

Measuring Network Sizes in the Context of Respondent Driven Sampling: Evidence from Two Independent Surveys

Sunghee Lee¹  · Jacob Fisher^{1,2} · Ai Rene Ong³ · Michael Elliott¹ · Kaidar Nurumov¹ · Jinseok Kim¹

¹University of Michigan, Institute for Social Research

²Duke University, Social Science Research Institute

³American Institutes for Research

Social networks are the fundamental premise of respondent driven sampling (RDS). Personal network sizes termed as “degrees” play an important role in the RDS literature, as RDS-specific point estimators incorporate degrees as an adjustment factor approximating the selection probabilities. For this reason, degrees relevant to RDS should consider *recruitability* rather than the state of connectedness.

This study examines various measures of degrees (standard degrees; degrees with priming recruitment requests; and degrees using naming stimuli that tap into specifics of social relationships) in two independent RDS surveys: one targeting people who inject drugs (PWID, $n=410$) and the other targeting Korean immigrants ($n=637$). The latter randomized interview language for bilingual English-Korean speakers.

There was greater noise in the standard degree measure compared to other degree measures. With a subtle hint about the recruitment request, respondents reported knowing fewer people, compared to the standard degree question, implying a mismatch between the standard degree question and recruitability. Degree reports were sensitive to interview language: reported degrees were smaller in Korean than English interviews and better explained the recruitment. Degrees measured in the contexts of close social relationships were shown to improve inference, while this was not true for the standard degree. This warrants scrutiny of network measures that reflect the RDS recruitment mechanism.

Keywords: respondent driven sampling; network size measurement; measurement error; question wording

1 Introduction

Central to respondent driven sampling (RDS) are social networks. Because RDS participants are recruited through tracing links between nodes in social networks, RDS is applicable for populations whose members are networked. As a participant is recruited by a node in its own network, the network size translates into chances to be recruited into RDS samples. Those with larger networks have higher chances of being recruited than those with smaller networks as illustrated in Fig. 1 with two ego-centric networks with different degrees, d_i . Ego 1 is linked to 8 alters ($d_1 = 8$),

Supplementary Information The online version of this article (<https://doi.org/10.18148/srm/2025.v19i1.8336>) contains supplementary material.

Corresponding author: Sunghee Lee, Institute for Social Research, University of Michigan, 426 Thompson St., Ann Arbor, MI 48104, USA (Email: sungheel@umich.edu)

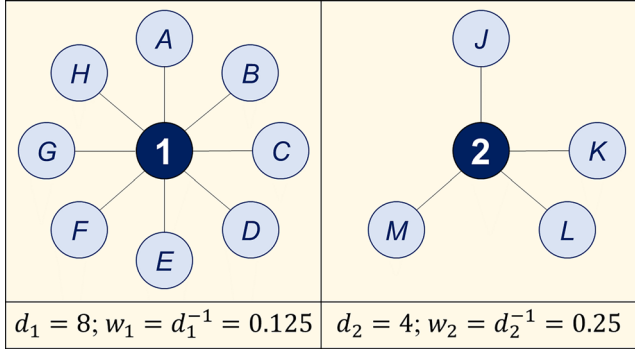


Fig. 1

Illustration of Degree-Based Weights of Two Egos in Respondent Driven Sampling

A through H, whereas Ego 2—4 alters ($d_2 = 4$), J through M. Under certain assumptions, summarized below, Ego 1 has twice the chance to be recruited into an RDS sample compared to Ego 2.

This rationale has led most existing RDS estimators to account for the network size, termed as *degree*, as a form of a factor adjusting for differential selection probabilities (Gile 2011; Heckathorn 2007; Ott et al. 2019; Rocha et al. 2017; Salganik and Heckathorn 2004; Volz and Heckathorn 2008). This is true for estimators for designs that combine RDS with probability sampling (Agans et al. 2021). Simply put, these estimators attempt to account for differential recruitment probabilities through ‘weights’ (w_1 and w_2) in the estimation. While this may appear analogous to the selection weights in probability sampling, RDS degree weights and probability sampling selection weights are not the same: probability sampling weights are a product of a sample design, known prior to sample selection, whereas degree weights are known only after obtaining degrees from participants.

There are implicit assumptions in using degree weights as a proxy for differential selection probabilities. First, it assumes reciprocity: for example, for Ego 1 to be assigned with the probabilities of being recruited by all alters, its alters A through H should consider Ego 1 as their own alter. In order for Ego 2 to be assigned with a weight twice larger than for Ego 1, three additional assumptions are necessary: every alter across egos (e.g., A and J) has the same propensity to make recruitment attempts; two egos have the same propensity to participate; and ego’s characteristics or relationship with the alters do not affect the selection probability. See Gile and Handcock (2010, Table 1) and Lee (2009) for assumption details. Finally, and most critically for this manuscript, it assumes that the reported degree equals the true degree that underlies the recruitment mechanisms at work in RDS.

When (1) putting these assumptions into one, (2) reversing the recruitment direction based on the reciprocity (i.e., egos recruit their alters), and (3) reflecting the RDS practice of issuing a fixed number of coupons, c , across all n participants (except for participant i with $d_i < c$, to whom d_i coupons are issued), the count of recruits from i ($i = 1, \dots, n$) can be expressed as: $m_i = c r_i a_{j|i}$, where r_i is i ’s propensity to attempt to recruit alters and $a_{j|i}$ is the participation propensity by alter j of participant i ($j = 1, \dots, d_i$). Under the same recruitment attempt propensity assumption ($r_i = r$) and the same participation propensity assumption ($a_{j|i} = a$), $m_i = c r a = m$. However, $m_i \neq m$ in reality (e.g., Forrest et al. 2016; Lachowsky et al. 2016; Lee et al. 2017). This means that the same propensity assumptions do not hold.

One may argue that degrees (d_i) do not directly affect m_i . It should be noted that, under the incentive system of RDS, m_i equates to the amount of the recruitment incentive. We counter-argue that, compared to participants with smaller degrees, those with larger degrees have more chances to maximize their alters’ participation propensities ($a_{j|i}$). This means that a larger network offers a larger pool of alters that will allow participant i to exercise selectivity in recruiting their alters by selecting alters who are more likely to participate, leading to systematically more recruits. This equates to increasing chances to obtain not only the recruitment incentive itself but also a higher incentive amount. When putting this to Fig. 1, Ego 1 has 8 alters to maximize $a_{j|i}$ with, while Ego 2 has only 4. This may be the impetus behind the recommendation in the RDS literature to start the data collection with seeds that have large social networks (Heckathorn 1997) and the statement that considered larger degrees as an ability to recruit more peers (Johnston and Sabin 2010).

While the importance of social networks for RDS is evident, its conceptualization and usage in inference are dependent on a valid degree measurement. This is dubious, as social network measurement is prone to error (Marsden 1990) due to its reliance on self-reports. Moreover, to our best knowledge, there is no established literature on degree measurement for RDS. To fill this gap, this paper focuses on the measurement of degrees in the context of RDS surveys. It starts with a broad theoretical overview of social network measurements along with the current practice in RDS and then delves into an empirical study that compares various types of degree measures in two independent RDS surveys as an attempt to provide recommendations on how to measure degrees pertinent to RDS.

Table 1*Participant Characteristics^a*

Project Attitude Towards Health		Health and Life Study of Koreans			
	<i>n</i>		Total	Los Angeles	Michigan
<i>Age</i>					
18–40 years old	93	18–29 years old	234	126	108
41–60 years old	159	30–39 years old	154	74	80
60 years old or older	158	40–49 years old	133	79	54
		50–59 years old	82	53	29
		60 years old or older	32	25	7
<i>Sex</i>					
Female	149		386	222	164
Male	261	Male	251	136	115
<i>Employment status</i>					
Employed	47	Employed	353	206	147
Not employed	363	Not employed	284	152	132
<i>Household income last year</i>					
≤ \$20,000	356	≤ \$50,000	357	224	133
> \$20,000	39	> \$50,000	267	128	139
<i>Race/ethnicity</i>					
Non-Hispanic White	120	<i>Ethnic identity</i>			
Other	289	Korean	420	225	195
		Korean American	187	118	69
		Other	29	15	14
<i>Living arrangement</i>					
Living alone	295	<i>Marital status</i>			
Living with someone else	105	Married	348	185	163
		Other	289	173	116
<i>Data collection site</i>					
Detroit	285	LA	358	NA	NA
Outside Detroit	125	MI	279	NA	NA
<i>Education</i>					
< High school	156	<i>Education</i>			
High school	144	<4-year college degree	255	148	107
> High school	110	4-year college degree	191	126	65
		>4-year college degree	191	84	107
<i>Interview language</i>					
		English	209	109	100
		Korean	428	249	179
Total Sample	410		637	358	279

^aThe number of participants under each variable may not sum up to the total participant counts due to item missing.

1.1 Errors in Social Network Measurement

Self-reported social networks are subject to errors, as evidenced by the efforts to model such measurement errors in the social network literature (e.g., Kadushin et al. 2006; McCormick et al. 2010). There are different types of errors in network measurement (e.g., Killworth et al. 2006). For

an easier illustration, we introduce each error using Korean American social networks as examples. The first error arises when people fail to recall their network members accurately and misreport (Sudman 1985). As relationships differ in type and strength (Kadushin 2012), some relationships are more salient than others, affecting their likelihood of being recalled. For example, a Korean American man

Degree Type		Question Wording
Total ^a	Standard	How many males in Greater Detroit area do you know who inject and you have seen in the past 30 days? {[IF R IS NOT A SEED] Please include the person who gave you the coupon if it was male.}
		How many females in Greater Detroit area do you know who inject and you have seen in the past 30 days? {[IF R IS NOT A SEED] Please include the person who gave you the coupon it was female.}
Naming Stimuli ^b	First Name, "Pat"	Among {NETSIZE} persons, how many have the first name "Pat", "Patrick", "Patsy", "Patty" or "Patricia"?
	Interact/Contact Frequency	Among {NETSIZE} persons, how many do you interact with (including talking to, visiting with, calling, emailing, texting, facebook, etc.) personally more than once week?
	Relation Closeness	Among {NETSIZE} persons, how many do you feel close to (that is, you feel at ease with, can talk to about what is on your mind, or call on for help)?
Total	Recruitment Prime	After this survey, we are going to give you green coupons that you can give to other injection drug users so that they can participate in this study. If they participate, you will be compensated with \$10 per participant. We are specifically interested in adults ages 18 years old or older living in Greater Detroit area who have injected drug in the last 6 months. How many such people do you know?

