations where there is peer selection, sampling errors, or spatial dependence in the network data, points we return to below.

Christakis and Fowler (2007, 2008) likewise undertake similar analyses for other types of social ties beyond ego-nominated friends and alter-nominated friends. They consider similar analyses with mutual friends (person i names person k as a friend, and person k names person i as a friend) and with spouses, neighbors, and siblings.

Using this methodology, Christakis and Fowler report evidence for social influence for smoking, obesity, alcohol consumption, happiness, loneliness, depression, drug use, and so forth (Christakis and Fowler 2007, 2008; Fowler and Christakis 2008; Cacioppo et al. 2009; Christakis and Fowler 2013). Thus, for example, using data from the Framingham Heart Study (Dawber 1980; Feinleib et al. 1975), Christakis and Fowler (2007) found that an individual's chances of being obese (body mass index > 30) increased by 57 % (95 % CI: 6–123 %) if he or she had a friend who was obese in a given interval. In these analyses, they controlled for an ego's age, sex, and education level; the ego's obesity status at the previous time point; and the alter's obesity status at the previous time point. Likewise, using the same data, Christakis and Fowler (2008) report that smoking cessation by a spouse decreased a person's chances of smoking by 67 % (95 % CI: 59–73 %).

Critiques and Responses

These analyses of Christakis and Fowler have come under some criticism. Such critiques have included incorrect estimation of standard errors (Lyons 2011), allegedly similar results using the same methodology for factors such as height, acne, and headaches for which social influence seems much less plausible (Cohen-Cole and Fletcher 2008), inadequate control for homophily (Shalizi and Thomas 2011), changes in friendship structure giving rise to spurious associations (Noel and Nyhan 2011), and issues with model inconsistency (Lyons 2011).

More specifically, Shalizi and Thomas (2011) argue that the possibility of latent (unmeasured) homophily threatens the validity of such longitudinal network analyses. As noted above, homophily refers to the tendency of individuals similar to one another to become friends with each other. It may, for example, be the case that two friends simultaneously become obese not because one influences the other but because they both enjoy excessive eating; this shared interest causes them to become friends and also causes them both to become obese over time. When control is not made for variables responsible for homophily in the analysis, it is difficult to attribute the association to social influence rather than homophily (Shalizi and Thomas 2011). Although control for alter's lagged obesity, as in Christakis and Fowler (2007), arguably does help somewhat, Shalizi and Thomas (2011) persuasively argue that the problem of latent homophily is still present in such analyses. If the latent factor giving rise to the formation of friendship ties affects present obesity even when controlling for past obesity, associations between the ego's and alter's current obesity can arise even when the alter has no social influence on the ego. Shalizi and Thomas (2011) also leverage this point to further critique the argument that Christakis and Fowler use against environmental confounding. Christakis and Fowler argue against environmental confounding as an explanation of their associations by noting that the effect estimates for ego-perceived friends are larger than those for alter-perceived friends, which would not occur if the associations were purely due to environmental confounding. Shalizi and Thomas (2011) show that in the presence of latent homophily, even if there is no unmeasured environmental confounding, the associations comparing ego-perceived friends and alter-perceived friends may differ in magnitude even when there is no social influence. This spurious causation can also arise just because of random sampling error to the extent that the friends group is only a subset of all the subjects (An 2011e). The basic reasoning used by Christakis and Fowler (2007) to argue against environmental confounding in this case breaks down, and we are then left with both latent homophily and environmental confounding as possible explanations of associations.

A somewhat related critique was also put forward by Noel and Nyhan (2011) concerning friendship retention. Through simulations, Noel and Nyhan (2011) show that if friends whose characteristics change to become different from one another are also more likely to end the friendship, then this can also lead to bias and could explain away associations between an ego's and an alter's states. Following the phenomenon on Facebook, they refer to this as the "unfriending" problem. They critique the type of longitudinal social network analysis undertaken by Christakis and Fowler (2007, 2008) on the grounds that such "unfriending" can give rise to spurious associations of the form reported by Christakis and Fowler, even in the absence of social influence.

