American Sociological Review
78(3) 339–360

DOI: 10.1177/0003122413482919
http://asr.sagepub.com
SAGE

Social Isolation in America: An Artifact

Anthony Paik[a] and Kenneth Sanchagrin[a]

Abstract

This article examines whether existing estimates of network size and social isolation, drawn from egocentric name generators across several representative samples, suffer from systematic biases linked to interviewers. Using several analytic approaches, we find that estimates of network size found in the 2004 and 2010 General Social Surveys (GSS), as well as other representative samples, were affected by significant interviewer effects. Across these surveys, we find a negative correlation between interviewer effects and mean network size. In the 2004 GSS, levels of social connectivity are strongly linked to interviewer-level variation and reflect the fact that some interviewers obtained highly improbable levels of social isolation. In the 2010 GSS, we observe larger interviewer effects in two versions of the questionnaire in which training and fatigue effects among interviewers were more likely. Results support the argument that many estimates of social connectivity are biased by interviewer effects. Some interviewers' failure to elicit network data makes inferences, such as the argument that networks have become smaller, an artifact. Overall, this study highlights the importance of interviewer effects for network data collection and raises questions about other survey items with similar issues.

Keywords

social isolation, social networks, interviewer effects, survey methods

Is social connectedness in the United States truly declining? Since McPherson, Smith-Lovin, and Brashears (2006) reported that Americans' discussion networks had shrunk, on average, from three partners in 1985 to two in 2004, and that social isolation more than doubled from 10 to 23 percent, a spate of articles has examined the social connectedness of adults in the United States. These additional estimates vary widely, with reported mean size of discussion networks ranging from 1.9 to 3.6 alters and social isolation falling between 2 and 12 percent (Brashears 2011; Cornwell, Laumann, and Schumm 2008; Cornwell et al. 2009; Cornwell and Waite 2009; Hampton, Sessions, and Her 2011; Marsden 2003; Wang and Wellman 2010). Indeed, many scholars, including McPherson and colleagues (2006), have been circumspect, suggesting that the shrunken-networks finding based on the 2004 General Social Survey (GSS) might be an artifact (Cornwell et al. 2009; Fischer 2009, 2011; Hampton et al. 2011). Despite the significant amount of time and resources devoted to this issue, however, the question of whether the shrunken-networks finding is an artifact remains unresolved.

Obtaining accurate estimates of social connectivity is more than just an academic exercise. Today, a key concern, even "panic," according to Wang and Wellman (2010:1148),

[a]University of Iowa

Corresponding Author:
Anthony Paik, Department of Sociology, University of Iowa, W140 Seashore Hall, Iowa City, IA 52242
E-mail: anthony-paik@uiowa.edu

is whether information technologies, such as the Internet, cell phones, and social media, have displaced face-to-face conversations and activities in favor of computer-mediated interaction. Although these fears appear to be overblown—information technologies appear to be *positively* associated with network size and interaction (Hampton et al. 2011; Hampton et al. 2009; Wang and Wellman 2010)—the ongoing moral panic highlights the significance of social connectivity for both sociologists and the public alike. Yet the strategy of generating additional point estimates of social connectivity may be flawed if the data on which these are based are biased as well. Thus, despite burgeoning interest, the substantial variability in estimates of personal network size and social isolation suggests it is not clear whether levels of social connectedness are stable, changing, or even reliable.

In this study, we examine whether a significant amount of the variation found in recent estimates of social connectivity can be attributed to a previously unexamined factor: measurement error at the interviewer level. Recent studies of core discussion networks (e.g., Brashears 2011; Hampton et al. 2011; Hampton et al. 2009; McPherson et al. 2006) focus exclusively on how egocentric name generators may produce measurement error related to respondents (Bailey and Marsden 1999; Bearman and Parigi 2004; Brewer and Webster 1999; Marin 2004; Pustejovsky and Spillane 2009; Ruan 1998; Straits 2000). The issue of interviewer-related measurement error, however, which research shows is substantial for name generators (Marsden 2003; Van Tilburg 1998), has received no attention. We thus examine whether interviewer-level measurement error affected not only network data collection in the 2004 GSS, which has received the majority of the scrutiny in recent years, but also other surveys that gathered data on core discussion networks.

Research on Social Connectivity in the United States

With the inclusion of a social networks module, the 1985 GSS offered sociologists their first glimpse into the size, structure, and composition of discussion networks among adults in the United States. The network module presented respondents with a name-generator question, which elicited the listing of individuals with whom they had discussed important matters in the preceding six months. The 2004 GSS again included the social network module, and it appeared that a major shift in the personal network structure of U.S. society had occurred. Specifically, McPherson and colleagues (2006, 2008) reported that the average American's discussion network had shrunk considerably, falling by almost one whole alter. If the 1985 and 2004 GSS data are accurate, social isolation increased massively, by 130 percent, suggesting a sharp increase in the fragmentation of U.S. society. The unexpected nature of these findings was not lost on the research team, who went to considerable lengths to determine whether the network changes they observed could be attributed to a training effect (that is, preceding sections of the survey may have clued in respondents that listing more names in the name generator would lead to more questions), respondent fatigue, uncooperative respondents, or demographic change. None of these factors accounted for the shrunken-networks finding.

The significance of this finding, if it is accurate, is twofold. First, it supports Putnam's (1995, 2000) argument that Americans are withdrawing from public life and becoming more isolated, and raises questions about what societal changes could lead to such a large increase in social isolation. Second, the findings suggest that an important dimension of people's social lives—one linked to the provision of social support (Fischer 1982; Wellman and Wortley 1990) and interpersonal influence (Berg 2009; Friedkin 1999)—is, in fact, under threat. Not surprisingly, the shrunken-networks finding (McPherson et al. 2006) garnered tremendous public and professional interest, generating two Associated Press stories (Crary 2006a, 2006b), countless articles and editorials by other national and local news outlets (e.g., Alfano 2009; Vedantam 2006), more than 500 citations to date in

Google Scholar, three replication studies (Brashears 2011; Hampton et al. 2011; Wang and Wellman 2010), and a book (Fischer 2011).

This result also inspired a robust scholarly debate (Fischer 2009, 2011; McPherson, Smith-Lovin, and Brashears 2009). Arguing that the shrunken-networks finding did not align with other parts of the 2004 GSS or with other surveys, Fischer (2009, 2011) attributed the finding to an unknown technical error.[1] More recently, Fischer (2012) blogged that results from the 2010 GSS, which asked the name-generator question at different points in three versions of the questionnaire, indicated that the 2004 finding was an artifact. Specifically, when distributions of network size were broken down by questionnaire, he found that, consistent with the training-effects and fatigue hypotheses, respondents reported higher levels of social isolation when the name generator was positioned later in the questionnaire and preceded by questions about membership in voluntary associations. We credit Fischer (2009, 2011, 2012) for raising the possibility that the shrunken-networks finding was an artifact and seek to resolve this debate by evaluating the importance of interviewer effects, which have not been considered as a potential explanation.

Other scholars have sought to address the shrunken-networks finding by analyzing name-generator data in new nationally representative samples of adults. Hampton and colleagues (2009, 2011) and Brashears (2011) both reported that the size of core networks has declined, but they found little evidence of dramatic increases in social isolation, leading them to suggest that estimates of the latter were unreliable. Yet both studies reported low response rates—21 and 7 percent in Hampton and colleagues (2011) and Brashears (2011), respectively—making nonresponse bias a concern. In contrast, based on data from the National Social Life, Health, and Aging Project (NSHAP), other scholars found evidence of increasing levels of network size and decreasing social isolation (Cornwell et al. 2008; Cornwell et al. 2009; Cornwell and Waite 2009).[2] Similarly, Wang and Wellman (2010), using an admittedly less restrictive definition of alters, found that only 5 percent of Americans did not interact with friends on a weekly basis. Taken together, Fischer's analysis and subsequent data points raise serious questions about whether social networks did, in fact, decrease in size over the past 25 years. Nevertheless, the extant literature has yet to identify the smoking gun that produced the shrunken-networks finding.

NETWORK DELINEATION AND MEASUREMENT ERROR

We argue that a significant amount of the variation found in survey-based assessments of network size and social isolation is an artifact of measurement error at the interviewer level. Drawing on the "total survey error approach," there are four sources of measurement error: (1) item wording, (2) questionnaire construction, (3) missing data due to item-level nonresponse, and (4) systematic interviewer effects (Groves et al. 2009; Weisberg 2005). Within the survey methodology literature, interviewer effects for specific items are generally small; however, even small interviewer effects can produce large cumulative biases (Groves 1989; Groves et al. 2009; Hox 1994). Possible sources of interviewer error include variation across interviewers in terms of their conduct, behavior, and probes as well as noncompliance and misreported data (Groves et al. 2009).

