An exploratory comparison of name generator content: Data from rural India

Holly B. Shakya [a,b,*], Nicholas A. Christakis [b,*], James H. Fowler [a,c]

[a] School of Medicine, University of California San Diego, 9500 Gilman Dr., La Jolla, CA 92093, United States
[b] Yale Institute for Network Science, PO Box 208263, New Haven, CT 06520, United States
[c] Political Science Department, University of California San Diego, SSB 301, 9500 Gilman Dr., La Jolla, CA 92093, United States

ABSTRACT

Since the 1970s sociologists have explored the best means for measuring social networks, although few name generator analyses have used sociocentric data or data from developing countries, partly because sociocentric studies in developing countries have been scant. Here, we analyze 12 different name generators used in a sociocentric network study conducted in 75 villages in rural Karnataka, India. Having unusual sociocentric data from a non-Western context allowed us to extend previous name generator research through the unique analyses of network structural measures, an extensive consideration of homophily, and investigation of status difference between egos and alters. We found that *domestic interaction* questions generated networks that were highly clustered and highly centralized. Similarity between respondents and their nominated contacts was strongest for gender, caste, and religion. We also found that *domestic interaction* name generators yielded the most homogeneous ties, while *advice* questions yielded the most heterogeneous. Participants were generally more likely to nominate those of higher social status, although certain questions, such as who participants *talk to* uncovered more egalitarian relationships, while other name generators elicited the names of social contacts distinctly higher or lower in status than the respondent. Some questions also seemed to uncover networks that were specific to the cultural context, suggesting that network researchers should balance local relevance with global generalizability when choosing name generators.

© 2016 Published by Elsevier B.V.

1. Introduction

Since the 1970s sociologists have explored the best means for measuring social networks. While survey questions are most commonly used, many scholars have experimented with different measurement tools, searching for the most valid and reliable methods including the reverse small-world technique and various forms of personal diaries, including a smart phone app that allows participants to enter social interactions in real time (Bernard et al., 1987, 1990; Fu, 2005; Lerner et al., 2014). The appropriate technique, however may depend upon the type of network study being conducted and the research question being asked (Knipscheer and Antonucci, 1990).

Sociocentric studies focus on a small population and attempt to ascertain all of the social relationships within a set of interconnected individuals (Marin and Wellman, 2011). In the case of such studies, it is important to be able to accurately connect nominated individuals in order to analyze the greater structure of the network. For large-scale sociocentric data collection efforts, the most practical means of eliciting the names of social contacts, therefore, may by through the administration of surveys that are administered uniformly to the entire population.

Surveys, of course, are composed of questions. The network data collection procedure most frequently used in the context of a survey is to simply ask participants (*egos*) the name of those people with whom they have social connections (*alters*) (Burt, 1984; Marin, 2004). This kind of question is called a "name generator". Although this method has its drawbacks, it is generally reliable, and is more efficient than other methods (Bien et al., 1991). The question can be hypothetical (with whom would you do something) or factual (with whom have you done something) (De Lange et al., 2004). While the way in which a question is asked is important in terms of eliciting network ties, the most crucial component is the content of

* Corresponding author at: University of California San Diego, 9500 Gilman Drive, Mail Code 0507, La Jolla, CA 92093-0041, United States.
E-mail addresses: hshakya@ucsd.edu (H.B. Shakya), nicholas.christakis@yale.edu (N.A. Christakis), jfowler@ucsd.edu (J.H. Fowler).

the question itself (Ferligoj and Hlebec, 1999). The ties elicited by the name generators create the structure of the network, and the specific questions asked to elicit those ties provide the context.

While friendship, as a social phenomenon, occurs throughout the world, the actions that define friendship may differ greatly (Hruschka, 2010). Fischer, for instance, determined that friendship in America is vaguely defined, although generally it refers to people to whom no other specific title, such as co-worker or relative, can be given (Fischer, 1982). Name generator questions, therefore, usually focus on the specific context of relationships. The content of the question determines the type of relationship depicted, which is a crucial component of understanding the significance of the network itself. Network contexts can be categorized as exchange (people with whom an ego engages in reciprocal service provision such as borrowing and lending money); role relation (specific relationships such as spouse or mother); interactive (people with whom an ego interacts during the day); and affective (people with whom an ego shares strong emotional bonds) (Knipscheer and Antonucci, 1990; Marin and Hampton, 2007; Van der Poel, 1993). Exchange, interactive, and role relation questions tend to elicit the largest networks, while affective networks the smallest but are comprised of the closest relationships. Obviously, there is likely to be some degree of overlap between any of these categories.