Yet another important set of critiques was put forward by Lyons (2011). Lyons (2011) argues that when using the repeated measures logistic regression model for binary outcomes used by Christakis and Fowler, if, in the network, there is a person i with a tie to person k and that person k has a tie to person $m \neq i$ then, when using contemporaneous ego and alter data, the models themselves imply that the coefficient for social influence $\beta = 0$; and that similar issues pertain to linear models. The models themselves effectively contradict the existence of the very effect Christakis and Fowler want to assess. The issue raised by Lyons is essentially that there are more equations than unknowns. Intuitively, the problem develops because the same variable at the same time period, for example, the

ego's state at time $t+1$, is the dependent variable in one regression and the independent variable in another regression. Lyons argues that the models themselves then effectively contradict the conjecture of social influence that Christakis and Fowler want to assess.

Lyons (2011) also criticizes the procedures Christakis and Fowler (2007, 2008) use for statistical inference in face of the complex statistical dependence structures that are generated by a social network. Christakis and Fowler (2007, 2008) use a method referred to as generalized estimating equations, clustering on the ego, to take into account the use of multiple time points for the ego. Unfortunately, as Lyons (2011) notes, this is not the only source of dependence in the data. If there is social influence (contagion), then the clusters defined by the ego will not be independent of one another. Moreover, even under the null of no contagion, when contemporaneous ego-alter data is used, the generalized estimating equations standard error is not always valid. In fact, it can be shown that because Christakis and Fowler (2007, 2008) use contemporaneous data for the ego and the alter and because one person's state at time $t+1$ is thus both an outcome in one regression and an independent variable in another, the standard errors for β obtained by Christakis and Fowler (2007, 2008) are too small whenever relationships are reciprocal (e.g., for mutual friends, spouses, siblings, and neighbors; see VanderWeele et al. 2012a).

This array of critiques has shed considerable doubt on the validity of the analyses undertaken by Christakis and Fowler. Although some of these critiques carry substantial weight, some progress has been made in responding to or at least partially circumventing some of these critiques.

For now, let us set aside the issues of homophily and environmental confounding, to which we will return later, and suppose that adequate control has been made for these. As noted above, Lyons argued that the models themselves effectively contradict the conjecture of social influence that Christakis and Fowler want to assess. However, an important exception arises when the null hypothesis of no contagion is in fact true. In this case, provided that homophily and environmental confounding have been properly controlled for, then $\beta = 0$, and if $\beta = 0$, then the models may be correctly specified, provided, for example, for a binary outcome, that the log odds of the ego's state is indeed linear in the covariates. Under the null hypothesis of no contagion, the problem of model inconsistency effectively vanishes. The estimate and confidence interval for β would not constitute a valid estimate of the contagion effect. However, whether the confidence interval for β contains 0 would constitute a valid test of the null hypothesis of no contagion, again provided the assumptions of no homophily and no environmental confounding conditional on the covariates and that of correct model specification with respect to the covariates held (VanderWeele et al. 2012a). Under these assumptions, we can in theory do testing, but not estimation.

This brings us to another critique of Lyons (2011), that of statistical modeling under the dependence structures that are generated by a social network. Even under the null of no contagion, when contemporaneous ego-alter data is used, the generalized estimating equations standard error is not always valid. Because Christakis and Fowler (2007, 2008) use contemporaneous data for the ego and the alter, the standard errors for β obtained by Christakis and Fowler (2007, 2008) are anti-conservative and the confidence intervals will be too narrow whenever relationships are reciprocal, for example, for mutual friends, spouses, siblings, and neighbors (VanderWeele et al. 2012a). However, for the purposes of testing, both the problem of model inconsistency and the problems of statistical dependence and standard error estimation can be easily addressed if the alter's state is lagged by an additional period in the regressions (VanderWeele et al. 2012a). The argument used by Lyons (2011) to show that the models are inconsistent in the presence of contagion is no longer applicable and, under the null of no contagion/social influence, the clusters defined by the ego are independent of one another leading to valid standard errors when using generalized estimating equations (VanderWeele et al. 2012a).

In fact, Christakis and Fowler (2007, 2008) report, in the online supplement to their papers, that they ran such analyses in which the alter's state was lagged by an additional period and that the results

of such analyses were similar to those of their main analyses using contemporaneous data for the ego and alter (i.e., they once again find evidence of significant contagion effects for smoking and obesity).