The important-matters name generator has long been considered an effective single-item approach for eliciting a central component of people's core networks (Bernard, Shelley, and Killworth 1987; Marin and Hampton 2007; Marsden 1993). Nevertheless, much of the extant literature focuses on the potential for respondent-level error in egocentric name generators (for a review, see Marsden 2005). For example, scholars have examined biases related to interpretation of the GSS important-matters question (Bailey and Marsden 1999; Bearman and Parigi 2004; Bernard et al. 1990; Straits 2000), recall issues (Bell, Belli-McQueen, and Haider 2007; Brewer 2000; Marin 2004), and question-order effects, such as respondent priming, fatigue, and training effects (Brashears 2011; McPherson et al. 2006; Pustejovsky and Spillane 2009). Indeed, the ongoing debate about the shrunken-networks finding has focused exclusively on respondent-level measurement error (Brashears 2011; Fischer 2009, 2011, 2012; McPherson et al. 2006, 2009).

However, egocentric name generators appear to be particularly susceptible to interviewer-level measurement error (Marsden 2003; Van Tilburg 1998).[3] In a study of social networks of older Dutch adults, Van Tilburg (1998) reported a zero-order intraclass correlation—that is, the proportion of the total variance in network size related to differences across interviewers (Hox 1994)—of .2, which is substantially higher than the average intraclass correlation of .01 found for many survey questions (Groves 1989). Similarly, Marsden (2003) analyzed the 1998 GSS name generator of "good friends" and found a substantial intraclass correlation of .15. Neither scholar could account for these sizeable interviewer effects, even after controlling for possible sources of interviewer error, nor did they assess whether this measurement error resulted in a negative bias in network size.

In this study, we address these issues related to unexplained interviewer effects and their impact on estimates of network size. Several possible types of interviewer effects, we argue, could lead to downward biases in estimates of network size. First, some interviewers may not be adept at motivating respondents to provide accurate estimates of their discussion networks (Fowler and Mangione 1990). Second, some interviewers may be less active probers, less motivated to go through repetitive follow-up items (known as name-interpreter questions) for each identified alter, or systematically engage in interview behaviors that cause respondents to report fewer network partners (Groves et al. 2009; Marsden 2003; Van Tilburg 1998). For example, if an interviewer displays a lack of enthusiasm, appears inattentive, or is disengaged, especially during the probing process, the respondent could interpret the interviewer's behavior as a cue to disengage.

Finally, interviewers may be subject to the same fatigue and training effects as respondents (Van Tilburg 1998). In the GSS, for example, an interview with a respondent reporting a full slate of alters can result in more than 100 additional name-interpreter questions compared to a respondent with no alters. Unlike respondents, interviewers have full knowledge about how reporting alters leads to these additional name-interpreter questions. Interviewers may thus seek to reduce the number of times they go through the repetitive question loops associated with name generators. Some interviewers may simply engage in data falsification by miscoding the name-generator question, indicating no discussion partners, to avoid all the follow-up questions. We thus see ample reasons to suggest there are interviewer effects on network size and that the impact of this error is a downward bias.

In summary, we argue that the 2004 GSS shrunken-networks finding is an artifact produced by extensive interviewer effects, which systematically biased downward the number of alters listed. Moreover, we suggest that biased estimates of social connectivity are not limited to the 2004 GSS; they are, in fact, common. This study addresses two gaps that exist in the current literature on social isolation. First, as outlined earlier, there appears to be a lack of awareness in the extant literature regarding the importance of interviewer effects when eliciting social networks via name generators. Second, the nature of these interviewer effects is not well understood, suggesting the need for research addressing this issue. This study assesses interviewer effects in estimates of social connectivity, while simultaneously attempting to account for mechanisms that make the collection of egocentric network data via name generators unreliable.

METHODS

We drew data primarily from the 2004 (N = 2,812) and 2010 (N = 2,044) GSS, which were nationally representative, face-to-face surveys of noninstitutionalized adults in the

United States. Social network modules were administered to subsamples of respondents in 2004 (n = 1,426) and 2010 (n = 1,276). In the 2010 GSS, the name-generator question was administered to respondents being re-interviewed as part of Wave III of the 2006 panel. Response rates for the 2004 and 2010 GSS were 70 percent. We also examined a number of comparable samples that met the following criteria: (1) employed representative, face-to-face designs; (2) were publicly available; (3) included interviewer identification numbers; and (4) utilized the important-matters name generator, or a closely related variant. Analyzing additional datasets allowed us to assess interviewer effects linked to egocentric name generators across a number of surveys, while largely holding constant the name-generator question and mode of administration. Interviewer identification numbers were not included with the public-use versions of the 1985 and 1998 GSS data, so we drew on results previously published by Marsden (1987, 2003).

Based on a search of the Inter-university Consortium for Political and Social Research (ICPSR) database as well as relying on our knowledge, we identified three additional surveys meeting these criteria: the 1995 Chicago Health and Social Life Survey (CHSLS), the 2005 National Social Life and Aging Project (NSHAP), and the 1992 Multi-City Study of Urban Inequality (MCSUI). We used the cross-section sample of the CHSLS (N = 890), which was representative of Cook County, IL, an area encompassing the city of Chicago and many of its suburbs, with a population well over five million (Van Haitsma, Paik, and Laumann 2004). The response rate for the CHSLS cross-section was 71 percent. NSHAP was a nationally representative, population-based study of older, community-dwelling Americans (N = 3,005). The response rate for NSHAP was 76 percent. MCSUI data were representative of four cities—Atlanta, Boston, Detroit, and Los Angeles—and had an average response rate of 68 percent. A social network module was administered in each urban sample with the exception of Detroit. Because the NSHAP focused on older Americans and the CHSLS and MCSUI samples were drawn from urban populations, it is worth noting that network size tends to be larger among the young and among urban residents (Marsden 1987). Table A1 in the Appendix presents descriptive statistics for the independent variables, drawn from the 2004 and 2010 GSS.

Dependent Measures

Network size and social isolation. We based measures of network size and social isolation on the important-matters name generator and several variants. The 2004 and 2010 GSS used the important-matters name generator, which was designed to elicit names or initials of up to six alters per respondent. The following question asked respondents to name their network alters:

> From time to time, most people discuss important matters with other people. Looking back over the last six months—who are the people with whom you discussed matters important to you? Just tell me their first names or initials.

If a respondent nominated fewer than five alters, interviewers were instructed to probe for additional partners.

A key difference between the 2004 and 2010 GSS was that the latter employed a split-ballot design: the name generator was positioned differently across three ballots (Fischer 2012), which were administered at random to respondents. The first ballot asked the name generator immediately after the GSS core questions and did not query respondents about membership in voluntary associations. The name generator in ballot two came later and was preceded by the GSS core; the aging, miscellaneous, and shared capitalism sections; and an extensive set of items related to voluntary associations. In ballot three, the name generator was also positioned later in the survey, compared to ballot one, and was preceded by the GSS core; the science, aging, miscellaneous, and shared capitalism sections; and a more limited set of questions on voluntary associations. According to Fischer (2012), ballot one was similar to the 1985 GSS, whereas the second and third ballots sought to replicate embedding of the name generator in ways similar to the 2004 and 1987 GSS, respectively. It is worth noting, however, that in the 2010 GSS, the name-generator question was followed by far fewer name-interpreter questions, compared to the 2004 GSS.

The NSHAP, MCSUI, CHSLS, and 1998 GSS utilized variants of the important-matters name generator. In the NSHAP, respondents were asked to recall discussion partners from the previous 12 months. The MCSUI name generator limited respondents to three alters from the previous six months. The CHSLS's name generator first asked respondents to provide names of three alters who were not sexual partners with whom they had spent a significant amount of free time in the past year. Next, respondents were asked the important-matters name generator, which also elicited up to three discussion partners from the previous 12 months. In both cases, interviewers were instructed to probe when respondents named fewer than three alters. Network size, ranging from zero to six alters, was constructed from these two questions and specified the number of unique alters named by each respondent. The 1998 GSS sought to elicit each respondent's network of good friends. Interviewers collected data on up to six alters, and the name generator was worded as follows:

> Many people have some good friends they feel close to. Who are your good friends (other than your spouse)? Just tell me their names.

Unlike the other surveys, the CHSLS and 1998 GSS name generators instructed respondents to exclude current sex partners and spouses, respectively. We assessed measures of network size for the various surveys. With the exception of the CHSLS and MCSUI surveys, which capped maximum network size at six and three, respectively, we followed convention and top-coded network size at 6.5 (Brashears 2011; Marsden 1987; McPherson et al. 2006, 2008). We also constructed a dichotomous measure of social isolation.