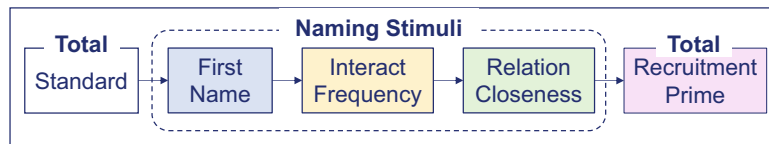


Fig. 2

Degree Measurement Question Sequence and Wording, Project Positive Attitudes Towards Health. a Question Wording (^aAnswers to these two standard questions were summed as NETSIZE used in naming stimuli questions. ^bIf NETSIZE = 1, naming stimuli questions were asked with binary response options, e.g., "Is that person's first name, 'Pat', 'Patrick', 'Patsy', 'Patty' or 'Patricia'?" If NETSIZE = 0 or missing, we did not ask naming stimuli questions.). **b Question Sequence**

who knows two Korean American adults, one, his wife and the other, an acquaintance, is less likely to recall the acquaintance than the wife. In order to report the network size, one needs to determine the scope of own network, i.e., who is in or out of the network. This choice, referred to as the boundary specification, is another source of error (Stork and Richards 1992). A boundary in RDS studies is the target population (e.g., Korean Americans) which, in the practice of RDS, depends on how participants define it. A discrepancy causes the boundary error. One may consider anyone with Korean heritage living in America as Korean American, while the next person may consider only single-race Koreans born in the U.S. as Korean American. The boundary of Korean Americans may be specified differently not only between participants and the study but also across participants. The third type of error, known as the transmission error, arises when people do not know relevant characteristics about their peers (Killworth et al. 2006). If

a person interacts with a friend of Korean descent but does not know this heritage, misreporting due to transmission error arises. Fourthly, in some cases, participants may simply provide careless responses, for example, reporting that they know 123,456 Korean Americans (Meade and Craig 2012) or rounding answers to multiples of 5 (Berchenko, Rosenblatt, and Frost 2017). These four types of errors examined thus far affect self-reports of degrees.

The last type is the barrier error which arises within a given social network when members of one subgroup consistently have a weaker connection with members of another subgroup than their own subgroup (Killworth et al. 2006). With this error, misreporting of degrees may not arise, but RDS recruitment chains may become stuck in a particular subgroup (Fisher and Merli 2014). Such subgroups within Korean Americans may be Korean speakers, who may not have a strong connection with English-speaking Korean Americans.

Degree Type		Question Wording
Total ^a	Standard	How many other Koreans, who are adults, born outside of U.S. and living in [LA/Michigan] do you know?
	Sequential	1. How many Koreans living in [LA/Michigan] do you know? 2. Among {ANSWER FROM 1} Koreans you know, how many are adults? 3. Among {ANSWER FROM 1} Korean adults you know in [LA/Michigan], how many are born outside of U.S.?
	Recruitment Prime	At the end of the survey, we will give you invitation notes that you can give to other Koreans just like yourself. With the invitation, people can participate in the survey and earn \$20. We will also provide \$5, if you recruit. This is because recruiting non-US-born Koreans is critical for this study but at the same time very difficult. We would like to express our appreciation with this monetary incentive. How many other Korean adults, who were born outside of U.S. and are living in [LA/Michigan] do you know?
Naming Stimuli ^b	Last Name, “Oh”	Now think about {TOTAL} Korean {adult/adults} born outside of U.S. you know in [LA/Michigan]. Among {TOTAL} persons, how many have the last name “Oh”?
	Family	Among {TOTAL} persons, how many are your immediate family members?
	Interact/Contact Frequency	Among {TOTAL} persons, how many do you interact with (including talking to, visiting with, calling, emailing, texting, kakao talk, facebook, etc) personally more than once week?
	Close	Among {TOTAL} persons, how many do you feel close to (that is, you feel at ease with, can talk to about what is on your mind, or call on for help)?

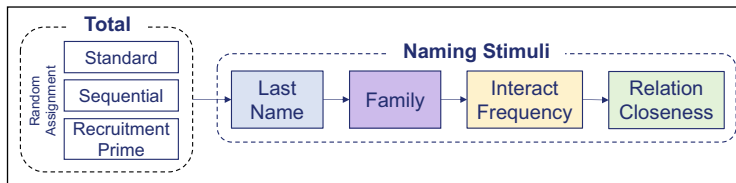


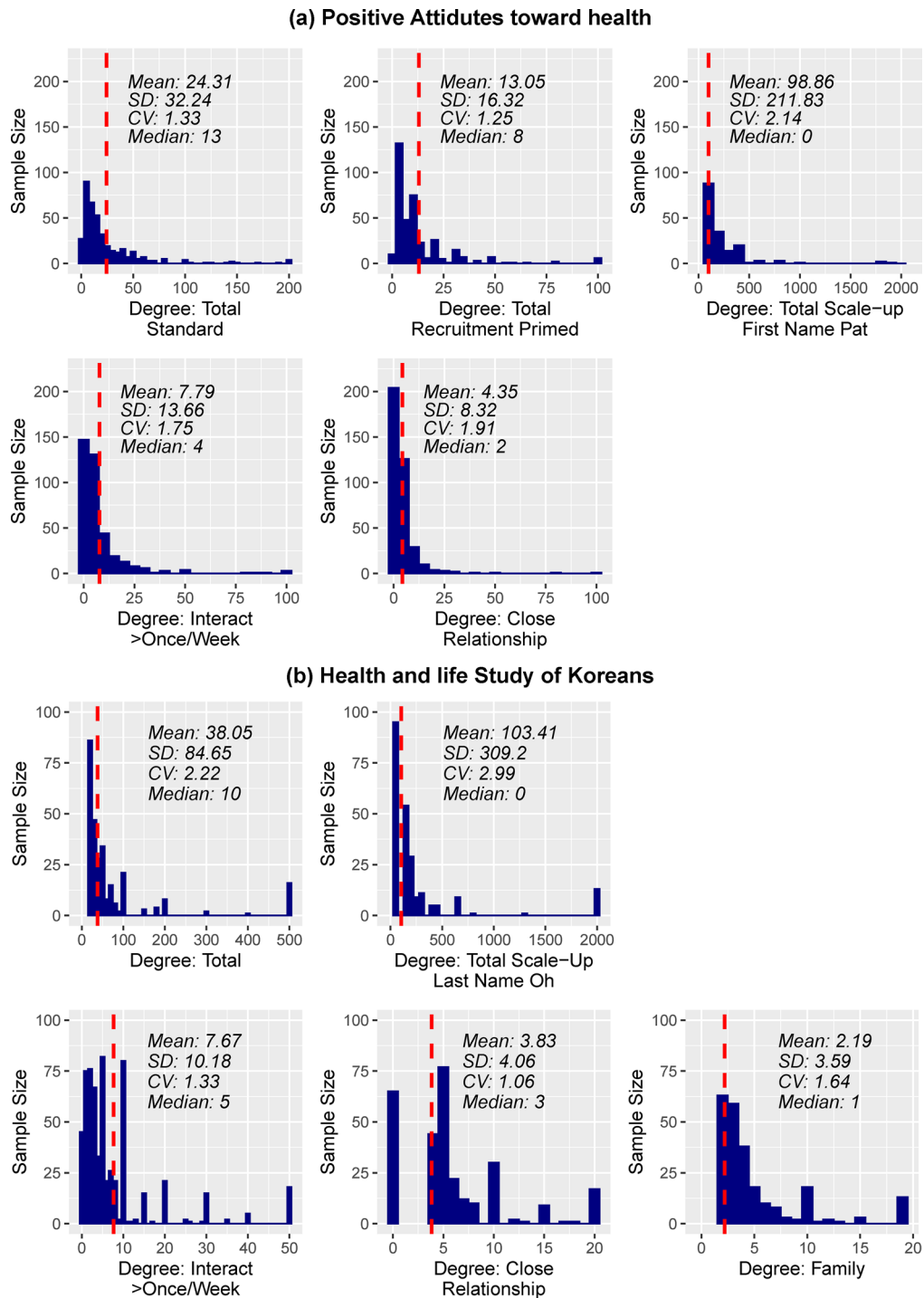
Fig. 3

Degree Measurement Question Sequence and Wording, Health and Life Study of Koreans.
a Question Wording (“Participants were randomly assigned to one of the three total degree questions. Answers to these questions were used as TOTAL in naming stimuli questions. ^bIf TOTAL = 1, naming stimuli questions were asked with binary response options, e.g., “Is that person’s last name, ‘Oh’”? If TOTAL = 0 or missing, we did not ask naming stimuli questions.).
b Question Sequence