All of our discussion thus far has assumed that adequate control has been made for homophily and environmental confounding. As noted by Lyons (2011) and by Shalizi and Thomas (2011), this assumption is very strong. To partially circumvent this issue, VanderWeele (2011) showed how sensitivity analysis techniques (see also VanderWeele and Arah 2011; Chap. 18 by Gangl, this volume) could be used to assess the extent to which an unmeasured factor responsible for homophily or environmental confounding would have to be related to both the ego's and the alter's state in order to substantially alter qualitative and quantitative conclusions. The sensitivity analysis technique is applicable to estimates obtained by lagging the alter's state by an additional period (VanderWeele 2011; VanderWeele et al. 2012a). Using the results of Christakis and Fowler (2007, 2008; cf. Fowler and Christakis 2008; Cacioppo et al. 2009), which are reportedly similar to what is obtained from lagged analyses, VanderWeele (2011) used such sensitivity analysis techniques to argue that the evidence reported by Christakis and Fowler (2007, 2008) for obesity among mutual friends and for smoking cessation among spouses was reasonably robust to potential latent homophily or environmental confounding; associations between other types of relational ties for smoking and obesity and those for happiness and loneliness were considerably less robust. The associations reported by Cohen-Cole and Fletcher (2008) concerning acne and headaches were not at all robust to potential latent homophily and environmental confounding; their associations concerning height were somewhat more robust, though their measure of height is self-reported height, and the AddHealth data that they used, unlike the Framingham Heart Study used by Christakis and Fowler, is for adolescents rather than adults. Self-reported height, among adolescents, may perhaps genuinely be subject to social influence (Christakis and Fowler 2013).

The final critique of those discussed above that has not yet been considered is that of the "unfriending" problem (Noel and Nyhan 2011). Noel and Nyhan argued that unfriending results in changes to social network structures that can lead to spurious associations between egos and alters even in the absence of social influence. However, the simulations of Noel and Nyhan also suggest that the degree of these potential biases depends largely on the extent of "unfriending." In the adolescent AddHealth data mentioned above, friendship retention across waves is only about 50 %, and in Noel and Nyhan's simulation, such low retention rates can generate substantial bias. However, in the Framingham Heart Study data used by Christakis and Fowler (2007, 2008), friendship retention is very high, and this unfriending problem does not seem, by Noel and Nyhan's own simulations, sufficiently common to result in substantial biases in the analyses of Christakis and Fowler (2007, 2008).

Considerable methodological development still needs to be done concerning such longitudinal social network analyses, perhaps especially in deriving valid estimators of the standard error that are not only applicable under the null hypothesis of no contagion but also in the presence of social influence. However, as argued here, a number of the existing critiques of previous longitudinal network analysis have at least partially been addressed.

An Alternative Stochastic Actor-Oriented Model

The approach employed by Christakis and Fowler (2007, 2008) makes use of well-established methods for longitudinal data and attempts to control for confounding and homophily by covariate control and the use of lagged states, but it does not explicitly model the mechanism for the selection of social ties itself. An alternative stochastic actor-oriented model (SAOM) has been developed (Snijders 2001, 2005; Steglich et al. 2010) that models social influence and the selection of social ties jointly.

The model assumes that at each instant an individual may either change the behavior/state under study or change a particular social tie. Such changes occur on the network with specific rates, which may vary across individuals. In the models, such as Steglich et al. (2010), these events are assumed to follow an exponential distribution. The rates may depend on the states of the individuals on the network, the existing network structure itself, individual actor-level covariates, or dyad-level covariates. Likewise, when an event occurs, the actual changes to either the behavior state or a social tie may depend on the states of the individuals on the network, the existing network structure itself, individual actor-level covariates, or dyad-level covariates and are also subject to random fluctuation. The magnitude of the effect of each of these components is fit with data.

Such stochastic actor-oriented models are typically too complex to allow for closed-form expressions for the probabilities of particular transitions and thus to employ traditional maximum-likelihood procedures. The models are instead fit with simulation techniques such as Markov chain Monte Carlo. Fitting such models can be computationally demanding which can limit the sample size to which the models can be employed. The fitting procedures of such models can also sometimes fail to converge.

These stochastic actor-oriented models are appealing in that they involve parameters corresponding to both social influence and homophily. They do, however, rely on stronger modeling assumptions. Modeling assumptions need to be made not only regarding the behavior states themselves, as in the Christakis-Fowler approach, but also for processes by which there are changes in social ties. This challenge may also be an advantage, however, in settings in which individuals change social ties frequently as in the AddHealth data mentioned above, in contrast to the Framingham Heart Study data. But the assumption that there is only one single change allowed at any particular instant might be problematic in cases where subjects make simultaneous changes in both behavior and social ties. For example, in the case of smoking, it is likely that some subjects sever their ties to smokers while stopping smoking at the same time. It is unclear how these models perform in the context of simultaneous changes.