Face-to-face contacts. We constructed a dependent variable from the 2004 GSS assessing the number of face-to-face contacts reported by respondents within the prior year. Respondents were asked the following questions: "Not counting people at work or family at home, about how many other friends or relatives do you keep in contact with at least once a year? Of these friends and relatives, about how many do you stay in contact with by seeing them socially, face-to-face?" This ordinal measure had the following values: 1 (none), 2 (1 to 2), 3 (3 to 5), 4 (6 to 10), 5 (11 to 15), 6 (16 to 25), 7 (26 to 50), and 8 (more than 50). This question was not asked in the social networks module.

Independent Measures from the 2004 and 2010 GSS

Respondent characteristics. We included a number of previously identified respondent characteristics associated with social connectivity and isolation, including race/ethnicity, age, gender, education, marital status, number of children, employment, and frequency of attending religious services (Brashears 2011; Cornwell et al. 2008; Cornwell et al. 2009; Cornwell and Waite 2009; Marsden 1987, 2003; McPherson et al. 2006; Van Tilburg 1998). Race/ethnicity was categorized as white, black, and other. Respondent's age was measured in years but top-coded for 89 years or more. Education reflected years of schooling. We included a dichotomous measure for being married and a continuous measure for the number of children. We also included two measures of respondents' ties to institutions outside the family: (1) frequency of religious service attendance, which was assessed on a nine-point scale, ranging from never to more than once per week, and (2) work status, which indicated whether respondents were employed or currently seeking work, retired, or homemakers.

Interviewer characteristics. Following Marsden (2003), we incorporated measures for several demographic characteristics of the interviewers, including age, gender, race/ethnicity (broken into three categories: white, black, and other), and years of experience interviewing. We also incorporated additional controls for the mode of administration (telephone versus in person) and the number of cases previously collected by interviewers during fielding of the survey.

Interview context. We included several measures assessing respondents' performance and the possibility of fatigue during the interview. Uncooperativeness ranged from friendly to restless/hostile, and noncomprehension assessed respondents' understanding of questions in the survey and ranged from good to poor. To address the possibility of respondent fatigue in the 2004 GSS, we also included a dichotomous variable measuring whether respondents were members of voluntary associations. We included this measure because the GSS voluntary association module was similar to the social network module in that its length was determined by the number of reported memberships. The structure of the 2004 GSS voluntary association module, which came before the network module, could have trained respondents to list fewer names in response to the network name generator in an effort to avoid the required follow-up questions. To assess the impact of interviewer fatigue and behaviors on network size, we included a lagged measure, the number of times the self-administered questionnaire (SAQ) on genes was skipped previously, for each interviewer. Following an interviewer-administered section on genes, interviewers were instructed to provide the SAQ on genes to respondents. Interviewers with a history of respondents who skipped this SAQ, which was administered prior to the network module in the 2004 GSS, may have engaged in systematic practices and behaviors associated with eliciting fewer network alters. Finally, the 2010 GSS included dummy variables distinguishing each respondent's assignment to one of the three ballots.

Analyses

We utilized several analytic approaches to examine the credibility of sample estimates of network size and social isolation. First, consistent with prior research (Hox 1994; Marsden 2003; Van Tilburg 1998), we assessed the size of interviewer effects using multilevel models. Specifically, we estimated maximum likelihood, random-intercept linear regression models of network size, which took the following form:

$$y_{ij} = \alpha + \boldsymbol{\omega}_{0j}\boldsymbol{\gamma} + \boldsymbol{x}_{ij}\boldsymbol{\beta} + \zeta_{0j} + \epsilon_{ij}$$

First-level units (respondents) are indexed $i = 1,\ldots, I$; second-level units (interviewers) are indexed $j = 1,\ldots, J$; α represents the fixed mean intercept; $\boldsymbol{\omega}_{0j}$ are interviewer characteristics; $\boldsymbol{\gamma}$ is a vector of estimated second-level coefficients; $\boldsymbol{x}_{ij}$ are first-level respondent characteristics; $\boldsymbol{\beta}$ is a vector of estimated first-level coefficients; ζ_{0j} is the second-level residual; and ε_{ij} is the first-level residual (Skrondal and Rabe-Hesketh 2004). The resulting linear random-intercept regression models allowed us to decompose the amount of variance in network size attributable to respondent and interviewer levels. We assessed the size of unexplained interviewer effects using the intraclass correlation, ρ_{int}, which we calculated as follows:

$$\rho_{int} = \frac{\sigma^2_{int}}{(\sigma^2_{int} + \sigma^2_{\varepsilon})}$$

σ^2_{ε} and σ^2_{int} represent variances at the respondent and interviewer levels, respectively. Put simply, the intraclass correlation calculates the proportion of the unexplained variance attributable to the interviewer level and captures the extent to which responses are correlated within interviewers. We estimated random-intercept models for the 2004 and 2010 GSS. The split-ballot design of the

2010 GSS provided an opportunity to assess how respondents *and* interviewers responded to the important-matters name generator across three ballots. We thus also estimated separate models for each ballot in the 2010 GSS. Note, though, that the split-ballot was not designed explicitly to study question-order effects on interviewers. Although respondents were only exposed to a single ballot, many interviewers administered all three ballots during fielding of the survey.

We next examined whether specific interviewers in the 2004 and 2010 GSS obtained unlikely levels of social isolation. To estimate the likelihood of obtaining a certain number of isolates across N interviews for each interviewer, we employed an exact binomial probability test. We ran this binomial test for each interviewer to obtain exact p-values based on the binomial distribution and an assumed probability of success. We calculated the upper one-sided p-value as follows:

$$\Pr(k \geq k_{obs}) = \sum_{m=k_{obs}}^{N} \binom{N}{m} p^m (1-p)^{N-m}$$

N represents the number of interviews, k_{obs} indicates the observed number of isolates, and p is an assumed probability of observing an isolate in the population. For example, we employed this test to determine the probability of obtaining $k_{obs} \geq 4$ isolates in $N = 10$ interviews with an assumed probability of observing an isolate of $p = .12$. In this instance, the $\Pr(k \geq 4)$ was .02, which indicates that the level of social isolation reported by the interviewer was above our 1 percent significance threshold.

Finally, we employed an alternative method for identifying outlier interviews and examined the impact of outliers on estimates of social isolation and associations with correlates of network size. First, based on random-intercept models of network size, we employed best linear unbiased predictions (BLUPs) of the random effects to examine whether interviewers were outliers. BLUPs can be generalized to the concept of residuals for the second-level effects in random-intercept models, and therefore can also be used to identify outliers (Robinson 1991; see also Fellner 1986). In this research, we employed a visual plot that ranks BLUP estimates according to their size with error bars designed to represent the confidence intervals around each estimate. To estimate these confidence intervals, we multiplied standard errors of BLUP estimates by 1.39 (as opposed to the conventional 1.96), because use of this multiplication factor results in confidence intervals that can identify significantly different residuals at the 5 percent error level (Hox 2010; see also Goldstein 2010). Second, using ordinary least squares regression, we estimated associations between several respondent characteristics and network size to examine whether expected associations held for outlier and nonoutlier interviewers in the 2004 and 2010 GSS. Proceeding in a similar fashion, we also utilized ordered logistic regression to analyze the relationship between social isolation and the number of face-to-face contacts reported by respondents in the prior year in the 2004 GSS.