Ideally, researchers would have the time and resources to collect data on every possible kind of relationship for every person within a given network. However, there are resource limitations both in the ability of most researchers to collect such an exhaustive amount of data and in the ability of the respondents to enumerate social contacts to that extent. For example, respondents might experience survey fatigue, and begin to increasingly underreport alters with each additional name generator question (Pustejovsky and Spillane, 2009). A critical decision faced by any network researcher, then, is how many name generators should be used. If there are too many questions, researchers may be collecting redundant information. De Lange and colleagues analyzed the results of name generator questions administered to employees in a small organization in Belgium, and they were able to uncover three primary conceptual factors as well as nomination overlap in a group of questions regarding relations within the organization (De Lange et al., 2004). This information provided crucial clues regarding what questions could potentially be cut from an instrument with multiple questions. The validity of network data that uses too few name generators, on the other hand, may be compromised by the fact that respondents tend to underreport their important ties, and instruments using only one name generator are not sensitive enough to capture the size and complexity of the real network (Burt, 1997; Marin, 2004; Marin and Hampton, 2007). For example, while one study in the US has shown that the sex and race of nominated alters are relatively invariant across name generators, other work has shown significant gender variation (Campbell and Lee, 1991; Ruan, 1998; Straits, 2000). Two of these studies also showed significant differences in tie characteristics, i.e. the strength and type of relationship (Campbell and Lee, 1991; Ruan, 1998). While McCallister and Fischer were the first to use multiple name generators to ascertain a more comprehensive and multi-dimensional network in 1978 (McCallister and Fischer, 1978), many network studies still limit the number of name generators to 1 or 2 general purpose name generators. One of the most commonly used is "with whom do you discuss important matters?".

An important shortcoming of using one general purpose name generator is the assumption that it can be correlated with diverse outcomes, such as finding a job, as the pathway by which this is supposed to occur is by no means clear (Parigi and Bearman, 2005). The *functional specificity hypothesis* is based upon the belief that individuals rely on different people for different types of support depending upon their need (Perry and Pescosolido, 2010).

A person with whom someone discusses politics may not be the person upon whom they rely for assistance with a sick child. Wellman and Worley, for instance, show that the types of social support given by a person's "strong ties" differs substantially from that given between parents and children (Wellman and Wortley, 1990). Milardo demonstrates that what he terms "significant other" relationships (which in his operationalization includes both close ties and kin) do not overlap significantly with exchange networks (Milardo, 1989, 1992). He also emphasizes the importance of understanding the difference in the structural elements of these different types of networks if we are to attribute significance to network structure overall. Given this viewpoint, the "important matters" name generator (or any general sort of name generator focused on non-specific social interaction) could be insufficient for eliciting ties that might be influential in a variety of contexts. Perry and colleagues (Perry and Pescosolido, 2010) show there is a significant difference between the "important matters" discussion network and a "health matters" discussion network both in structure and ability to predict relevant health outcomes (the health matters network could predict health outcomes in this study but not the important matters network). Researchers in Mali collected network data on women from two different ethnic groups using questions to elicit support networks on four different dimensions: emotional, cognitive (meaning information sharing), material, and practical (help with childcare, etc.) (Adams et al., 2002). Their research demonstrated that the correlation between network composition and child mortality differed between network types as well as across ethnic groups. Ruan's study (Ruan, 1998) showed that there was considerable overlap between the names of people generated by the important matters question and those with whom the respondent socializes, but not those from whom the respondent expects help or goes to regarding family matters.

One of the main purposes of applied network research is to understand how behaviors cluster and shift within communities. By using network methods to implement interventions, researchers and program specialists can increase their reach and impact. Often this approach involves strategies to magnify possible social effects so that a behavior adopted by one person spreads to others in her social network. In order to most successfully exploit these potential social effects, it is important that network researchers use questions that are the most likely to elicit the network in which these effects occur.

In some contexts this may be best achieved by finding individuals with whom participants are similar, as individuals may be more likely to adopt a new behavior if that behavior has already been adopted by someone with whom she is similar (Centola, 2011). In other contexts this may be best achieved by finding individuals with whom participants are different. High levels of *homophily* (similarity between socially connected individuals) can cause norms to become entrenched making social change difficult. For instance, past research has suggested that smoking behavior may become solidified within small pockets of smokers as a result of increasing anti-smoking sentiment (Christakis and Fowler, 2008). In contrast, in other contexts, lower status individuals tend to emulate higher status individuals (DiMaggio and Garip, 2011), suggesting that it may be important to elicit network ties between individuals of different status. Often behavior change interventions will utilize a "peer educator" model, in which highly connected, high status individuals are educated to disseminate new ideas within their focal communities (Valente, 2012). Recent research has also suggested that targeting friends of friends might be beneficial, given that nominated individuals tend to be more central than those that nominated them (Cobb et al., 2010; Kim et al., 2015). These sorts of strategies require an approach that goes beyond considering just network size, and that takes into consideration the function of the name generators themselves. Nevertheless, few studies have