Another strong assumption made by the SAOM is that actors have full information of their local networks. This may be true for some actors but will most likely not be true for all actors due, for example, to cognitive constrains (e.g., actors cannot recall their contacts correctly). One of the stronger assumptions made by the SAOM is that actors' decisions to make changes in either behavior or social ties are not reactive (i.e., not taking into account other actors' potential reactions). This is particularly problematic in business settings where strategic behaviors and alliances are mostly reactive. A game-theoretic strategic element may enhance the performance of SAOM in such settings. One other strong assumption made by the SAOM is that the coevolution of network and behavior follows a Markov process in which only the immediate past states matter. This assumption greatly simplifies the mathematical derivations behind it but would be in tension with social relationships and social behaviors that are "sticky" and may have longer lasting impact on future network and behavior change. Whether it is important and how to incorporate these features into SAOM merits further research. Finally, the stochastic actor-oriented models are also subject to the same limitations concerning environmental confounding as the longitudinal models of Christakis and Fowler. This issue of environmental confounding needs to be critically investigated and evaluated within the SOAM context.

The use of a SAOM is also appealing in that it provides users with a substantial degree of flexibility to choose all kinds of parameters to specify their models. But this can pose challenges to investigators as well, as it is not clear which sets of parameters the users should choose and it can be difficult to select a final set of parameters through a trial-and-error process. Model selection tools could improve the usefulness of the SAOM.

Further work could better consider the relationship and relative advantages and disadvantages of these different approaches to handling observational longitudinal social network data.

Future Directions for Causal Inference and Social Networks

There is considerable scope for further methodological development, and for further empirical and experimental studies, related to causal inference and social networks. Work remains to be done in providing a more rigorous foundation for causal inference from observational social network data. As was noted above, in the longitudinal social network analyses currently being employed, the estimation procedures used for statistical inference are often valid only for testing, not estimation. Further work is required in developing variance estimators that are applicable in the presence of social influence and when statistical dependence may be present between the states of all individuals within a social network. Possible alternative methods that can be useful in such efforts include instrumental variable regressions, regression discontinuity designs, and adaption of existing spatial regressions (e.g., Lee 2009). As has been noted, even in a simpler setting with well-defined clusters, formal statistical inference in the presence of social influence can be challenging (Hudgens and Halloran 2008; Tchetgen Tchetgen and VanderWeele 2012), and these issues are likely to be further complicated within the context of a social network.

At a more conceptual level, further theoretical development remains to be done in attempting to formulate longitudinal social network analyses within a counterfactual framework. Within the counterfactual or potential outcomes framework, causation is generally conceived in terms of counterfactual contrasts, and the counterfactuals are themselves generally tied to hypothetical interventions. Within the context of studying the possibility of social influence in a social network of a state, such as obesity, say, it is not entirely clear how to appropriately tie discussion of causation to such hypothetical interventions or what such hypothetical interventions might be. Moreover, different possible interventions (e.g., exercise, diet, or liposuction) may have different effects in terms of influencing other individuals within the network (cf. Hernán and VanderWeele 2011). One of the key advantages of experimental studies of interventions on social networks is that it is clear what it is that the causal effects estimated actually correspond to. With observational social network data concerning a particular trait, state, or characteristic, the "exposure" of interest is often not well-defined and does not necessarily clearly correspond to a particular intervention. Such issues pertain to both the longitudinal models of Christakis and Fowler (2007, 2008) and the stochastic actor-oriented models of Snijders (2001, 2005) and Steglich et al. (2010). Work could also be done formalizing the confounding/selection assumptions required to give the parameters of these models a causal interpretation. See Shalizi and Thomas (2011) who provide a preliminary sketch at formalizing aspects of the longitudinal models of Christakis and Fowler (2007, 2008) in terms of causal diagrams.