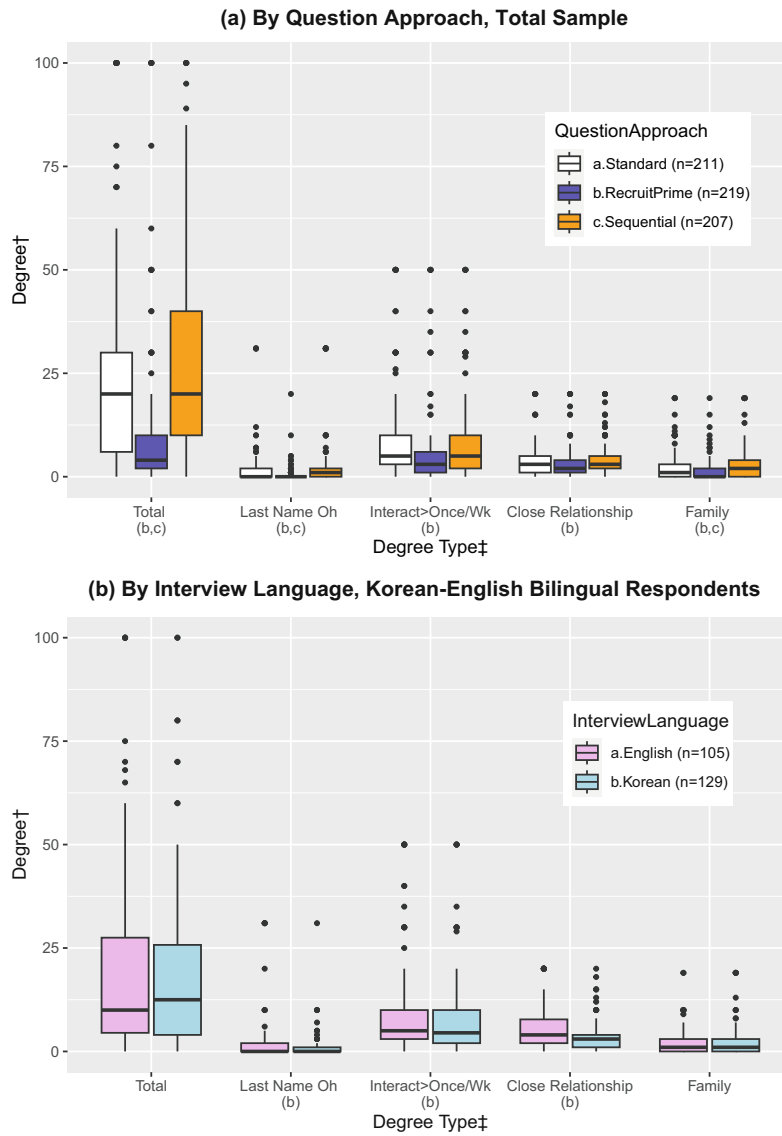
1.2 Culture, Language, and Social Networks

“Knowing someone” may take different forms depending on the culture (Ajrouch, Antonucci, and Janevic 2001). In a survey of US versus Korean college students (Kim, Sohn, and Choi 2011), US students reported far larger social networks than Korean students. A similar observation was made in a comparative study between China and the US (Qiu, Lin, and Leung 2012). Cultural values of individualism versus collectivism (Hofstede 2001; Kim et al. 1994) explain this difference with the US being an exemplary individualistic country and Korea/China a collectivistic one.

These two cultures contrast in how their members form, maintain and view social relationships. In-group membership carries a higher value in collectivistic than individualistic culture, which leads to fewer but more intimate and enduring relationships in collectivistic than individualistic culture (Triandis et al. 1988). Without specifying a boundary of social networks, it is likely that this cultural difference becomes pronounced. Social networks in collectivistic than individualistic culture already imply tighter relationships, and a “close” relationship may require even a greater level of closeness in collectivistic than individualistic culture.

**Fig. 4**

Distribution of Degrees. (Note. The sample size for each figure varies slightly due to differential item missing on degree variables. The red dotted lines indicate the location of the mean for each variable. SD Standard deviation, CV Coefficient of variation)

**Fig. 5**

*Distribution of Degrees by Experiment, Health and Life Study of Koreans. (Note. (b) and (c) indicate a significant coefficient ($p < 0.05$) of the recruitment prime degree question and of the sequential degree question, respectively, from the standard degree question estimated from gamma regression. †The total degree is shown up to 100 for an easier illustration. This resulted in 34 cases not displayed in the plot. ‡The total scale-up degree using the last name, “Oh”, is replaced with the non-scaled up version in this plot. As the former is a simple arithmetic proportional manipulation of the latter, the gamma regression results remain the same. *The sample size for each figure varies slightly due to differential item missing on degree variables.). (Note. (b) in parentheses indicates a significant coefficient ($p < 0.05$) of the Korean interview language compared to English estimated from gamma regression. †The total degree is shown up to 100 for an easier illustration. This resulted in 17 cases not displayed in the plot. ‡The total scale-up degree using the last name, “Oh”, is replaced with the non-scaled up version in this plot. As the former is a simple arithmetic proportional manipulation of the latter, the gamma regression results remain the same. *The sample size for each figure varies slightly due to differential item missing on degree variables.)*

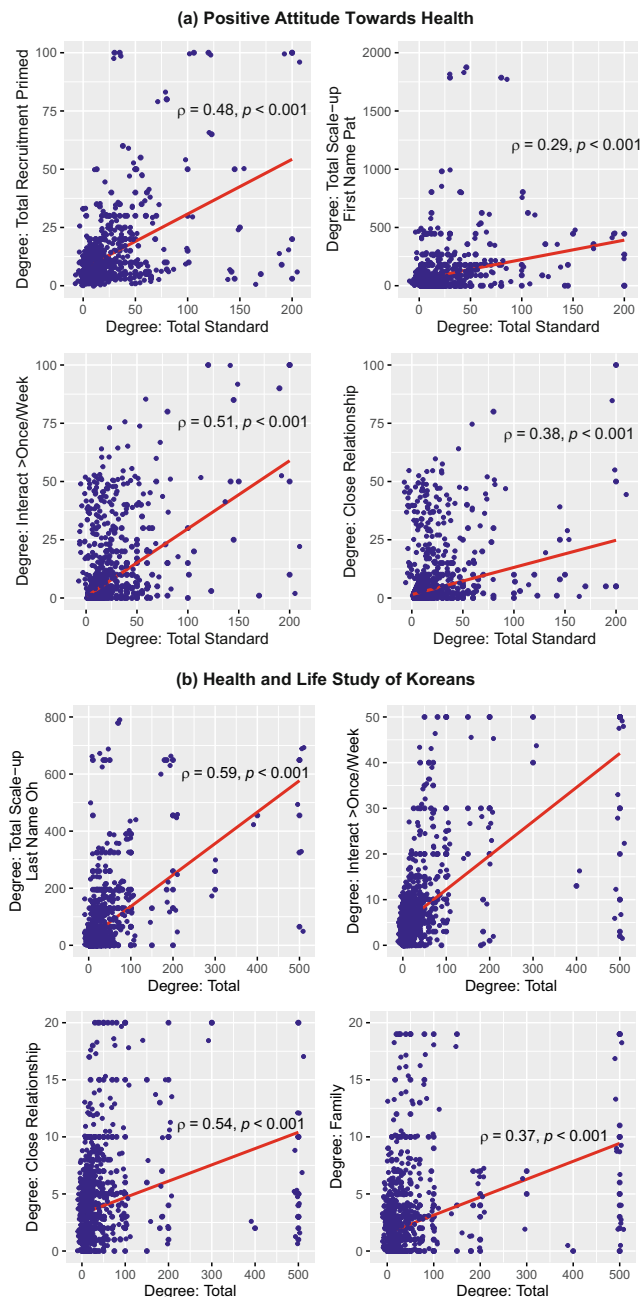


Fig. 6

Relationship between Degrees. (Note. The sample size for each figure varies slightly due to differential item missing on degrees. A small amount of jitter was introduced to avoid overlapping data points. Regression lines between the x and y axes are in red.)

This effect of culture may imply an interesting proposition: bilinguals' conceptualization of social networks differs by the language used in the conversation. Because language itself brings associated cultural norms to the conversation (Chen and Bond 2010; Ross, Xun, and Wilson 2002), those

fluent in a language of collectivistic culture as well as a language of individualistic culture may conceptualize their social networks with a tighter in-group boundary when using the language of collectivistic culture than its counterpart. This language effect has also been demonstrated with survey responses (Lee and Schwarz 2014). Therefore, in social network measurement, this may be expressed with reports of smaller network sizes when interviewed in languages associated with collectivistic than individualistic culture.

1.3 Social Networks and RDS

We argue that the social networks for RDS should be conceptualized in the context of *peer recruitment* rather than the general social network context. While being connected with a peer is a state, recruiting a peer to a research study is a behavior that may not conform to the norms of social interactions (Argyle 1969). In fact, in order for a person to recruit a peer, their relationship needs to be safe enough to bring in this unusual interaction, evidenced by nonrandomness in the RDS recruitment that favors certain types of relationships (e.g., kinship and close relationships in Lee et al. 2017, 2020; Young et al. 2014). The networks under these two contexts overlap to some extent but are not the same. The boundary of being connected is likely to be broader and more fluid than that of being able to recruit. The social network measurement in the literature goes beyond the first-order relationship, i.e., the links between an ego and their immediate alters; however, what matters for RDS recruitment is this first-order relationship. In sum, social networks in RDS are applicable to the first-order relationship in the context of peer recruitment, and the measurement error should be conceptualized around *recruitability*.

1.4 Measurement of Degrees in RDS

Bias is an obvious impact of inaccurate measurement of degrees in RDS (Mills et al. 2014). At the same time, degree questions in the practice of RDS vary widely. Some may use one simple question that covers the total network, such as, "How many Latino gay, bisexual or transgender over 18 in Chicago do you know?" (Ramirez-Valles 2013). Sometimes, boundary specificities are added (noted with in italics), such as, "How many people *whose cell number you know* use methamphetamine?" (Wendel et al. 2014), or "How many people do you *know personally* (i.e., *you know their name, you know who they are, and they know you, and you have seen them in the last 6 months*) who use heroin, methamphetamines, and/or powder or crack cocaine or who inject some other drug?" (Iguchi et al. 2010). The persons who inject drugs (PWID) component of the National

Table 2*Reported Degrees by Number of Recruits—Positive Attitude Towards Health^a*

Degree Type ^b	Number of Recruits							
	0		1		2		3	
	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.
Total Standard	23.2	2.5	29.1	4.1	24.6	4.1	26.3	4.6
Total Primed	11.8	1.2	15.5	1.8	12.0	1.9	14.4	2.8
Total Scale-Up First Name Pat	84.3	15.7	97.9	17.5	133.9	36.5	105.8	43.9
Interact > Once/Week	7.6	1.0	8.8	1.6	8.8	2.3	9.0	1.6
Close Relationship	4.2	0.6	5.7	1.3	3.9	0.8	3.8	0.8
<i>N</i>	158		99		65		45	

^aThose who did not receive coupons were excluded from the analysis; ^bSee Fig. 2 for the details about the degree types

HIV Behavioral Surveillance (NHBS) by the U.S. Centers for Disease Control and Prevention uses a series of questions as follows and takes the sum of the answers of those questions as a degree: “Please tell me how many males in [AREA] you know who inject and whom you have *seen in the past 30 days*” and “How many females in [AREA] do you know who inject and whom you have *seen in the past 30 days*?” This practice echoes a recommendation to “break the question into several parts” (Johnston and Sabin 2010). If PWID’s social networks favor the same sex, this question type may reveal a barrier error.