RESULTS

To begin examining whether estimates of network size may be influenced by systematic biases related to interviewers, Figure 1 presents a scatterplot of mean network size and estimated intraclass correlations, ρ_{int}, for the 2004 and 2010 GSS, the 1995 CHSLS, and the 2005 NSHAP. We also report Marsden's (2003) previously published estimate from the 1998 GSS, but exclude MCSUI samples, which did not elicit up to six network partners. We estimated intraclass correlations using random-intercept models without covariates. Figure 1 shows substantial interviewer effects associated with elicitation of egocentric networks. All samples had statistically significant intraclass correlations in excess of .1, and two, the CHSLS and 2004 GSS, exceeded .25. Random-intercept models of network size in the Atlanta ($\rho_{int} = .32$), Boston ($\rho_{int} = .32$), and Los Angeles ($\rho_{int} = .37$) samples of the MCSUI also had significant intraclass correlations (results available upon request). Overall, these findings are consistent with Marsden's (2003) analysis of

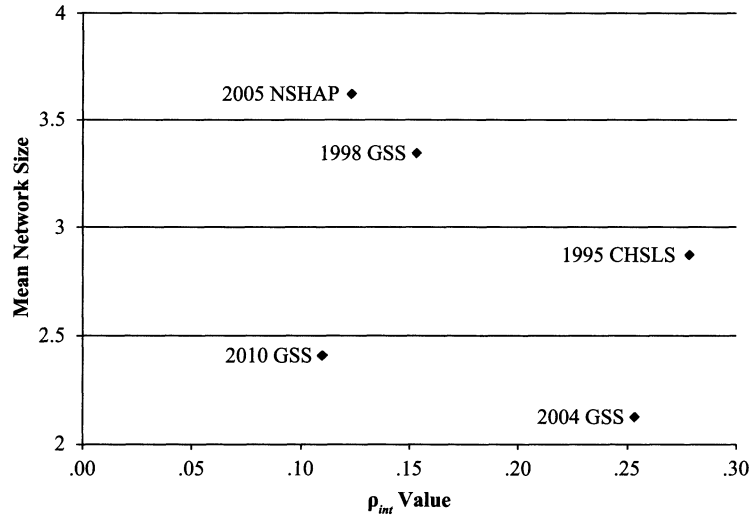

Figure 1. Mean Network Size and Intraclass Correlations across Selected Samples

the 1998 GSS, which had a high intraclass correlation for network size. We also found a negative correlation between network size and estimated intraclass correlations ($r = -.4$). Interviewer-related measurement error associated with the important-matters name generator appears to exert downward pressure on the number of ties elicited.

Compared to the other samples, the 2004 GSS has extreme values on both dimensions. In Figure 1, only the CHSLS has a slightly higher intraclass correlation. Although the observed correlation in the scatterplot is sensitive to the relatively few data points available, interviewer effects associated with elicitation of egocentric networks are a significant concern in all the selected samples. Nevertheless, the 2004 GSS estimate of network size in the United States appears to be an outlier warranting further investigation.

Do Interview Characteristics Explain these Interviewer Effects?

One possible explanation is that the substantial interviewer effects observed in Figure 1 were produced by the social context of interviewer–respondent interactions, so we examined whether respondent, interviewer, and interview characteristics could account for the high intraclass correlations. Table 1 presents random-intercept linear regressions of network size with covariates for the 2004 GSS as well as the 2010 GSS, a comparable sample with a lower intraclass correlation and larger mean network size. Compared to the zero-order random-effects models presented in Figure 1, including covariates for respondent, interviewer, and interview characteristics had little impact, rendering only 13 and 8 percent reductions in the intraclass correlations for the 2004 and 2010 GSS, respectively. Thus, most of the interviewer effect on network size in the 2004 and 2010 GSS cannot be explained by respondent, interviewer, and interview contexts.

Still, these models reveal several informative associations. In the 2004 GSS model, reporting membership in a voluntary association was positively associated with network size. This provides little support for the training-effect hypothesis, which suggested that respondents' prior experiences with questions similar in structure to egocentric name generators and name interpreters, such as the preceding voluntary associations section, would train them to report lower numbers of network partners. This model also shows that interviewers who had at least four prior respondents skip the SAQ gene section tended

Table 1. Random-Intercept Linear Regressions of Network Size

Variables	2004 GSS *b*	2004 GSS (SE)	2010 GSS *b*	2010 GSS (SE)
Respondent Characteristics				
Age	.003	(.003)	.009*	(.004)
Education	.088***	(.016)	.077***	(.018)
Female	.312***	(.088)	.203*	(.100)
Race (White)				
Black	–.416**	(.135)	–.539**	(.157)
Other	–.671***	(.182)	–.531**	(.197)
Currently Married	.071	(.090)	.062	(.100)
Number of Children	–.047	(.030)	–.003	(.032)
Work Status (Employed/Looking)				
Retired	–.128	(.149)	–.350*	(.153)
Homemaker	–.045	(.152)	–.168	(.161)
Religious Service Attendance	–.002	(.017)	.048**	(.018)
Interviewer Characteristics				
Age	–.006	(.008)	–.009	(.006)
Female	–.288	(.233)	.005	(.167)
Race (White)				
Black	–.204	(.250)	–.269	(.200)
Other	–.832*	(.362)	–.442**	(.180)
Experience (Years)	.003	(.026)	–.004	(.047)
Number of Previous Interviews	.007	(.005)	.003	(.003)
Interview Context				
Telephone Mode	–.064	(.119)	–.510***	(.120)
Uncooperativeness	–.238**	(.099)	–.378**	(.117)
Poor Comprehension	–.267*	(.122)	–.444**	(.128)
Voluntary Association Member	.382***	(.096)		
Prior SAQ Skips (None)				
1 Prior Skip	.088	(.151)		
2 to 3 Prior Skips	–.090	(.175)		
4+ Prior Skips	–.319*	(.193)		
Ballot Version (Ballot One)				
Ballot Two			–.410***	(.115)
Ballot Three			.002	(.114)
Constant	1.639**	(.609)	2.625***	(.517)
Intraclass Correlation (ρ_{int})	.218***		.101***	
N	1,369		1,189	
Number of Interviewers[a]	131		175	
Mean Interviews per Interviewer	10.5		6.8	

[a]Due to missing data, we excluded five interviewers from the 2004 GSS and eight interviewers from the 2010 GSS in these analyses.
$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$ (one-tailed test).

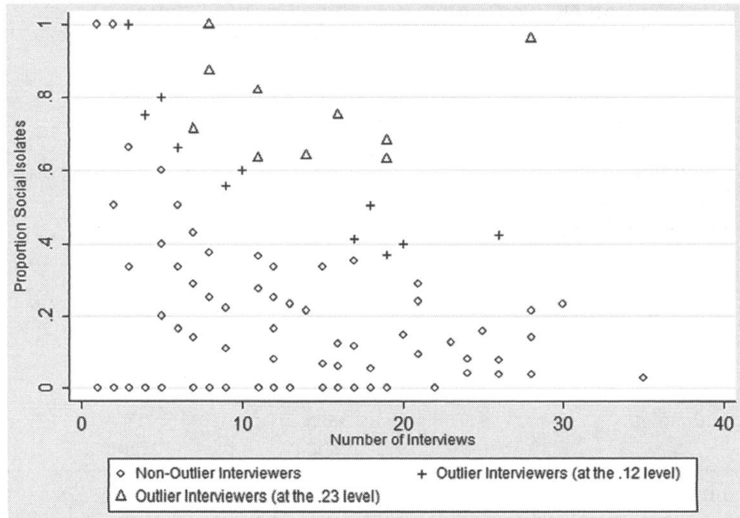

Figure 2. Scatterplot of Proportion Social Isolates and Number of Interviews by Interviewer

to elicit smaller networks in the *current* interview, compared to those with no prior skips.[4] It thus appears that interviewers who were more prone to generating missing data in the SAQ gene section in the 2004 GSS, for whatever reason, tended to elicit smaller networks.

Finally, the model for the 2010 GSS shows that ballot two is negatively associated with network size, but the coefficient for ballot three is nonsignificant. In contrast to these two, the name generator for ballot one was both earlier and not immediately preceded by questions about voluntary associations. This latter finding is partially consistent with fatigue and training effects, but it is not clear whether they affect respondents, interviewers, or both. Taken together, the significant associations for the SAQ gene section in 2004 and ballot two in 2010 are consistent with the notion that some interviewers may have been prone to eliciting smaller networks and generating high levels of social isolation, but, as demonstrated by the significant intraclass correlations, these findings largely do not explain the substantial interviewer effects observed.

Identifying Outlier Interviewers

Given the significant interviewer effects, even after introducing controls, we next examined the extent to which interviewers obtained highly improbable levels of social isolation. Assuming that the observation of a social isolate by an interviewer is a Bernoulli trial, we employed binomial probability tests to determine the likelihood of observing at least k social isolates in N interviews, compared to an expected level of social isolation. We assumed two different levels of social isolation for our population: 23 and 12 percent. The first assumed level of isolation corresponds to the 2004 GSS; the second reflects the level of social isolation in the 2010 GSS and is consistent with the estimate in Hampton and colleagues (2009).