Further work also remains to be done in better explicating the relationship between the phenomenon sometimes referred to as "interference" on the one hand and that which is often called "contagion" on the other. The former term is generally used for settings in which the exposure of one individual may influence the outcomes of another, whereas "contagion" is typically used for the phenomenon whereby the outcome or state of one individual influences the same outcome or state of another individual. Contagion may be one mechanism by which interference occurs. Consider a study of obesity in which a particular weight loss intervention is assigned to certain persons within a network. The intervention may affect other persons to whom the intervention was not explicitly assigned in at least two distinct ways. First, information from the weight loss intervention may be passed from one individual to another, leading to weight loss even among those to whom the intervention was not explicitly assigned. Second, the intervention may lead to weight loss for those to whom it was explicitly assigned, and such weight loss may influence the norms of other persons to whom the intervention was not assigned, motivating them also to lose weight. The second mechanism might be conceived of as one of "contagion," whereas the first might be conceived of as one of "direct interference" (i.e., "direct" with respect to, not through, the obesity state of the person assigned the

intervention). Further work remains to be done in better explicating and formalizing the relationship between these concepts. See VanderWeele et al. (2012b) for one such formalization in the infectious disease context.

Generalizing the stochastic actor-oriented models for dynamic analysis of weighted networks is another important area that calls for further research, as most of previous network methods are developed to study relationship ties that are binary (present versus absent), while in reality, it may be that the strength of relationships, the intensities of interactions, and the depth of attachment, rather than just the existence of social ties, matter for the outcomes of interest. Such a generalization would help to draw more precise causal inference about social networks. The work of Desmarais and Cranmer (2012) on generalizing the exponential random graph models to study static weighted networks could be illuminating in this line of research.

Causal inference in multiplex networks is another area that demands more research. Current techniques in social network analysis usually focus on modeling only one network or one kind of relationship at a time. But people are embedded in multiple networks with multiplex relationships. The neglect of this may often lead to obscured inferences about causation. One research direction in this area would be to develop statistical methods to study how multiple types of intimate relationships simultaneously affect subjects' outcomes.

Social network data does not only create challenges for causal inference but gives rise to a number of opportunities for better modeling and more precise inferences. Much of the causal inference literature on interference and social influence treats all individuals within a cluster as essentially exchangeable concerning any particular individual's outcome (Hong and Raudenbush 2006; Hudgens and Halloran 2008; Tchetgen Tchetgen and VanderWeele 2012). Social network data gives the investigator the opportunity to more realistically and accurately model the relationships between one individual's outcome and the exposures or states of the other individuals in the group. Moreover, as was noted above, social network data may also be useful in determining on which person's interventions may be most effective for achieving changes in outcomes across an entire social network.

Better understanding of the generation process of social network data is another area that demands further research. Most studies have assumed that subjects report their social connections correctly and completely upon inquiry. But it is possible that some social ties are not reported and are not missing at random. In particular, it is likely that some subjects may purposely under report certain kinds of social ties. This might be apparent in studies of peer effects on delinquent behaviors among adolescents, such as smoking and drinking – smokers may purposefully under report their connections to other smokers and over report their connections to nonsmokers, thus diluting estimates of peer effects. If this is true, the resulting data will make it very difficult to reject the null that there are no peer effects.

Within the context of randomized experiments on social networks, careful thought also needs to be given to the relationship between the effects of interventions on outcomes directly and their effects through changing the network itself. Methodologically, in some cases, networks may be only an intermediary factor, which, like a catalyst, needs to interact with other factors to realize its effects and so may not be the ultimate causes of the outcomes that we are interested in. In addition, even though networks may play an important role in determining some outcomes of interest, it may be the case that effective interventions do not lie in changing the networks, either because it is difficult, not cost-effective, too time-consuming, or unethical to do so. For example, to improve the social life in a ghetto, a more effective strategy than attempting to change social ties might be to provide more resources to the community. Then both social networks and the social behaviors of the residents in the community might change in response.

Considerable methodological development for randomized network experiments will also be important. Several methodological issues which have been addressed within the context of individually randomized experiments remain to be worked out within the social network setting. Such issues include nonrandom attrition or addition of subjects, treatment compliance problems, a counterfactual-based formulation of instrumental variable estimands, measurement error, and missing data problems.

Another area that calls for more experimental study is the coevolution of networks and behaviors. The work done by Martin Nowak and colleagues (Ohtsuki et al. 2006; Pacheco et al. 2006) on the coevolution of networks and cooperation laid a good foundation for this (e.g., Ohtsuki et al. 2006; Pacheco et al. 2006). Their work could be extended in three respects. First, outcomes other than just cooperation could be studied. Second, field experiments, rather than just lab experiments, could be conducted to evaluate the robustness of the results found in lab experiments. Third, the size of networks could be allowed to change over time.