Although not explicit, these questions make certain assumptions about the social networks of a given population. For example, NHBS assumes that the PWID networks should be restricted to PWID that respondents have seen in the last 30 days and that live in the same area. However, these assumptions are rarely examined (Rudolph, Fuller, and Latkin 2013) and how well these degrees capture the RDS recruitment behavior is not known. It is not surprising that the empirical utility of degree weights is reported unclear (Lee et al. 2017; Martin, Johnson, and Hughes 2015) and that characterizing the RDS recruitment as a function of degree alone has been criticized as “fundamentally flawed” (Crawford 2016). One may argue that degrees do not need to be measured precisely, because the degree weights account for relative magnitudes. In other words, if everyone misreports the degree to be twice the true size, this does not affect the estimation. How realistic equal misreporting is certainly debatable. Even when not affecting point estimates, mismeasurement of degrees has been reported to lower estimate precision (Fellows 2022).

This study is the first of its kind and attempts to explore various ways to measure first-order degrees that tap into various aspects of social networks. We hypothesize that degrees reflecting close ties rather than general ties are advantageous for mirroring recruitment behaviors in RDS

and for population inference. For doing so, we use data from two RDS surveys that targeted two distinctive hard-to-reach groups: (1) PWID and (2) foreign-born Korean Americans. The nature of the social networks for these two groups is likely to differ. The PWID networks lack stability, lasting for a short term (Neaigus et al. 1995). PWID do not necessarily benefit from their social networks (Suh et al. 1997), although having a PWID partner may decrease substance use (Latkin et al. 1995). On the other hand, for ethnic minorities such as Korean Americans, social networks are persistent (Mollica, Gray, and Treviño 2003) and have positive effects, functioning as social capital (Anthias 2007) and a protective mediator of health conditions, such as depression (Ikram et al. 2016). Differences in the nature of social networks between these two groups imply that the scope of recruitable members may need to be understood with specific to the target population. Further, for the survey of Korean Americans, bilingual English-Korean speakers may report larger degrees when asked in English rather than in Korean. With Korean as a language provoking collectivistic cultural norms, degrees measured in Korean than English are expected to reflect closer relationships equating to higher recruitability.

2 Data and Methods

Two RDS surveys were implemented with degree measurement features unique to each survey. Hence, data sets are described separately. A more detailed operational description of these surveys is available elsewhere (Lee et al. 2021, 2020).

The common feature between the two surveys is the use of standard degree questions similar to those that measure a total network size. These surveys also asked degrees by priming respondents about peer recruitment and by adding

Table 3*Reported Degrees by Number of Recruits—Health and Life Study of Koreans^a*

Degree Type ^b	Number of Recruits					
	0		1		2	
	Mean	S.E.	Mean	S.E.	Mean	S.E.
Total	42.3	5.2	42.5	8.0	30.5	3.3
Total Scale-Up Last Name Oh	115.3	19.3	114.8	26.0	70.6	17.6
Interact > Once/Week	7.7	0.6	7.2	0.7	8.5	0.8
Close Relationship	3.7	0.2	3.7	0.3	4.6	0.4
Family	2.1	0.2	2.1	0.3	2.7	0.3
<i>N</i>	323		151		131	

^aThose who did not receive coupons were excluded from the analysis; ^bSee Fig. 3 for the details about the degree types

various naming stimuli that tapped into specific aspects of social relationships, such as relationship closeness, and by putting the relationship into certain contexts. One of the stimuli used people's names that can be applied to the scale-up method (Killworth et al. 1998; Lubbers, Molina, and Valenzuela-García 2019; McCormick et al. 2010). This may allow us to project the “true” total degree discussed more in detail in the Analysis Steps.

There are important features unique to the Korean American survey. First, we implemented two randomized experiments, a rare feature in the social network literature (Sudman 1985, 1988): (1) with the total sample, we experimented three different wordings of the total degree question; and (2) with bilingual Korean-English speakers, we randomly assigned questionnaire language. The first experiment offers an opportunity to examine the question wording sensitivity. With the second experiment, we examine the language effect on degree reports. Another feature to note is an external validation dataset that informs degree question choices for population inference. These will be discussed more shortly.

2.1 Survey 1: Project Positive Attitudes Towards Health

Project Positive Attitudes Towards Health (PATH) was an in-person RDS survey of PWID in Southeast Michigan, conducted from May to November 2017. From 46 seeds, a total of 410 PWID were recruited. For each participant, 3 coupons were issued unless they reported having fewer than 3 PWID peers; in the latter case, the number of issued coupons corresponded to the reported count of PWID peers. A total of 43 participants reported having 0 PWID peers and did not receive coupons. PATH study protocols followed those of NHBS-PWID (Centers for Disease Control and Prevention 2015). The socio-demographic characteristics of 410 participants recruited through 14 waves are available in Table 1. PATH data are publicly available (Lee and Roddy 2021). The University of Michigan institutional review board approved the study protocol (IRB # HUM00118004).

Table 4*Comparison of Quasi-Binomial Models Predicting Recruitment Success—Positive Attitude Towards Health (*n* = 410^c)*

Model	Deviance Δ	Standardized β	S.E.	<i>p</i>
Baseline (BL) ^a	663.0 ^b			
BL + Log(Degree): Total Standard	−0.890	0.068	0.094	0.469
BL + Log(Degree): Total Primed	−0.005	0.005	0.096	0.955
BL + Log(Degree): Total Scale-Up First Name Pat	−1.264	0.078	0.090	0.387
BL + Log(Degree): Interact > Once/Week	−3.565	0.138	0.095	0.148
BL + Log(Degree): Close Relationship	−0.012	−0.008	0.093	0.932

^aCovariates of the baseline model are in Table 1. See Lee et al. (2020) for details about estimated coefficients of the predictors of recruitment success; ^b *df* = 313, ^cThe sample size varies slightly by model due to differential item missing on degree variables

Table 5

Comparison of Quasi-Binomial Models Predicting Recruitment Success—Health and Life Study of Koreans, Total Sample ($n = 637^c$)

Model	Deviance Δ	Standardized β	S.E.	p
1. Baseline (BL) ^a	1042.3 ^b			
2. BL + Log(Degree): Total Standard	−0.387	0.042	0.081	0.631
3. BL + Log(Degree): Total Scale-Up Last Name Oh	−1.087	0.069	0.079	0.385
4. BL + Log(Degree): Interact > Once/Week	−6.979	0.183	0.083	0.028
5. BL + Log(Degree): Close Relationship	−18.911	0.298	0.083	<0.001
6. BL + Log(Degree): Family	−10.180	0.230	0.088	0.009

^aCovariates of the baseline model are in Table 1. See Lee et al. (2020) for details about estimated coefficients of the predictors of recruitment success; ^b $df = 565$; ^cThe sample size varies slightly by model due to differential item missing on degree variables

Table 6

Comparison of Quasi-Binomial Models Predicting Recruitment Success—Health and Life Study of Koreans, Korean-English Bilingual Respondents ($n = 234^c$)

Model	English ($n = 105$)				Korean ($n = 129$)			
	Deviance Δ	Std β	S.E.	p	Deviance Δ	Std β	S.E.	p
1. Baseline (BL) ^a	109.1 ^b				191.6 ^d			
2. BL + Log(Degree): Total Standard	−0.038	−0.044	0.243	0.857	−5.269	0.364	0.195	0.065
3. BL + Log(Degree): Total Scale-Up Last Name Oh	−0.443	0.139	0.224	0.537	−0.651	0.127	0.192	0.509
4. BL + Log(Degree): Interact > Once/Week	−0.166	0.089	0.235	0.706	−16.066	0.647	0.205	0.002
5. BL + Log(Degree): Close Relationship	−2.113	0.298	0.225	0.189	−19.479	0.743	0.212	0.001
6. BL + Log(Degree): Family	−0.805	0.195	0.235	0.409	−7.708	0.448	0.195	0.023

^aCovariates of the baseline model are in Table 1. See Lee et al. (2020) for details about estimated coefficients of the predictors of recruitment success; ^b $df = 79$; ^cThe sample size varies slightly by model due to differential item missing on degree variables; ^d $df = 104$

2.1.1 Degree Questions

The degree in PATH was the number of PWID who participants knew in Southeast Michigan. For a given respondent, a total of five types of degrees were asked. The total degree was measured in two contexts: (1) standard NHBS questions introduced earlier; and (2) a single question that primed about the recruitment request. We measured three specific types of degrees with naming stimuli: (3) a first name, “Pat,” and its variants (e.g., Patrick, Patricia, etc.) which pertaining to the scale-up method; (4) interacting more than once a week; and (5) close relationships. These stimuli were applied to degrees measured without recruitment priming, by asking questions such as, “Among [TOTAL DEGREE], how many do you interact with (including talking to, visiting with, calling, emailing, texting, etc.) personally more than once week?” See Fig. 2 for the exact wording of all degree questions and the sequence. These

five measures will be termed, “total standard”, “total recruitment primed”, “total scale-up”, “frequent interaction/contact”, and “close relationship” degrees throughout the paper.