Figure 2 presents a scatterplot of the proportion of isolates reported and the number of interviews conducted for each interviewer in the 2004 GSS. Based on the binomial probability tests, we flagged interviewers who observed levels of social isolation with a probability less than 1 percent ($p < .01$), assuming that the population value for social isolation was both 12 and 23 percent. A number of interviewers occupying the upper region of the figure are outliers—that is, they reported highly improbable levels of social isolation. Of the flagged interviewers, 13, accounting for 122 interviews, reported levels of social isolation above 70 percent. One

Table 2. Size of Discussion Networks after Removing Outlier Interviewers

	Full Sample	23%	12%
2004 GSS			
Percent Interviewers with Extreme Values[a]		8.8%***	18.4%***
Intraclass Correlation (ρ_{int})[b]	.22***	.14***	.13**
Mean Network Size[c]	2.12	2.32	2.43
Proportion Social Isolates[c]	22.6%	16.2%	12.0%
N	1,426	1,262	1,116
2010 GSS			
Percent Interviewers with Extreme Values[a]		.5%	3.3%***
Intraclass Correlation (ρ_{int})[b]	.10***	.10***	.10***
Mean Network Size[c]	2.41	2.41	2.44
Proportion Social Isolates[c]	13.0%	12.7%	11.6%
N	1,272	1,268	1,230

[a]Binomial probability test: $*p < .05$; $**p < .01$; $***p < .001$ (one-tailed test).
[b]Intraclass correlation reported for linear models with respondent, interviewer, and interview context variables included; due to missing data, sample sizes for these estimates are different from those reported for network size. Likelihood ratio test of $\rho_{int} = 0$ ($*p < .05$; $**p < .01$; $***p < .001$) (one-tailed test).
[c]Network size estimates calculated using weights.

outlier interviewer reported social isolates in 27 out of 28 interviews ($p < .000000$).

To examine the extent to which flagged interviewers contributed to interviewer effects, Table 2 reports the percentage of interviewers reporting improbable levels of social isolation ($p < .01$), assuming that the true population level was 23 or 12 percent. The table shows that a larger proportion of interviewers in the 2004 GSS, compared to the 2010 GSS, obtained highly improbable levels of social isolation. For the 2004 GSS, we flagged 9 percent of interviewers (12 out of 136) at the 23 percent level. Assuming each interviewer is a trial, the probability of observing 12 outliers out of 136 interviewers, according to a binomial test, is less than .001. In contrast, for the 2010 GSS, we do not observe a statistically significant number of outlier interviewers ($p > .05$). If we assume that the level of social isolation in the United States is 12 percent—an entirely plausible level given estimates of isolation reported in the 1985, 1998, and 2010 GSS—then nearly one-fifth of interviewers in 2004 obtained improbable levels of social isolation ($p < .001$), but only a small, albeit significant, percentage of interviewers, 3.3 percent, are outliers in the 2010 GSS. To check the robustness of our approach for identifying outliers, we examined whether the identified outlier interviewers were also outliers based on the BLUPs[5] (see Figure A1 in the Appendix). This figure shows that most of the outlier interviewers we identified at the 12 percent level, 88 and 83 percent for the 2004 and 2010 GSS, respectively, had BLUPs with extreme values, suggesting that both approaches yielded convergent results.

Table 2 also shows the impact of outlier interviewers on estimates of interviewer effects and network size. In the 2004 GSS, the intraclass correlation steadily declines as we remove interviewers with extreme levels of social isolation. In fact, once outliers are removed in the 12 percent scenario, intraclass correlations and the level of social isolation are reduced by 41 and 47 percent, respectively, in the 2004 GSS. Put another way, significant amounts of the interviewer effect and social isolation observed in the 2004 data—slightly less than half—can be traced to a number of interviewers who obtained highly unlikely or extreme values for social isolation across all of their interviews. In contrast, the 2010 GSS indicates that a small

Table 3. Variability in Network Size by Geographic Area

	Mean Network Size	Inconsistent with Nonflagged	N
Nonflagged	2.29		909
Nonflagged in Flagged PSU	2.25	ns	353
Flagged Interviewers	.45	***	164

Note: The mean number of PSU's for the 136 interviewers was 2.14, with a standard deviation of 1.81.
$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$ (Wald χ^2 test; one-tailed).

proportion of interviewers obtained improbable levels of social isolation and their removal has little impact on explaining interviewer effects. Further analyses of the CHSLS and the MCSUI urban samples were consistent with the 2004 results: removing interviewers who collected improbable levels of social isolation, which ranged between 5 and 20 percent of interviewers at the .12 level, led to reductions in intraclass correlations from 15 to 54 percent (see Table A2 in the Appendix). More similar to the 2010 GSS, we did not identify any outlier interviewers in the NSHAP data. Taken together, these results indicate that substantial amounts of the interviewer effect observed in the 2004 GSS are attributable to interviewers who obtained highly improbable levels of social isolation.

Does Geography Matter?

A potential alternative explanation for these results is that some interviewers were assigned to geographic regions where social isolation is endemic. To examine this issue, we compared the flagged outlier interviewers against two categories: interviews conducted by nonoutlier interviewers assigned to the same primary sampling units (PSU) and the remaining nonflagged interviews. As Table 3 shows, the mean network size reported by nonflagged interviewers conducting interviews in PSUs with outlier interviewers was 2.3 if the assumed level of isolation was 23 percent, which is not significantly different from nonflagged interviews in other PSUs ($m = 2.3, p > .05$). In other words, there is no significant difference in estimates of network size reported by nonoutlier interviewers across geographic space. There is, however, a significant difference between mean estimates of network size among nonoutliers (both inside and outside the flagged sampling units) and outlier interviewers, who reported a mean network size of .45, nearly two alters fewer. This significant pattern holds at the .12 level as well (results available upon request). Overall, comparison of these mean levels of network size indicates that the existence of outlier interviewers cannot be attributed to geographic variation in the size of social networks.

Convergent Validity

We also examined effects of these outlier interviewers on convergent validity. Table 4 presents two models assessing differences between outlier and nonoutlier interviewers. First, we estimated ordinary least squares models of network size on demographic characteristics by nonoutlier and outlier interviewers, based on the 23 percent level in the 2004 GSS.[6] These models show that education is positively associated with network size among respondents who were assigned nonoutlier interviewers, which is consistent with prior research (Marsden 1987), but we do not observe this association in the subsample of respondents interviewed by outlier interviewers. This result indicates why the association between education and network size was weak, which was the subject of considerable debate (Fischer 2009; McPherson et al. 2006, 2009). If outlier interviewers biased estimates of network size downward and inflated estimates of social isolation, then the expected association between the former and education should be weakened. This is exactly what we observe.

Table 4. Tests for Convergent Validity with Outlier Interviewers Removed

	OLS Regression of Network Size			
	Nonoutliers		Outliers	
Variables	*b*	(SE)	*b*	(SE)
Education	.146***	(.017)	.064	(.041)
N	1,257		163	
	Ordered Logistic Regression of Number of Face-to-Face Contacts			
	Nonoutliers		Outliers	
	Log-odds	(SE)	Log-odds	(SE)
No Discussion Partners	–.511**	(.208)	–.165	(.521)
N	835		94	

Note: Each model controls for age, education, gender, race, and marital status.
*p < .05; **p < .01; ***p < .001 (one-tailed test).

Second, Table 4 presents an ordered logistic regression of an ordinal measure of face-to-face contacts reported within the past 12 months on social isolation and the previously used control variables. The measure of social isolation is negatively associated with the number of face-to-face contacts within the past 12 months only in the nonoutlier sample. We do not observe any association between social isolation and face-to-face contacts in interviews conducted by outlier interviewers. Taken together, these findings are consistent with the notion that respondents interviewed by nonoutlier interviewers have expected associations with education and face-to-face contacts, whereas these associations are largely nonexistent among respondents who were assigned to outlier interviewers. Such a finding is possible to the extent that reported levels of network size and social isolation reflect substantial measurement error due to interviewers who failed to elicit egocentric networks.

Interviewer Effects and Survey Design

The analyses so far show that some interviewers, the outliers, obtained improbable levels of social isolation in the 2004 GSS and were responsible for substantial amounts of interviewer effects associated with the important-matters name generator. One possible explanation is interviewer fatigue and training effects, induced by the name generator's placement in the 2004 GSS near the end of the survey following extensive questions on voluntary associations. To assess whether question order increased interviewer nonperformance, we estimated random-intercept models for ballots one, two, and three of the 2010 GSS, which, as discussed earlier, were designed to replicate how the important-matters name generator was asked in the 1985, 2004, and 1987 GSS, respectively.