Many other aspects of network causation have received limited empirical and experimental study. As noted above, there has been limited empirical and experimental study of positional effects and structural effects. Further work is also needed in developing specific experimental designs well suited to answer particular network questions and in understanding which designs are best suited and most efficient in addressing which class of questions or which designs can be implemented in a cost-effective manner. Internet- or lab-based experiments in which investigators themselves fix the network structure may prove an efficient and cost-effective means for studying network phenomenon (Centola 2010; Fowler and Christakis 2010). In addition, more experimental network research could be done on the booming social networking sites like Twitter and Facebook. The scope for further theoretical and empirical advancement concerning causal inference and social networks is indeed considerable.

Acknowledgements The authors thank Stephen Morgan for helpful comments. This work was supported by NIH grant ES017876.

References

An, W. (2011a). Models and methods to identify peer effects. In J. Scott & P. J. Carrington (Eds.), *The Sage handbook of social network analysis* (pp. 514–532). London: The Sage Publications.

An, W. (2011b). *Peer effects on adolescent smoking and social network-based interventions*. PhD dissertation, Department of Sociology, Harvard University.

An, W. (2011c). *Instrumental variable estimates of peer effects*. Working paper, Department of Sociology, Harvard University.

An, W. (2011d). *Algorithms for social network-based interventions and policies*. Working paper, Department of Sociology, Harvard University.

An, W. (2011e). *On the directionality test of peer effects*. Working paper, Department of Sociology, Harvard University.

Anagnostopoulos, A., Kumar, R., & Mahdian, M. (2008). Influence and correlation in social networks. In *Proceeding of the 14th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 7–15). New York: ACM.

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton: Princeton University Press.

Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks, 29*, 555–564.

Borgatti, S. P. (2005). *Creating knowledge: Network structure and innovation*. Available at http://www.socialnetworkanalysis.com/knowledge_creation.htm

Borgatti, S. P. (2006). Identifying sets of key players in a network. *Computational, Mathematical and Organizational Theory, 12*(1), 21–34.

Bramoullé, Y., Djebbari, H., & Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics, 150*(1), 41–55.

Burt, R. S. (1987). Social contagion and innovation: Cohesion versus structural equivalence. *American Journal of Sociology, 92*(6), 1287–1335.

Burt, R. S. (1995). *Structural holes: The social structure of competition*. Cambridge: Harvard University Press.

Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology, 110*, 349–399.

Cacioppo, J. T., Fowler, J. H., & Christakis, N. A. (2009). Alone in the crowd: The structure and spread of loneliness in a large social network. *Journal of Personality and Social Psychology, 97*(6), 977–991.

Callahan, O. D., & Robin, S. S. (1969). A social system analysis of preferred leadership role characteristics in high school. *Sociology of Education, 42*(3), 251–260.

Camargo, B., Stinebrickner, R., & Stinebrickner, T. R. (2010). *Interracial friendships in college* (NBER Working Paper No. 15970). http://www.nber.org/papers/w15970.pdf

Campbell, R., Starkey, F., Holliday, J., Audrey, S., Bloor, M., Parry-Langdon, N., Hughes, R., & Moore, L. (2008). An informal school-based peer-led intervention for smoking prevention in adolescence (ASSIST): A cluster randomised trial. *Lancet, 371*, 1595–1602.

Carrington, P. J., Scott, J., & Wasserman, S. (2005). *Model and methods in social network analysis*. New York: Cambridge University Press.

Centola, D. (2010). The spread of behavior in an online social network experiment. *Science, 329*, 1194–1197.

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.

Christakis, N. A., & Fowler, J. H. (2008). The collective dynamics of smoking in a large social network. *New England Journal of Medicine, 358*, 2249–2258.

Christakis, N. A., & Fowler, J. H. (2013). Social contagion theory: Examining dynamic social networks and human behavior. *Statistics in Medicine 32*(4), 556–577.

Cipollone, P., & Rosolia, A. (2007). Social interactions in high school: Lessons from an earthquake. *The American Economic Review, 97*(3), 948–965.