2.2 Survey 2: Health and Life Study of Koreans

Health and Life Study of Koreans (HLSK), a Web-RDS survey of Korean immigrant adults in Los Angeles County, California and the State of Michigan, was conducted between June 2016 and March 2018. The data collection included 222 seeds from whom 637 participants were recruited (358 in Los Angeles and 279 in Michigan) through 12 waves. Participants were issued with two coupons, unless they reported knowing fewer than two peers or participated at the end of the survey period. Out of 637 participants, 54 received no or one coupon. The questionnaire was

Table 7

Comparison of Unweighted and Various Degree Weighted Estimates from the Health and Life Study of Koreans with Population Characteristics from the American Community Survey 2012–2016^a

Age 18–29 yrs		Male		Married		4 year college		HH inc. > \$50T		Employed	
%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
<i>American Community Survey (n = 8772)</i>											
12.3	0.2	43.9	0.2	63.8	0.3	35.3	0.4	54.4	0.5	56.2	0.3
<i>Health and Life Study of Koreans (n = 637)</i>											
Unweighted											
36.9	1.9	39.4	1.9	54.6	2.0	40.0	1.9	42.8	2.0	55.4	2.0
Weighted—Degree: Total											
40.7	6.8	38.2	7.3	51.3	7.2	42.3	6.9	39.7	7.5	51.5	7.4
Weighted—Degree: Total Scale-Up Last Name Oh											
32.3	5.3	39.3	5.1	57.2	5.5	34.7	5.6	40.5	5.4	57.2	5.2
Weighted—Degree: Interact > Once/Week											
40.2	4.4	37.4	4.8	53.7	4.7	41.5	4.7	42.8	4.7	53.5	4.5
Weighted—Degree: Close Relationship											
37.9	3.7	38.5	3.7	55.9	4.0	40.8	3.7	43.3	3.9	53.1	3.8
Weighted—Degree: Family											
29.1	4.2	37.9	4.4	67.9	4.5	39.8	4.3	42.3	4.7	52.4	4.4

^aThe ACS sample included foreign-born Korean adults in Los Angeles County, California and State of Michigan to match the target population of the HLSK

prepared in both English and Korean. Participants started the questionnaire in the language of their choice and, shortly after that, were asked about their language use. We considered those who reported speaking English and Korean equally well ($n = 234$) as “bilingual” and randomly assigned the questionnaire language for the remainder of the survey. Among the bilinguals, 129 completed in Korean and 105 in English. Table 1 includes characteristics of the HLSK sample, and the data are available publicly (Lee 2020). The University of Michigan institutional review board approved the study protocol (IRB # HUM00114530).

2.2.1 Degree Questions

The degree in HLSK was the number of Korean immigrant adults who participants knew in respective sites. See Fig. 3 for the exact wording of all degree questions and the sequence. The total degree was measured in an experiment where a given respondent was asked one of the three versions: a direct standard question asking how many foreign-born Korean adults a respondent knew; a sequential version with three questions, starting with how many Koreans a respondent knew in the corresponding site, followed by how many were adults and then how many were foreign-born; and a single question with priming about the recruitment

request that added, “We will give you invitation notes that you can give to other Koreans,” to the direct standard question. Through the randomization, 211, 207 and 219 respondents were assigned to the first, second and third versions, respectively.

Additionally, we asked specific degrees with four naming stimuli: the last name, “Oh”, pertaining to the scale-up method; frequent interactions/contacts; relationship closeness; and family relationships. Naming stimuli questions were not an experiment and were asked of all respondents. Overall, a total of five questions on degrees were asked per respondent. They will be termed, “total”, “total scale-up”, “frequent interaction/contact”, “close relationship”, and “family” degrees throughout the paper.

2.2.2 External Validation Data

Because RDS is conducted typically with hard-to-reach populations, an empirical examination of population inference is not always feasible. Fortunately, for HLSK, its population data are available from the American Community Survey (ACS). In particular, we use the weighted sample of the foreign-born Korean adults in Los Angeles County and the State of Michigan in ACS 2012–2016 Public Use Microdata Sample (PUMS) as a validation dataset. The focus

is on six characteristics available in both ACS and HLSK: age, sex, marital status, education, income and employment.

2.3 Analysis Steps

The analyses are done in four steps. The first three steps use the common features (i.e., various degree measures) in both surveys with some variations associated with experimental features in the second survey. The last step focuses on the utility of degrees for population inference with the second survey (HLSK).

The first step focuses on the degrees. In both surveys, for a given respondent, there are five versions of degrees as shown in Figs. 2 and 3. Specifically, in PATH, the total degree of a given respondent can be derived from three sources: the standard degree using NHBS questions; the recruitment prime degree; and the scale-up degree based on the “Pat” stimulus; and in HLSK, from two sources: one randomly selected from the three total degree questions; and the scale-up degree based on the last name, “Oh”. The first/last name measures require scaling up the responses. For PATH, while the exact share of the first name, “Pat,” in the PWID population is unknown, data from Tzioumis (2018) and the U.S. Social Security Administration (<https://www.ssa.gov/oact/babynames/>) suggest the share to be around 1.04–1%. This means that, for a given person who knows one person with such a first name, we can expect that person’s total network size to be 83 (= 1/1%) to 96 (= 1/1%). For HLSK, the scale-up method uses the number of Korean immigrants with the last name, “Oh,” that a participant reported to know. According to the Statistics Korea (2017), as of 2015, 2% of its population have this last name. This allows us to project, for example, people who report to know 1 Korean person with this last name to know about 65 (= 1/2%) Koreans in total. Using these statistics, we derive a scale-up version of total degrees for all respondents in both surveys. Further, there are degrees tapping into specific relationships: the degrees in frequent interactions and the degrees in close relationships in both surveys and the family degrees in HLSK.

We examine the distribution of all degrees, including their relative variance through coefficients of variation (CVs) and compare different versions of total degrees through Wilcoxon signed-rank tests in order to accommodate the non-normal distribution of the reported degrees.

For HLSK, we examine degrees as a function of two experimental factors: total degree question approaches for the overall sample; and interview languages for the bilingual respondents. To compensate for positively skewed degree distributions, the effect of these experimental factors on degree reports is examined through gamma regression. The analysis of question approaches will show how sensitive

degree distributions are to the choice of question wording, and the analysis of interview language will show the existence and direction of language effects on degree reports.

In the second step, we examine the relationship between the total, scale-up total and remaining degrees estimated through Spearman correlation coefficients and scatterplots. If degree measures are similar in nature, the correlation coefficients will not only be positive but also similar in their magnitudes.

The third analysis step examines degrees in relation to actual recruitment, which is used as a validation measure based on the rationale of recruitability. We first examine reports on various degrees by the number of recruits among participants who received coupons. We then focus on the recruitment success, Y , a binomial variable indicating whether each issued coupon was redeemed by a peer. Because of its dispersion issues, we let $Y \sim \text{quasibinom}(n, p)$, where n is the number of issued coupons and p is the probability of each coupon being redeemed and use quasi-binomial regression for $\log\left(\frac{p}{1-p}\right) = \beta X$. For the covariates (X), we start with a baseline model that includes socio-demographic characteristics listed in Table 1. We then add one of the five different measures of degrees introduced above one by one to the baseline model and compare their slope coefficient estimates as well as their model fit through deviance. In these models, we log-transform degrees to mitigate outlier effects. For easier comparisons, we use standardized coefficient estimates. The degree with a significant slope estimate and/or with the best model fit (i.e., the smallest deviance) is considered as advantageous, as these indicate degree’s utility for explaining recruitment success. With HLSK, we examine whether the model assessment done with the overall sample holds the same between the two interview languages using data from the bilingual respondents.

Note that our focus is on the relationships among the individuals sampled through the RDS mechanism—in a repeated sampling context, we are attempting to understand the associations among the individuals within a network that would appear if the RDS were to be conducted repeatedly. The use of weights in multivariate models using RDS data is a debatable topic (Avery et al. 2019), and it is less clear how to interpret the results of our models in this context (as opposed to a multivariate substantive model focused on population associations of interest). For this reason, we present regression analysis unweighted by degrees. However, for the sake of completeness, the first step of the analysis was done using non-transformed degrees and the second step with degree weights. The results were consistent between transformed and non-transformed degrees as well as between weighted and unweighted regression models.

The last step compares estimates from HLSK against population estimates from ACS, focusing not only the biases but also the variance of HLSK estimates. HLSK point estimates are computed with the RDS-II estimator (Volz and Heckathorn 2008) and their standard errors are estimated using the bootstrap method (Salganik 2006), available in the R RDS package (Handcock et al. 2017). We apply various degree weights to the RDS-II estimator in order to examine the changes in the point estimates and how they stand against ACS estimates and the changes in the precision. RDS-II estimator was selected because, with its resemblance to the Horvitz-Thompson estimator, it is commonly used and the performance of existing RDS estimators is reported to be similar when applied to real-world data (e.g., Crawford 2016; Lee et al. 2017; Lee, Ong, Chen, et al. 2020; Martin et al. 2015). We also include unweighted HLSK estimates with their standard errors calculated using the delta method (Ver Hoef 2012), which provides another angle for assessing validity of the degree weights.

3 Results

3.1 Degree Distribution

Not all participants reported degrees. Out of 410 PATH participants, the total degree was missing at 5%; the primed total degree at 10%; the first name, “Pat”, degree at 7%; the frequent interaction degree at 7%; and the close relation degree at 7%. No specific patterns were observed in these missing rates (e.g., by wave). In HLSK, missing rates on the degree questions were small ranging from 1% on the total degree to 4% across all four specific degrees, with no specific patterns.

It should be noted that there were extremely large and potentially unrealistic values reported in HLSK. For example, the total degree ranged to 300,000; the frequent contact degree to 250,000; and the family degree to 100,000 (see Appendix 1). The close-relationship degree, however, ranged only to 50. Consequently, their CVs were large, well above 10, except for the close relationship degree (CV = 1.28). To mitigate issues with potentially unrealistic values, we treated values above the 98th percentile under each measure as outliers and top-coded them at the 98th percentile values reported in Appendix 1 for the remainder of the analysis.