Table 5 shows significant differences in estimates of network size across the three ballots. Compared to ballot one, ballots two and three have significantly higher proportions of respondents reporting no discussion partners at 21 (p < .001) and 13 percent (p < .001), respectively, whereas mean network size is lower only in ballot two (p < .01). In other words, ballot two, which sought to replicate the 2004 name generator, produced network statistics in-line with the 2004 numbers (23 percent social isolation and mean network size of 2.12). Results for ballot three are more mixed, but this is not too surprising because it elicited considerably less information about voluntary associations compared to the second ballot. Taken together, these results support the notion that there are question-order

Table 5. Network Statistics and Random-Intercept Linear Regression Models by Ballot in the 2010 GSS

Estimated Parameter	Ballot 1	Ballot 2	Ballot 3
Mean Network Size[a]	2.55	2.13**	2.52
Proportion Social Isolates[a]	5.3%	21.1%***	13.0%***
Random-Intercept Models			
Interviewer-Level Variance (σ_u^2)	.00	.30	.36
Respondent-Level Variance (σ_ε^2)	2.52	2.23	2.47
Intraclass Correlation (ρ_{int})[b]	.00	.12*	.13**
N	402	379	408

[a]Network size estimates calculated using weights; we calculated significance tests using Ballot 1 as the reference group.
[b]Intraclass correlation reported for models with respondent, interviewer, and interview context variables included.
*$p < .05$; **$p < .01$; ***$p < .001$ (Wald χ^2 test; one-tailed).

effects: introducing the important-matters name generator late in the GSS and asking it after questions about voluntary associations produced higher levels of social isolation. Based on these results alone, however, it is not clear whether we are seeing interviewer effects.

To answer this question, Table 5 presents random-intercept linear regression models by ballot. Intraclass correlations clearly indicate significant interviewer effects for ballots two and three (ρ_{int} = .12 and .13, respectively) but not for ballot one. Respondent-level variation varies slightly across these three ballots; it is 10 to 11 percent smaller in ballot two compared to ballots one and three. This provides some evidence that respondents in ballot two, on average, reported numbers of network partners in a more restricted and reduced range, which is consistent with the respondent fatigue and training hypotheses. Importantly, the significant interviewer effects still indicate that part of the shrunken-networks finding is attributable to interviewers. This result thus supports the notion that fatigue or training effects on interviewers, attributable to survey design, resulted in high levels of social isolation and substantial interviewer effects. In short, results from the split-ballot design, while not designed explicitly to test for question-order effects on interviewers, support the argument that survey design can induce interviewer nonperformance in elicitation of egocentric networks.

DISCUSSION

The central question posed in this study was whether social connectivity is truly declining in the United States. Our research suggests that the amount of social isolation in the 2004 GSS is an artifact of interviewer effects. Indeed, this conclusion pertains not only to the 2004 GSS, but to most of the surveys examined here. We found that nearly 50 percent of the interviewer effect observed in the 2004 GSS was tied to some interviewers collecting highly improbable levels of social isolation. These outlier interviewers, many reporting levels of social isolation in excess of 70 percent, were responsible for both the high level of social isolation reported in the 2004 GSS and the weakened associations between network size and two variables—education and the number of face-to-face contacts—that should have been strongly correlated (Fischer 2009; McPherson et al. 2006, 2009). We also showed that interviewer effects linked to enumeration of social networks partially result from training and fatigue effects related to interviewers. Taken together, these analyses support our hypothesis: estimates of social connectivity based on name generators were biased due to the presence of high levels of measurement error attributable to systematic interviewer-effects. Results are consistent with the unfortunate interpretation that survey design played a crucial role in producing training and fatigue effects among interviewers and that some interviewers were unskilled, unmotivated, or worse, intentionally avoided collecting ego-centric network data. Our research thus suggests there is no evidence of increasing levels of social isolation among Americans and raises the specter that the extant literature has yet to render any estimate of social connectivity that is not biased by interviewer effects.

This research has a number of additional implications. First, interviewer effects appear to be a major issue associated with name generators. Every study that has looked for them (Marsden 2003; Van Tilburg 1998), including this one, has observed substantial interviewer effects for these types of survey questions. Yet a significant amount of time, effort, and resources have been devoted to clarifying the question of social connectivity through the acquisition of additional point estimates of network size and social isolation. The evidence presented here, however, suggests that the strategy of collecting additional data points is flawed to the extent that these data also suffer from interviewer effects. Our approach should be replicated and applied to other name generators and other modes of administration, particularly telephone surveys. Put simply, our findings indicate that researchers should be circumspect and approach estimates of network size obtained via name generators with a certain degree of caution.

Second, biases traceable to interviewer effects may have led to a number of incorrect inferences. Although we do not test this proposition directly, it is possible that researchers in other areas are drawing incorrect inferences between social isolation and other variables. If this is true, the effects could be wide ranging. Beyond the seeming confirmation of Putnam's (1995, 2000) warnings, the notion that social isolation has been growing in the United States has been incorporated into theories of marriage and health (Liu and Umberson 2008), overall health in relation to network size (Berkman 2009), and the relationship between loneliness and physical activity (Hawkley, Thisted, and Cacioppo 2009). Furthermore, as outlined earlier, the popular media has taken the shrunken-networks finding as evidence of large-scale, significant changes in the structure of U.S. society and has promoted the idea that technological changes are severing these ties.

Third, in spite of high levels of measurement error found in this study, results for the 2010 GSS and NSHAP seem to indicate that we may have turned the corner. Compared to the other measurements of interviewer effects, ballot one of the 2010 GSS and the NSHAP have the lowest. This indicates that efforts undertaken during collection of these recent datasets have resulted in a reduction of this type of measurement error. In fact, we view the 2005 NSHAP, which had no outlier interviewers and reported a relatively low intraclass correlation, a low level of social isolation, and a robust measure of network size, as the current gold standard in large-scale, survey-based network delineation. We believe that estimates obtained by NSHAP can be attributed, in part, to the placement of the name generator (which was found at the very beginning of the survey) and the fact that multiple questions were used to assess discussion network size (questions that we did not include here).

Fourth, this research has methodological implications that extend beyond the world of name generators used in social network analysis. We believe that extensions of the methodology, first proposed by Hox (1994), should be applied to other types of survey questions similar to the name generator analyzed here. Specifically, interviewer effects could be pervasive in instances where filter questions lead to a myriad of additional items. For example, enumeration of household members, self-answered sections, and retrospective items, like those used to assess partnership histories, schooling, and employment experiences, could be rife with interviewer effects. Future research could further study the 2004 GSS to identify whether interviewer nonperformance affected other parts of the survey; this may help survey researchers understand how survey design may contribute to interviewer measurement error. At a minimum, our findings suggest that survey researchers should be cognizant of the potential threat of interviewer effects on data quality.

Finally, based on our findings, we renew the call advanced by Billiet and Loosveldt (1988) and Marsden (2003) for additional and stricter training regimes for interviewers engaged in personal network elicitation. We acknowledge, however, like Marsden (2003), that our findings with regard to the persistently high levels of interviewer effects are particularly troubling given the high-quality training and extensive quality controls associated with the fielding of the GSS. It appears that simple calls for additional training for interviewers who construct personal networks may have reached their ceiling, and that more serious approaches may need to be taken. Podolny and Baron (1997) have advocated for complete automation of the network construction process, which would certainly eliminate influences of uninterested or unmotivated interviewers. We suggest that when interviewers are used, they could be digitally recorded for subsequent analysis if and when changes of the magnitude reported in the 2004 GSS appear. If the 2004 GSS interviewers had been recorded, for example, we could have examined the outlier interviewers we identified in this study to determine whether their estimates were correct. Alternatively, a low-tech strategy could involve field staff flagging interviewers producing extreme values near the beginning of the fielding of a survey. Respondents of flagged interviewers could be re-interviewed to check for discordant responses. If such a pattern were found, the re-interviews could replace the original interviews.

This research suggests that estimates of social connectivity and social isolation that rely on egocentric name generators suffer from substantial bias related to interviewer measurement error. The magnitude of this bias raises the possibility that inferences in the literature, based on prior estimates of social connectivity, are likely mistaken. It is incumbent upon researchers in this area, therefore, to acknowledge this issue and to address it in ways that seek to minimize its effects. An accurate rendering of social connectivity in the United States remains to be seen.