Cohen-Cole, E., & Fletcher, J. M. (2008). Detecting implausible social network effects in acne, height, and headaches: Longitudinal analysis. *British Medical Journal, 337*, a2533.

Copic, J., Jackson, M. O., & Kirman, A. (2009). Identifying community structures from network data via maximum likelihood methods. *The B.E. Journal of Theoretical Economics*, *9*(1), Article 30.

Cornwell, B. (2009). Good health and the bridging of structural holes. *Social Networks, 31*, 92–103.

Cox, D. R. (1958). *The planning of experiments*. New York: Wiley.

Dawber, T. R. (1980). *The Framingham study: The epidemiology of atherosclerotic disease*. Cambridge: Harvard University Press.

Desmarais, B. A., & Cranmer, S. J. (2012). Statistical inference for valued-edge networks: The generalized exponential random graph model. *PLoS ONE, 1*(7), e30136.

Duflo, E., & Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *The Quarterly Journal of Economics, 118*(3), 815–842.

Duncan, O. D., Haller, A. O., & Portes, A. (1968). Peer influences on aspirations: A reinterpretation. *The American Journal of Sociology, 74*(2), 119–137.

Falk, A., & Ichino, A. (2006). Clean evidence on peer effects. *Journal of Labor Economics, 24*(1), 39–56.

Feinleib, M., Kannel, W. B., Garrison, R. J., McNamara, P. M., & Castelli, W. P. (1975). The Framingham offspring study: Design and preliminary data. *Preventive Medicine, 4*, 518–525.

Fowler, J. H., & Christakis, N. A. (2008). Estimating peer effects on health in social networks. *Journal of Health Economics, 27*(5), 1386–1391.

Fowler, J. H., & Christakis, N. A. (2010). Cooperative behavior cascades in human social networks. *PNAS: Proceedings of the National Academy of Sciences, 107*, 5334–5338.

Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of National Academy of Sciences of the United States of America, 99*(12), 7821–7826.

Goldenberg, A., Zheng, A. X., Fienberg, S. E., & Airoldi, E. M. (2009). A survey of statistical network models. *Foundations and Trends in Machine Learning, 2*, 129–233.

Graham, B. (2008). Identifying social interactions through conditional variance restrictions. *Econometrica, 76*, 643–660.

Halloran, M. E., & Struchiner, C. J. (1991). Study designs for dependent happenings. *Epidemiology, 2*, 331–338.

Hernán, M. A., & VanderWeele, T. J. (2011). Compound treatments and transportability of causal inference. *Epidemiology, 22*, 368–377.

Hudgens, M. G., & Halloran, M. E. (2008). Towards causal inference with interference. *Journal of the American Statistical Association, 103*, 832–842.

Hong, G., & Raudenbush, S. W. (2006). Evaluating kindergarten retention policy: A case study of causal inference for multilevel observational data. *Journal of the American Statistical Association, 101*, 901–910.

Kremer, M., & Levy, D. (2008). Peer effects and alcohol use among college students. *Journal of Economic Perspectives, 22*(3), 189–206.

Lee, L.-f. (2009). Identification and estimation of spatial econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics, 140*(2), 333–374.

Lyons, R. (2011). The spread of evidence-poor medicine via flawed social-network analyses. *Statistics, Politics and Policy*, *2*(1), Article 2, 1–26.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies, 60*, 531–542.

Manski, C. F. (2013). Identification of treatment response with social interactions. *The Econometric Journal 16*(1), S1–S23.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology, 27*, 415–444.

Moody, J. (2001). Peer influence groups: Identifying dense clusters in large networks. *Social Networks, 23*, 261–283.

Morgan, S. L., & Sørensen, A. B. (1999). Parental networks, social closure, and mathematics learning: A test of Coleman's social capital explanation of school effects. *American Sociological Review, 64*, 661–681.

Morgan, S. L., & Todd, J. J. (2009). Intergenerational closure and academic achievement in high school: A new evaluation of Coleman's conjecture. *Sociology of Education, 82*(July), 267–286.

Noel, H., & Nyhan, B. (2011). The 'unfriending' problem: The consequences of homophily in friendship retention for causal estimates of social influence. *Social Networks, 33*, 211–218.

Ohtsuki, H., Hauert, C., Lieberman, E., & Nowak, M. A. (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature, 441*, 502–505.