Fig. 4a, b summarize degree distributions. All measures are positively skewed, with few respondents reporting large degrees. The close-relationship degrees showed the smallest gap between the means and medians in both surveys. On average, PATH respondents reported knowing a total of 24.31 (SD: 32.24) PWID in the area who they saw in the

past 30 days. Once primed about recruitment, respondents reported a total degree of 13.05 (SD: 16.32), significantly smaller than the standard version (Wilcoxon signed-rank tests statistic (W) = 30,941; $p < 0.001$). In fact, 57% of the participants reported smaller degrees on the primed version than the standard one. When using the scale-up method on the “Pat” degree, the average total degree is projected to be 98.86 (SD: 211.83), significantly larger than both the standard (W = 24,165; $p < 0.001$) and primed versions of the total degree (W = 18,923; $p < 0.001$). The scale-up total degree was larger than the standard total degree for 44% of the respondents.

In HLSK where the respondents reported knowing an average of 38.05 (SD: 84.65) foreign-born Korean adults, a similar pattern emerged with the scale-up total degree (103.41, SD: 309.2) marginally significantly larger than the reported total degree (W = 83,239; $p = 0.022$). In fact, for 35% of the HLSK respondents, the scale-up version was larger than the reported total degree.

Once the relationship types were specified, the degrees were reported as much smaller. On average, PATH respondents reported interacting with 7.79 (SD: 13.66) PWID more than once a week and feeling close to 4.35 (SD: 8.32) PWID, and HLSK respondents reported interacting with 7.67 (SD: 10.18) Korean immigrant adults more than once a week, feeling close to 3.83 (SD: 4.06) Korean immigrant adults and having 2.19 (SD: 3.59) Korean immigrant family members. The CVs ranged 1.06–2.99 with the largest CV observed for the scale-up degrees in both surveys and the smallest for the close relationship degree in HLSK.

We examined the five different measures of degrees in HLSK as a function of three total question approaches introduced earlier in Fig. 5a: total degrees were reported to be significantly smaller when asked in a single question with recruitment priming than the standard approach. The total degree averaged at 45.30 (SD: 89.86) under the standard approach and at 17.49 (SD: 55.83) under recruitment prime. The sequential approach did not produce different degrees from the standard approach, except for the family degree, where the average degree was significantly larger at 3.03 (SD: 4.28) under the sequential approach and 2.21 (SD: 3.65) under the standard approach.

When comparing the total degrees between Korean and English questionnaires among bilingual respondents in Fig. 5b, the total degree was smaller for those who responded in Korean (36.78, SD: 80.97) than in English (50.90, SD: 105.23) but not statistically significant. When examining specific types of degrees, the language significantly affected the “Oh” degree, the frequent interaction degree and the close relationship degree. Bilinguals assigned to Korean reported significantly smaller degrees: 1.22 (SD: 3.30) for the last name; “Oh”, 7.66 (SD: 10.35) for the frequent contact and 3.54 (SD: 3.74) for the close

relationship, compared to those assigned to English who reported 2.69 (SD: 6.56), 10.52 (SD: 13.43) and 5.96 (SD: 5.77) on these respective degrees.

3.2 Relationship across Degrees

The relationships between the total degree and other degrees were significant and positive in both surveys as shown in Fig. 6a, b. However, the magnitude of these relationships varied, from $\rho = 0.29$ between the total degrees of PWID and the total degree of PWID scaled up using the first name, “Pat,” to $\rho = 0.69$ between the total degree of Korean immigrants and the degrees coming from the number of Korean immigrant adults with whom participants interacted more than once a week.

3.3 Degrees and Recruitment Success

On average, an issued coupon was redeemed at 34% of the time in the PATH study. Among 367 participants to whom coupons were issued, 158 did not recruit anyone (Table 2). Only 12% of participants ($n = 45$) successfully recruited 3 alters. Among all degrees, only the frequent contact degree was associated with the number of recruits in that this degree was an increasing fashion as the number of recruits increased.

Similar to PATH, 35% of the coupons issued were redeemed in HLSK. At the participant level, out of 605 who received coupons, less than half recruited at least one alter (Table 3). Total degrees based on the standard questions and on the scale-up method were the smallest for those who recruited 2 alters than the counterparts. On the other hand, the frequent interaction degree, the close relationship degree and the family degree were the largest among participants who recruited 2 alters compared to those who recruited no or 1 alter. (Note that details of PATH and HLSK recruitment success are reported by Lee, Ong, and Elliott (2020).)

We examined the coupon redemption pattern as a function of the degrees, while controlling for socio-demographic characteristics listed in Table 1. Tables 4, 5 and 6 includes the comparisons between the baseline model and the models where each degree was added as an independent variable to the baseline model. For PATH (Table 4), the deviance of the baseline model was 663.0 ($df = 313$), and the model using the frequent contact degree showed the largest decrease in deviance. This model also showed the largest standardized regression coefficient of the degree at 0.138. However, none of the degrees significantly predicted recruitment success, and the decrease in deviance was not large across all models including degrees.

The baseline model of HLSK showed a deviance of 1042.3 (Table 5). Adding the standard total degree or the scale-up total degree to the baseline model decreased the deviance only minimally with Δ deviance far less than -1 . The decrease in the deviance was the largest when adding the close-relationship degree (Δ deviance = -18.911) to the baseline model, followed by the family degree (Δ deviance = -10.180) and the frequent contact degree (Δ deviance = -6.979). Among all five types of degrees, these three showed a significant and positive relationship with recruitment success.

We examined the predictive value of degrees focusing on the bilingual respondents of HLSK in Table 6. For those interviewed in English, none of the degrees added much to the baseline model in predicting recruitment success. However, in Korean, the close relationship degree, the frequent contact degree and the family degree when added to the baseline model, reduced the deviance substantially (Δ deviance = -19.472 for the close contact degree; Δ deviance = -16.066 for the frequent contact degree; Δ deviance = -7.708 for the family degree) and showed a significant and positive relationship with recruitment success. Notably, in Korean interviews, even the standard degree was a marginally significant predictor of recruitment.

3.4 Population Inference

Table 7 compares HLSK estimates against ACS estimates for the Korean immigrant population. Overall, HLSK estimates differed from the ACS population estimates across all six characteristics examined. For example, in the population, 12% of foreign-born Korean adults were aged 18–29 years old. Based on the HLSK sample, this estimate was 37% (S.E.: 2%) without weights. Weighted HLSK estimates ranged from 29% (S.E.: 4%) to 41% (S.E.: 7%) depending on the type of degree used. Compared to the unweighted estimates, weighted estimates with the total standard degree showed larger biases on all characteristics, although not statistically significant. This statistical insignificance was partially due to large standard errors of estimates that were weighted by the total degree, which were almost four times larger than those of unweighted estimates. For example, the standard error of the proportion of ages 18–29 years old was 2% when unweighted and 7% when weighted by the total degree; and with weights based on degrees tapping into specific social relationships, the standard errors were around 4%.

Biases of the specific degree weighted estimates were similar to those of unweighted estimates. Importantly, among weighted estimates, those weighted by the close relationship degree were the most precise (i.e., smallest standard errors) across all six characteristics. In fact, they

were twice more precise than the standard degree weighted estimates, as standard errors of the former were almost roughly half of those of the latter.

4 Summary and Discussions

This paper examined degrees, an important feature of RDS, using data from two surveys that measured respondents' network sizes in various ways. In both surveys, all degrees, regardless of their types, were positively skewed. While impossible to verify with the available data, some HLSK respondents reported degrees that were not entirely realistic (e.g., having 100,000 family members who are Korean immigrants), which may imply careless reports and a weight approaching 0. Relatively speaking, there was a larger variance in the standard total degrees than the specific degrees with naming stimuli that tap into close and specific social relationships (e.g., frequent contacts), in turn, implying a greater level of noise in the total than specific degree measures.

Regardless of whether examined experimentally or non-experimentally, respondents reported knowing significantly fewer people when primed about the recruitment request by 12.0 in PATH and 28.8 in HLSK. Clearly, a subtle hint about the peer recruitment request prompted them to report much smaller network sizes. This implies not only that degrees are sensitive to question wording but also that the boundary of recruitable peers in one's social network is smaller than the boundary of the standard/total social network, a clear mismatch between recruitability and the typical degree questions. This last point echoes our results indicating the total degree not being predictive of recruitment success and raises a question about its utility as an adjustment factor in RDS estimators. Rather, it was the specific degrees such as those measuring peers with whom respondents interacted frequently in the survey of PWID and to whom respondents felt close in the survey of Korean immigrants that were associated with recruitment success. Referring back to the differences in the nature of the social networks between PWID and Korean immigrants, different groups may leverage different types of relationships when it comes to recruiting peers into RDS studies.

Contribution of weight adjustments based on the standard total degrees to population inference was unclear. Reflecting the large variation in the standard degrees, corresponding point estimates were subject to large variances, while this type of degrees showed no clear utility to correcting biases in the characteristics examined with HLSK and ACS. Compared to estimates weighted by the standard degrees, when weighted by degrees tapping into specific social relationships (close relationships, family relationships, having frequent contacts), the estimates in the direction of

improved inference as the biases were smaller and variances were smaller. However, they were not better than unweighted estimates, which echoes empirical assessments of weighted RDS estimates (Crawford 2016; Lee et al. 2021; Martin et al. 2015). When putting together, while degrees are conceptually relevant to RDS and its inference, our study did not support that the degrees as currently measured in RDS studies and used in RDS estimators provided utility to the inference.

Due to its exploratory nature, there are a number of limitations in this study. First, our illustration of the recruitment process around Fig. 1 and m_i is overly simplified, compared to real-world recruitment. Characteristics of recruits and alters, recruitment with vs. without replacement, and the point where recruiters stop their recruitment effort are some of the examples of factors affecting the recruitment process. Neither does the extant literature provide empirical evidence about their roles nor incorporate them into mathematical illustrations. Second, our analysis ignored recruiter-recruit dependence, which may have resulted in underestimated precision (and thus overstated statistical significance) of our recruitment success models. We note that the recruiter-recruit dependence existed in both surveys at varying levels depending on the characteristics (e.g., high dependence on age and race/ethnicity in PATH and on age and marital status in HLSK but low on education and employment status in both PATH and HLSK). Still, this pattern did not vary by recruit counts. Third, while specific degrees were favored over the total degrees for constructing degree weights in our examination of the population inference, our comparison was limited to six characteristics. Whether this result holds beyond these characteristics remains to be answered. Despite potential limitations, our study signals that the degrees reflecting recruitability outperform the standard degrees currently measured and used in the RDS literature.