APPENDIX

Table A1. Unweighted Descriptive Statistics of Independent Variables

	2004 GSS		2010 GSS	
	Mean/Prop.	SD	Mean/Prop.	SD
Respondent Characteristics				
Age	45.89	(16.64)	51.25	(16.44)
Education	13.67	(2.89)	13.72	(2.98)
Female	55.7%		59.3%	
Race/Ethnicity				
White	79.2%		78.6%	
Black	14.5%		14.4%	
Other	6.4%		7.0%	
Currently Married	53.5%		50.4%	
Number of Children	1.84	(1.60)	2.01	(1.66)
Work Status				
Employed or Looking	75.4%		68.9%	
Retired	15.0%		19.8%	
Homemaker	9.6%		11.4%	
Religious Service Attendance	3.80	(2.69)	3.80	(2.81)
Interviewer Characteristics				
Age	52.77	(11.56)	52.30	(11.05)
Female	83.6%		83.7%	
Race				
White	81.8%		68.9%	
Black	14.5%		15.3%	
Other	3.7%		15.8%	
Experience (years)	2.66	(3.51)	1.07	(1.59)
Number of Previous Interviews	16.93	(13.16)	17.53	(15.64)
Interview Context				
Telephone Mode	20.1%		23.7%	
Uncooperativeness	1.20	(.46)	1.18	(.43)
Poor Comprehension	1.15	(.39)	1.16	(.42)
Voluntary Association Member	61.4%			
Prior SAQ Skips				
Zero Prior Skips	50.1%			
1 Prior Skip	20.3%			
2 to 3 Prior Skips	11.7%			
4+ Prior Skips	17.9%			
Ballot in 2010				
Ballot One			33.8%	
Ballot Two			31.9%	
Ballot Three			34.3%	
N	1,369		1,189	

Table A2. Size of Discussion Networks after Removing Outlier Interviewers

	Full Sample	23%	12%
2005 NSHAP			
Percent Interviewers with Extreme Values		0%	0%
Intraclass Correlation (ρ_{int})[a]	.12***	.12***	.12***
Mean Network Size[b]	3.62	3.62	3.62
Proportion Social Isolates[b]	1.7%	1.7%	1.7%
N	3,001	3,001	3,001
1995 CHSLS: Cook County Sample			
Percent Interviewers with Extreme Values		2.0%***	5.9%***
Intraclass Correlation (ρ_{int})[a]	.28***	.17***	.13**
Mean Network Size[b]	2.87	3.07	3.17
Proportion Social Isolates[b]	12.9%	7.5%	5.7%
N	883	827	768
1992 MCSUI: Atlanta Sample			
Percent Interviewers with Extreme Values		2.3%*	4.7%***
Intraclass Correlation (ρ_{int})[a]	.32***	.30***	.27***
Mean Network Size[b,c]	2.14	2.15	2.22
Proportion Social Isolates[b]	6.4%	5.9%	4.4%
N	1,276	1,255	1,153
1992 MCSUI: Los Angeles Sample			
Percent Interviewers with Extreme Values		12.3%***	19.6%***
Intraclass Correlation (ρ_{int})[a]	.37***	.24***	.22***
Mean Network Size[b,c]	1.95	2.26	2.34
Proportion Social Isolates[b]	25.3%	13.6%	11.0%
N	4,025	2,553	1,777
1992 MCSUI: Boston Sample			
Percent Interviewers with Extreme Values		13.0%***	18.5%***
Intraclass Correlation (ρ_{int})[a]	.32***	.16***	.17***
Mean Network Size[b,c]	2.25	2.35	2.35
Proportion Social Isolates[b]	12.3%	9.0%	8.8%
N	1,604	1,036	872

[a]Intraclass correlation reported for intercept-only linear models of network size.
[b]Network size estimates calculated using weights.
[c]Total network size for MCSUI does not exceed three alters.
$*p < .05$; $**p < .01$; $***p < .001$ (one-tailed tests); likelihood ratio test of $\rho_{int} = 0$.

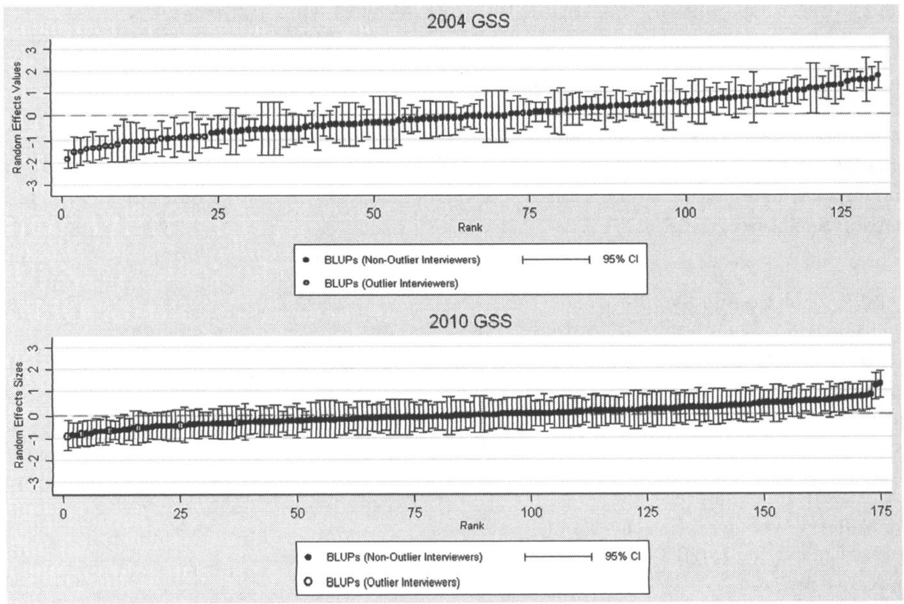

Figure A1. 2004 and 2010 GSS BLUPs

Authors' Note

Coding syntax and statistical commands used in this research are available on Anthony Paik's website (http://clas.uiowa.edu/sociology/people/anthony-paik).

Acknowledgments

We thank the reviewers of the *American Sociological Review,* Matthew Brashears, Benjamin Cornwell, Claude Fischer, Jennifer Glanville, Karen Heimer, Edward Laumann, Kevin Leicht, Freda Lynn, Jamie Sanchagrin, Lynn Smith-Lovin, Barry Wellman, and participants of the Theory Workshop at the University of Iowa for their comments and suggestions.

Notes

1. Drawing from a variety of sources, including Gallup Polls and multiple waves of the GSS, Fischer (2011) reported that consistent patterns or trends associated with increased social isolation simply could not be found.
2. Moreover, analyses of Wave II of NSHAP show an increase in network size from 3.5 to 3.7 (personal communication with Benjamin Cornwell).
3. Indeed, concern about interviewer effects associated with elicitation of social networks is a long-standing concern. In his book *To Dwell among Friends,* Fischer (1982) adopted procedures to minimize interviewer effects (personal communication).
4. Based on a reviewer's suggestion, we also estimated this model with a lagged variable indicating the proportion of interviews in which the SAQ was skipped previously. This alternative measure, which was also statistically significant, produced similar results (results available upon request).
5. BLUP estimates were rank-ordered and then examined (along with each estimate's corresponding confidence interval) to determine whether an individual estimate was significantly different from zero. We used BLUPs as a secondary method for identifying outlier interviewers because the identification process relies on deviations from the mean, which assumes that the relevant sample means are representative of the true population. As we argued elsewhere in this article, however, the mean values of network size and social isolation are suspect (especially in the 2004 GSS), which leads us to view the BLUP estimates with a certain degree of caution.
6. Analyses based on the 12 percent scenario (available upon request) produced results similar to those discussed here.

References

Alfano, Sean. 2009. "The Lonely States of America." CBSNews, February 11. Retrieved January 10, 2012 (http://www.cbsnews.com/stories/2006/06/28/opinion/meyer/main1762234.shtml).

Bailey, Stefanie and Peter V. Marsden. 1999. "Interpretation and Interview Context: Examining the General Social Survey Name Generator Using Cognitive Methods." *Social Networks* 21:287–309.

Bearman, Peter and Paolo Parigi. 2004. "Headless Frogs and Other Important Matters: Conversation Topics and Network Structure." *Social Forces* 83:535–57.

Bell, David C., Benedetta Belli-McQueen, and Ali Haider. 2007. "Partner Naming and Forgetting: Recall of Network Members." *Social Networks* 29:279–99.

Berg, Justin A. 2009. "Core Networks and Whites' Attitudes toward Immigrants and Immigration Policy." *Public Opinion Quarterly* 73:7–31.

Berkman, Lisa F. 2009. "Social Epidemiology: Social Determinants of Health in the United States: Are We Losing Ground?" *Annual Review of Public Health* 30:27–41.

Bernard, H. Russell, Eugene C. Johnson, Peter D. Killworth, Christopher McCarty, Gene A. Shelley, and Scott Robinson. 1990. "Comparing Four Different Methods for Measuring Personal Social Networks." *Social Networks* 12:179–215.

Bernard, H. Russell, Gene Ann Shelley, and Peter Killworth. 1987. "How Much of a Network Does the GSS and RSW Dredge Up?" *Social Networks* 9:49–61.

Billiet, Jacques and Geert Loosveldt. 1988. "Improvements of the Quality of Responses to Factual Survey Questions by Interviewer Training." *Public Opinion Quarterly* 52:190–211.

Brashears, Matthew E. 2011. "Small Networks and High Isolation? A Reexamination of American Discussion Networks." *Social Networks* 33:331–41.

Brewer, Devon D. 2000. "Forgetting in the Recall-Based Elicitation of Personal and Social Networks." *Social Networks* 22:29–43.