O'Malley, A. J., Elwert, F., Rosenquist, J. N., Zaslavsky, A. M., & Christakis, N. A. (2013). *Estimating peer effects in longitudinal dyadic data using instrumental variables* (Working Paper). Department of Health Care Policy, Harvard Medical School.

Pacheco, J. M., Traulsen, A., & Nowak, M. A. (2006). Coevolution of strategy and structure in complex networks with dynamical linking. *Physical Review Letters, 97*(25), 258103.

Podolny, J. M. (2001). Networks as the pipes and prisms of the market. *American Journal of Sociology, 107*, 33–60.

Portes, A. (1998). Social capital: Its origins and applications in modern sociology. *Annual Review of Sociology, 24*, 1–24.

Rosenbaum, P. R. (2007). Interference between units in randomized experiments. *Journal of the American Statistical Association, 102*, 191–200.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology, 66*, 688–701.

Rubin, D. B. (1980). Comment on: 'Randomization analysis of experimental data in the Fisher randomization test' by D. Basu. *Journal of the American Statistical Association, 75*, 591–593.

Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics, 116*, 681–704.

Shalizi, C. R., & Thomas, A. C. (2011). Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods and Research, 40*, 211–239.

Snijders, T. A. B. (2001). The statistical evaluation of social network dynamics. *Sociological Methodology, 31*, 361–395.

Snijders, T. A. B. (2005). Models for longitudinal network data. In P. J. Carrington, J. Scott, & S. S. Wasserman (Eds.), *Models and methods in social network analysis*. New York: Cambridge University Press. Chap. 11.

Sobel, M. E. (2006). What do randomized studies of housing mobility demonstrate? Causal inference in the face of interference. *Journal of the American Statistical Association, 101*, 1398–1407.

Steglich, C. E., Snijders, T. A., & Pearson, M. (2010). Dynamic networks and behavior: Separating selection from influence. *Sociological Methodology, 40*, 329–393.

Tchetgen, T., Eric, J., & VanderWeele, T. J. (2012). On causal inference in the presence of interference. *Statistical Methods in Medical Research – Special Issue on Causal Inference, 21*, 55–75.

Uzzi, B., & Spiro, J. (2005). Collaboration and creativity: The small world problem. *American Journal of Sociology, 111*(2), 447–504.

Valente, T. W. (2005). Network models and methods for studying the diffusion of innovations. In P. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis* (pp. 98–116). New York: Cambridge University Press.

Valente, T. W., & Davis, R. L. (1999). Accelerating the diffusion of innovations using opinion leaders. *The ANNALS of the American Academy of Political and Social Science, 566*, 55–67.

Valente, T. W., & Pumpuang, P. (2007). Identifying opinion leaders to promote behavior change. *Health Education and Behavior, 34*, 881–896.

Valente, T. W., Hoffman, B. R., Ritt-Olson, A., Lichtman, K., & Johnson, C. A. (2003). Effects of a social-network method for group assignment strategies on peer-led tobacco prevention programs in schools. *American Journal of Public Health, 93*(1), 1837–1843.

VanderWeele, T. J. (2011). Sensitivity analysis for contagion effects in social networks. *Sociological Methods and Research, 40*, 240–255.

VanderWeele, T. J., & Arah, O. A. (2011). Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments and confounders. *Epidemiology, 22*, 42–52.

VanderWeele, T. J., & Tchetgen Tchetgen, E. J. (2011). Effect partitioning under interference for two-stage randomized experiments. *Statistics and Probability Letters, 81*, 861–869.

VanderWeele, T. J., Ogburn, E. L., & Tchetgen Tchetgen, E. J. (2012a). Why and when "flawed" social network analyses still yield valid tests of no contagion. *Statistics, Politics, and Policy, 3*, Article 4, 1–11.

VanderWeele, T. J., Tchetgen Tchetgen, E. J., & Halloran, M. E. (2012b). Components of the indirect effect in vaccine trials: Identification of contagion and infectiousness effects. *Epidemiology, 23*, 751–761.

Vissa, B. (2011). A matching theory of entrepreneurs' tie formation intentions and initiation of economic exchange. *Academy of Management Journal, 54*(1), 137–158.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. New York: Cambridge University Press.

Wing, R. R., & Jeffery, R. W. (1999). Benefits of recruiting participants with friends and increasing social support for weight loss and maintenance. *Journal of Consulting and Clinical Psychology, 67*(1), 132–138.