One may question the fact that this study relied on two surveys targeting special populations, which may cause concerns of external validity. We note that, despite the two target populations being distinct (Korean immigrants vs. persons who inject drugs), coherent findings emerged. This, in fact, strengthens our point that the measurement of social networks in the contexts of RDS should reflect recruitability regardless of the population types and that questions that reflect close relationships may prove.

This calls for thorough examinations of degree measurement in the context of RDS recruitment. At a minimum, our surveys suggest that the scope of *recruitable* members within a person's social network is different than the degrees currently measured in RDS practice and may need to be understood with considerations specifically on the characteristics of a given population. This, in turn, may serve as the foundation for developing questions measuring degrees. Recall that Korean-English bilingual HLSK respondents re-

ported fewer close relationships when given a Korean than English questionnaire and that the degrees measured in Korean interviews were a significant predictor of recruitment but not in English interviews. This corresponds to the literature on social network in collectivistic and individualistic cultures (Kim et al. 2011; Qiu et al. 2012; Triandis et al. 1988). “Knowing someone” or being “close to someone” in a collectivistic society may require a stronger relationship than in an individualistic society. Further, for unusual social behaviors such as peer recruitment to take place without repercussions, this may be even more pronounced in the collectivistic than individualistic culture. This significant language effect illustrates the level of attention that needs to be given to the degree measurement.

Noteworthy are the implications of the limitations of the network scale-up method and the similarities between degrees measured with a single question or multiple sequential questions. The network scale-up method is a relatively new concept shown to be promising for overcoming difficulties in estimating network or population sizes (Maltiel et al. 2015; McCormick et al. 2010). This, however, did not translate particularly well for RDS degrees. It may be that, at least in the case of PATH, the composition of first names in PWID’s networks does not align well with that of the general population. While we expected asking network sizes through multiple questions sequentially funneling to a specific boundary (from Korean Americans to foreign-born Korean American adults) to produce smaller sizes than asking a single question imposing the same specific boundary, to our surprise, that did not make a difference. It may be that, in the case of Korean Americans, whether US- or foreign-born and whether adults or not are important and salient characteristics of alters forming a social network, making a single question as effective as multiple questions sequentially narrowing the scope.

Importantly, as degree measurement relies on respondents’ self-reports, questions ascertaining degrees should be easy to answer. Smaller variability of the specific degrees and total degrees under the recruitment prime, compared to the standard total degrees, observed in this study may be evidence for higher reporting difficulties with for the standard degree question than the counterparts.

As these specific degrees were more predictive of recruitment success (and better at correcting for biases in the case of HLSK) than the standard total degrees, questions tapping into different aspects of social relationships pertinent to peer recruitment, such as close relationship, through naming stimuli that are easy to answer and better reflect recruitability may prove to be a fruitful direction for future investigations. Further, acknowledging that accurate measurement may not be attainable, it may be worthwhile to pursue developing pseudo-weights through modelling reported degrees (Elliott 2013). In order to minimize the effect of

unnecessary loss of precision (Fellows 2022), such pseudo-weights may need to balance the trade-off between bias and variance.

Acknowledgments This work was supported by the National Science Foundation [grant number SES-1461470] and the National Institute on Aging of the National Institutes of Health [grant numbers R01 AG060936-01; R21 AG062844-01].

References

- Agans, R.P., Zeng, D., Shook-Sa, B.E., Boynton, M.H., Brewer, N.T., Sutfin, E.L., Goldstein, A.O., Noar, S.M., Vallejos, Q., Queen, T.L., Bowling, J.M., & Ribisl, K.M. (2021). Using social networks to supplement RDD telephone surveys to oversample hard-to-reach populations: a new RDD+RDS approach. *Sociological Methodology*, 51(2), 270–289. <https://doi.org/10.1177/00811750211003922>.
- Ajrouch, K.J., Antonucci, T.C., & Janevic, M.R. (2001). Social networks among blacks and whites: the interaction between race and age. *Journals of Gerontology—Series B Psychological Sciences and Social Sciences*, 56(2), 112–118. <https://doi.org/10.1093/geronb/56.2.S112>.
- Anthias, F. (2007). Ethnic ties: social capital and the question of mobilisability. *The Sociological Review*, 55(4), 788–805. <https://doi.org/10.1111/j.1467-954X.2007.00752.x>.
- Argyle, M. (1969). *Social interaction: process and products*. London: Methuen.
- Avery, L., Rotondi, N., McKnight, C., Firestone, M., Smylie, J., & Rotondi, M. (2019). Unweighted regression models perform better than weighted regression techniques for respondent-driven sampling data: results from a simulation study. *BMC Medical Research Methodology*, 19(1), 202. <https://doi.org/10.1186/s12874-019-0842-5>.
- Berchenko, Y., Rosenblatt, J.D., & Frost, S.D.W. (2017). Modeling and analyzing respondent-driven sampling as a counting process. *Biometrics*, 73(4), 1189–1198. <https://doi.org/10.1111/BIOM.12678>.
- Centers for Disease Control and Prevention (2015). *National HIV behavioral surveillance, injection drug use—round 4: operations manual*. Atlanta: Centers for Disease Control and Prevention.
- Chen, S.X., & Bond, M.H. (2010). Two languages, two personalities? Examining language effects on the expression of personality in a bilingual context. *Personality and Social Psychology Bulletin*, 36(11), 1514–1528. <https://doi.org/10.1177/0146167210385360>.

- Crawford, F.W. (2016). The graphical structure of respondent-driven sampling. *Sociological Methodology*, 46(1), 187–211. <https://doi.org/10.1177/0081175016641713>.
- Elliott, M.R. (2013). Combining data from probability and non-probability samples using pseudo-weights. *Survey Practice*, 2, 6.
- Fellows, I.E. (2022). On the robustness of respondent-driven sampling estimators to measurement error. *Journal of Survey Statistics and Methodology*, 10(2), 377–396. <https://doi.org/10.1093/JSSAM/SMAB056>.
- Fisher, J.C., & Merli, M.G. (2014). Stickiness of respondent-driven sampling recruitment chains. *Network Science*, 2(2), 298–301. <https://doi.org/10.1017/nws.2014.16>.
- Forrest, J.I., Lachowsky, N.J., Lal, A., Zishan, C., Sereda, P., Raymond, H.F., Ogilvie, G., Roth, E.A., Moore, D., & Hogg, R.S. (2016). Factors associated with productive recruiting in a respondent-driven sample of men who have sex with men in Vancouver, Canada. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, 93(2), 379–387. <https://doi.org/10.1007/S11524-016-0032-2>.
- Gile, K.J. (2011). Improved inference for respondent-driven sampling data with application to HIV prevalence estimation. *Journal of the American Statistical Association*, 106(493), 135–146. <https://doi.org/10.1198/jasa.2011.ap09475>.
- Gile, K.J., & Handcock, M.S. (2010). Respondent-driven sampling: an assessment of current methodology. *Sociological Methodology*, 40(1), 285–327. <https://doi.org/10.1111/j.1467-9531.2010.01223.x>.
- Handcock, M.S., Gile, K.J., Fellows, I.E., & Whipple, W.N. (2017). *Package “RDS”: respondent-driven sampling*. <https://doi.org/10.1111/j.1467-9531.2010.01223.x>.
- Heckathorn, D.D. (1997). Respondent-driven sampling: a new approach to the study of hidden populations. *Social Problems*, 44(2), 174–199. <https://doi.org/10.2307/3096941>.
- Heckathorn, D.D. (2007). Extensions of respondent-driven sampling: analyzing continuous variables and controlling for differential recruitment. *Sociological Methodology*, 37(1), 151–208. <https://doi.org/10.1111/j.1467-9531.2007.00188.x>.
- Hofstede, G. (2001). Culture’s recent consequences: using dimension scores in theory and research. *International Journal of Cross Cultural Management*, 1(1), 11–17. <https://doi.org/10.1177/147059580111002>.
- Iguchi, M.Y., Sandra, H.B., Ober, A.J., Fain, T., Heckathorn, D.D., Gorbach, P.M., Heimer, R., Kozlov, A., Ouellet, L.J., Shoptaw, S., & Zule, W. (2010). *Sexual acquisition and transmission of HIV cooperative agreement program (SATHCAP), 2006–2008 [United States] restricted use files*. ICPSR29181-V1. Ann Arbor: Inter-university Consortium for Political and Social Research. <https://doi.org/10.3886/ICPSR29181>.
- Ikram, U.Z., Snijder, M.B., de Wit, M.A.S., Schene, A.H., Stronks, K., & Kunst, A.E. (2016). Perceived ethnic discrimination and depressive symptoms: the buffering effects of ethnic identity, religion and ethnic social network. *Social Psychiatry and Psychiatric Epidemiology*, 51(5), 679–688. <https://doi.org/10.1007/S00127-016-1186-7/TABLES/3>.
- Johnston, L.G., & Sabin, K. (2010). Sampling hard-to-reach populations with respondent driven sampling. *Methodological Innovations Online*, 5(2), 38.1–38.48. <https://doi.org/10.4256/mio.2010.0017>.
- Kadushin, C. (2012). *Understanding social networks: theories, concepts, and findings*. New York: Oxford University Press.
- Kadushin, C., Killworth, P.D., Bernard, H.R., & Beveridge, A.A. (2006). Scale-up methods as applied to estimates of heroin use. *Journal of Drug Issues*, 36(2), 417–440. <https://doi.org/10.1177/002204260603600209>.
- Kelly, M., Johnson, T.P., & Hughes, T.L. (2015). Using respondent driven sampling to recruit sexual minority women. *Survey Practice*, 8(1), 1–10. <https://doi.org/10.29115/sp-2015-0004>.
- Killworth, P.D., McCarty, C., Bernard, H.R., Shelley, G.A., & Johnsen, E.C. (1998). Estimation of seroprevalence, rape, and homelessness in the United States using a social network approach. *Evaluation Review*, 22(2), 289–307. <https://doi.org/10.1177/0193841x9802200205>.
- Killworth, P.D., McCarty, C., Johnsen, E.C., Bernard, H.R., & Shelley, G.A. (2006). Investigating the variation of personal network size under unknown error conditions. *Sociological Methods & Research*, 35(1), 84–112. <https://doi.org/10.1177/0049124106289160>.
- Kim, U., Triandis, H.C., Kâğıtçıbaşı, Ç., Choi, S.-C., & Yoon, G. (1994). *Individualism and collectivism: theory, method, and applications*. Thousand Oaks: SAGE.
- Kim, Y., Sohn, D., & Choi, S.M. (2011). Cultural difference in motivations for using social network sites: a comparative study of American and Korean college students. *Computers in Human Behavior*, 27(1), 365–372. <https://doi.org/10.1016/J.CHB.2010.08.015>.
- Korean Statistical Information Service, & Statistics Korea (2017). Population counts by last name. <https://kosis>.