Brewer, Devon D. and Cynthia M. Webster. 1999. "Forgetting of Friends and its Effects on Measuring Friendship Networks." *Social Networks* 21:361–73.

Cornwell, Benjamin, Edward O. Laumann, and L. Philip Schumm. 2008. "The Social Connectedness of Older Adults: A National Profile." *American Sociological Review* 73:185–203.

Cornwell, Benjamin, L. Philip Schumm, Edward O. Laumann, and Jessica Graber. 2009. "Social Networks in the NSHAP Study: Rationale, Measurement, and Preliminary Findings." *Journal of Gerontology: Social Sciences* 64B(S1):i47–i55.

Cornwell, Erin Y. and Linda J. Waite. 2009. "Measuring Social Isolation among Older Adults Using Multiple Indicators from the NSHAP Study." *Journal of Gerontology: Social Sciences* 64B(S1):i38–i46.

Crary, David. 2006a. "With Isolation More Common, Americans Try to Connect." *Associated Press,* August 5. Retrieved February 1, 2012 (http://www.lexisnexis.com).

Crary, David. 2006b. "Isolated Americans Trying to Connect." *Associated Press,* August 5. Retrieved February 1, 2012 (http://www.lexisnexis.com).

Fellner, William H. 1986. "Robust Estimation of Variance Components." *Technometrics* 28:51–60.

Fischer, Claude S. 1982. *To Dwell among Friends: Personal Networks in Town and City*. Chicago: University of Chicago Press.

Fischer, Claude S. 2009. "The 2004 GSS Finding of Shrunken Social Networks: An Artifact?" *American Sociological Review* 74:657–69.

Fischer, Claude S. 2011. *Still Connected: Family and Friends in America since 1970*. New York: Russell Sage Foundation.

Fischer, Claude S. 2012. "The Loneliness Scare Is Back." Made in America: A Social History of American Culture and Character Blog. Retrieved June 18, 2012 (http://madeinamericathebook.wordpress.com/2012/04/24/the-loneliness-scare-is-back/#more-2352).

Fowler, Jr., Floyd J. and Thomas W. Mangione. 1990. *Standardized Survey Interviewing: Minimizing Interviewer-Related Error*. Newbury Park, CA: Sage Publications.

Friedkin, Noah E. 1999. "Choice Shift and Group Polarization." *American Sociological Review* 64:856–75.

Goldstein, Harvey. 2010. *Multilevel Statistical Models*. Hoboken, NJ: John Wiley & Sons.

Groves, Robert M. 1989. *Survey Error and Survey Costs*. Hoboken, NJ: John Wiley & Sons.

Groves, Robert M., Floyd Fowler Jr., Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau. 2009. *Survey Methodology,* 2nd ed. Hoboken, NJ: John Wiley & Sons.

Hampton, Keith N., Lauren F. Sessions, and Eun Ja Her. 2011. "Core Networks, Social Isolation, and New Media: How Internet and Mobile Phone Use Is Related to Network Size and Diversity." *Information, Communication & Society* 14:130–55.

Hampton, Keith N., Lauren F. Sessions, Eun Ja Her, and Lee Rainie. 2009. "Social Isolation and New Technology: How the Internet and Mobile Phones Impact Americans' Social Networks." Pew Internet & American Life Project Report (http://pewresearch.org/pubs/1398/internet-mobile-phones-impact-american-social-networks).

Hawkley, Louise C., Ronald A. Thisted, and John T. Cacioppo. 2009. "Loneliness Predicts Reduced Physical Activity: Cross Sectional and Longitudinal Analysis." *Health Psychology* 28:354–63.

Hox, Joop J. 1994. "Hierarchical Regression Models for Interviewer and Respondent Effects." *Sociological Methods & Research* 22:300–318.

Hox, Joop J. 2010. *Multilevel Analysis: Techniques and Applications*. New York: Routledge.

Liu, Hui and Debra J. Umberson. 2008. "The Times They Are a Changin': Marital Status and Health Differentials from 1972 to 2003." *Journal of Health and Social Behavior* 49:239–53.

Marin, Alexandria. 2004. "Are Respondents More Likely to List Alters with Certain Characteristics? Implications for Name Generator Data." *Social Networks* 26:289–307.

Marin, Alexandria and Keith N. Hampton. 2007. "Simplifying the Personal Network Name Generator: Alternatives to Traditional Multiple and Single Name Generators." *Field Methods* 19:163–93.

Marsden, Peter V. 1987. "Core Discussion Networks of Americans." *American Sociological Review* 52:122–31.

Marsden, Peter V. 1993. "The Reliability of Network Density and Composition Measures." *Social Networks* 15:399–421.

Marsden, Peter V. 2003. "Interviewer Effects in Measuring Network Size Using a Single Name Generator." *Social Networks* 25:1–16.

Marsden, Peter V. 2005. "Recent Developments in Network Measurement." Pp. 8–30 in *Models and Methods in Social Network Analysis,* edited by P. J. Carrington, J. Scott, and S. Wasserman. New York: Cambridge University Press.

McPherson, Miller, Lynn Smith-Lovin, and Matthew E. Brashears. 2006. "Social Isolation in America: Changes in Core Discussion Networks over Two Decades." *American Sociological Review* 71:353–75.

McPherson, Miller, Lynn Smith-Lovin, and Matthew E. Brashears. 2008. "ERRATA: Social Isolation in America: Changes in Core Discussion Networks over Two Decades." *American Sociological Review* 73:1022.

McPherson, Miller, Lynn Smith-Lovin, and Matthew E. Brashears. 2009. "Models and Marginals: Using Survey Evidence to Study Social Networks." *American Sociological Review* 74:670–81.

Podolny, Joel M. and James N. Baron. 1997. "Resources and Relationships: Social Networks and Mobility in the Workplace." *American Sociological Review* 62:673–93.

Pustejovsky, James E. and James P. Spillane. 2009. "Question-Order Effects in Social Network Name Generators." *Social Networks* 31:221–29.

Putnam, Robert D. 1995. "Bowling Alone: America's Declining Social Capital." *Journal of Democracy* 6:65–78.

Putnam, Robert D. 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon and Schuster.

Robinson, G. K. 1991. "That BLUP Is a Good Thing: The Estimation of Random Effects." *Statistical Science* 6:15–51.

Ruan, Danching. 1998. "The Content of the General Social Survey Discussion Networks: An Exploration of the General Social Survey Discussion Name Generator in a Chinese Context." *Social Networks* 20:247–64.

Skrondal, Anders and Sophia Rabe-Hesketh. 2004. *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. Boca Raton, FL: Chapman & Hall.

Straits, Bruce C. 2000. "Ego's Important Discussants or Significant People: An Experiment in Varying the Wording of Personal Network Name Generators." *Social Networks* 22:123–40.

Van Haitsma, Martha, Anthony Paik, and Edward O. Laumann. 2004. "The Chicago Health and Social Life Study Design." Pp. 39–65 in *The Sexual Organization of the City,* edited by E. O. Laumann, S. Ellingson, J. Mahay, A. Paik, and Y. Youm. Chicago: University of Chicago Press.

Van Tilburg, Theo. 1998. "Interviewer Effects in the Measurement of Personal Network Size: A Nonexperimental Study." *Sociological Methods & Research* 26:300–328.

Vedantam, Shankar. 2006. "Social Isolation Growing in U.S., Study Says." *Washington Post,* June 23. Retrieved January 10, 2012 (http://www.washingtonpost.com/wp-dyn/content/article/2006/06/22/AR2006062201763.html).

Wang, Hua and Barry Wellman. 2010. "Social Connectivity in America: Changes in Adult Friendship Network Size from 2002 to 2007." *American Behavioral Scientist* 53:1148–69.

Weisberg, Herbert F. 2005. *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. Chicago: University of Chicago Press.

Wellman, Barry and Scot Wortley. 1990. "Different Strokes from Different Folks: Community Ties and Social Support." *American Journal of Sociology* 96:558–88.

Anthony Paik is Associate Professor of Sociology and Gender, Women's and Sexuality Studies and Director of the Social Science Policy Program in the Public Policy Center at the University of Iowa. His research focuses in the areas of social networks, social demography, sexuality, the legal profession, and formal models. His publications include *The Sexual Organization of the City* (2004), which he co-edited.

Kenneth Sanchagrin is a doctoral candidate in the Department of Sociology at the University of Iowa and holds a juris doctor (JD). He is interested in social networks and social capital, and the role they play in a variety of outcomes, including career trajectories, juvenile delinquency, and deviant behavior. His dissertation focuses on relationships between different forms of social capital and the career trajectories of early-career lawyers. His other current projects include an examination of the effects of mentoring on early career transitions among new lawyers as well as a study of peer networks and violence.