- kr/statHtml/statHtml.do?orgId=101&tblId=DT_1IN15SB&conn_path=I2. Accessed 16 Feb 2021.
- Lachowsky, N.J., Sorge, J.T., Raymond, H.F., Zishan, C., Sereda, P., Rich, A., Roth, E.A., Hogg, R.S., & Moore, D.M. (2016). Does size really matter? A sensitivity analysis of number of seeds in a respondent-driven sampling study of gay, bisexual and other men who have sex with men in Vancouver, Canada. *BMC Medical Research Methodology*. <https://doi.org/10.1186/S12874-016-0258-4>.
- Latkin, C., Mandell, W., Oziemkowska, M., Celentano, D., Vlahov, D., Ensminger, M., & Knowlton, A. (1995). Using social network analysis to study patterns of drug use among urban drug users at high risk for HIV/AIDS. *Drug and Alcohol Dependence*, 38(1), 1–9. [https://doi.org/10.1016/0376-8716\(94\)01082-V](https://doi.org/10.1016/0376-8716(94)01082-V).
- Lee, S. (2009). Understanding respondent driven sampling from a total survey error perspective. *Survey Practice*. <https://doi.org/10.29115/SP-2009-0029>.
- Lee, S. (2020). *Health and life study of Koreans, United States, 2016–2018*. <https://doi.org/10.3886/ICPSR37635.v1>.
- Lee, S., & Roddy, J.K. (2021). *Project positive attitudes towards health, Michigan, 2017*. <https://doi.org/10.3886/ICPSR37957.v1>.
- Lee, S., & Schwarz, N. (2014). Question context and priming meaning of health: effect on differences in self-rated health between hispanics and non-hispanic whites. *American Journal of Public Health*, 104(1), 179–185.
- Lee, S., Suzer-Gurtekin, T., Wagner, J., & Valliant, R. (2017). Total survey error and respondent driven sampling: focus on nonresponse and measurement errors in the recruitment process and the network size reports and implications for inferences. *Journal of Official Statistics*. <https://doi.org/10.1515/jos-2017-0017>.
- Lee, S., Ong, A.R., & Elliott, M. (2020). Exploring mechanisms of recruitment and recruitment cooperation in respondent driven sampling. *Journal of Official Statistics*, 36(2), 339–360. <https://doi.org/10.2478/JOS-2020-0018>.
- Lee, S., Ong, A.R., Chen, C., & Elliott, M.R. (2021). Respondent driven sampling for immigrant populations: a health survey of foreign-born Korean Americans. *Journal of Immigrant & Minority Health*, 23(4), 784–792. <https://doi.org/10.1007/s10903-020-01077-4>.
- Lubbers, M.J., Molina, J.L., & Valenzuela-García, H. (2019). When networks speak volumes: variation in the size of broader acquaintanceship networks. *Social Networks*, 56, 55–69. <https://doi.org/10.1016/j.socnet.2018.08.004>.
- Maltiel, R., Raftery, A.E., McCormick, T.H., & Baraff, A.J. (2015). Estimating population size using the network scale up method. *Annals of Applied Statistics*, 9(3), 1247–1277. <https://doi.org/10.1214/15-AOAS827>.
- Marsden, P.V. (1990). Network data and measurement. *Annual Review of Sociology*, 16(1), 435–463. <https://doi.org/10.1146/annurev.so.16.080190.002251>.
- McCormick, T.H., Salganick, M.J., & Zheng, T. (2010). How many people do you know?: efficiently estimating personal network size. *Journal of the American Statistical Association*, 105(489), 59. <https://doi.org/10.1198/JASA.2009.AP08518>.
- Meade, A.W., & Craig, S.B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17(3), 437–455. <https://doi.org/10.1037/a0028085>.
- Mills, H.L., Johnson, S., Hickman, M., Jones, N.S., & Colijn, C. (2014). Errors in reported degrees and respondent driven sampling: implications for bias. *Drug and Alcohol Dependence*, 142, 120–126. <https://doi.org/10.1016/J.DRUGALCDEP.2014.06.015>.
- Mollica, K.A., Gray, B., & Treviño, L.K. (2003). Racial Homophily and its persistence in newcomers' social networks. *Organization Science*, 14(2), 123–136. <https://doi.org/10.1287/orsc.14.2.123.14994>.
- Neaigus, A., Friedman, S.R., Goldstein, M., Ildefonso, G., Curtis, R., & Jose, B. (1995). Using dyadic data for a network analysis of HIV infection and risk behaviors among injecting drug users. *NIDA Research Monograph*, 151, 20–37.
- Ott, M.Q., Gile, K.J., Harrison, M.T., Johnston, L.G., & Hogan, J.W. (2019). Reduced bias for respondent-driven sampling: accounting for non-uniform edge sampling probabilities in people who inject drugs in Mauritius. *Journal of the Royal Statistical Society. Series C: Applied Statistics*. <https://doi.org/10.1111/rssc.12353>.
- Qiu, L., Lin, H., & Leung, A.K.Y. (2012). Cultural differences and switching of in-group sharing behavior between an American (Facebook) and a Chinese (Renren) social networking site. *Journal of Cross-Cultural Psychology*, 44(1), 106–121. <https://doi.org/10.1177/0022022111434597>.
- Ramirez-Valles, J. (2013). *Latino MSM community involvement: HIV protective effects*. ICPSR34385-V1. <https://doi.org/10.3886/ICPSR34385.v1>.
- Rocha, L.E.C., Thorson, A.E., Lambiotte, R., & Liljeros, F. (2017). Respondent-driven sampling bias induced by community structure and response rates in social networks. *Journal of the Royal Statistical Society*.

- ety: *Series A (Statistics in Society)*, 180(1), 99–118. <https://doi.org/10.1111/rssa.12180>.
- Ross, M., Xun, W. Q. E., & Wilson, A. E. (2002). Language and the bicultural self. *Personality and Social Psychology Bulletin*, 28(8), 1040–1050. <https://doi.org/10.1177/01461672022811003>.
- Rudolph, A. E., Fuller, C. M., & Latkin, C. (2013). The importance of measuring and accounting for potential biases in respondent-driven samples. *AIDS and Behavior*, 17(6), 2244–2252. <https://doi.org/10.1007/s10461-013-0451-y>.
- Salganik, M. J. (2006). Variance estimation, design effects, and sample size calculations for respondent-driven sampling. *Journal of Urban Health*, 83(S1), 98–112. <https://doi.org/10.1007/s11524-006-9106-x>.
- Salganik, M. J., & Heckathorn, D. D. (2004). 5. Sampling and estimation in hidden populations using respondent-driven sampling. *Sociological Methodology*, 34(1), 193–240. <https://doi.org/10.1111/j.0081-1750.2004.00152.x>.
- Stork, D., & Richards, W. D. (1992). Nonrespondents in communication network studies. *Group & Organization Management*, 17(2), 193–209. <https://doi.org/10.1177/1059601192172006>.
- Sudman, S. (1985). Experiments in the measurement of the size of social networks. *Social Networks*, 7(2), 127–151. [https://doi.org/10.1016/0378-8733\(85\)90002-4](https://doi.org/10.1016/0378-8733(85)90002-4).
- Sudman, S. (1988). Experiments in measuring neighbor and relative social networks. *Social Networks*, 10(1), 93–108. [https://doi.org/10.1016/0378-8733\(88\)90012-3](https://doi.org/10.1016/0378-8733(88)90012-3).
- Suh, T., Mandell, W., Latkin, C., & Kim, J. (1997). Social network characteristics and injecting HIV-risk behaviors among street injection drug users. *Drug and Alcohol Dependence*, 47(2), 137–143. [https://doi.org/10.1016/S0376-8716\(97\)00082-3](https://doi.org/10.1016/S0376-8716(97)00082-3).
- Triandis, H. C., Bontempo, R., Villareal, M. J., Asai, M., & Lucca, N. (1988). Individualism and collectivism: cross-cultural perspectives on self-ingroup relationships. *Journal of Personality and Social Psychology*, 54(2), 323–338. <https://doi.org/10.1037/0022-3514.54.2.323>.
- Tzioumis, K. (2018). Demographic aspects of first names. *Scientific Data*, 5(1), 1–9. <https://doi.org/10.1038/sdata.2018.25>.
- Ver Hoef, J. M. (2012). Who invented the delta method? *The American Statistician*, 66(2), 124–127. <https://doi.org/10.1080/00031305.2012.687494>.
- Volz, E., & Heckathorn, D. D. (2008). Probability based estimation theory for respondent driven sampling. *Journal of Official Statistics*, 24(1), 79–97.
- Wendel, T., Curtis, R., Khan, B., & Dombrowski, K. (2014). *Dynamics of retail methamphetamine markets in New York City, 2007–2009*. <https://doi.org/10.3886/ICPSR29821.v1>.
- Young, A. M., Abby, E. R., Quillen, D., & Havens, J. R. (2014). Spatial, temporal and relational patterns in respondent-driven sampling: evidence from a social network study of rural drug users. *Journal of Epidemiology and Community Health*, 68(8), 792–798. <https://doi.org/10.1136/jech-2014-203935>.