

Respondent-driven sampling (RDS) method: Introduction and its potential use for social psychological research

Andrian Liem* & Brian J. Hall

Department of Psychology, University of Macau, Macau.

Abstract

Contrary to other non-probability sampling methods in which researchers actively recruit potential participants, respondent-driven sampling (RDS) relies on connection and trust within social networks to access hidden or hard-to-reach populations through a peer-to-peer recruitment process. These interdependency relations aligned with Indonesian communal culture. Personal network size calculation in the RDS method makes this innovative sampling method approximate random sampling methods, where findings from the sample could be generalized to the target population. Considering its superiority, the RDS method could be applied in social psychology research in Indonesia to explore current sensitive social issues among hidden or hard-to-reach Indonesian sub-populations, for instance, radical religious groups. The current article aimed to concisely describe the RDS method; discuss ethical considerations, strengths, and weaknesses of the RDS method; and outline the potential use of the RDS method in improving the contribution of social psychology research in Indonesia, for example, by advancing strategies for social intervention programs. It is followed by a brief step-by-step process to conduct a study using the RDS method.

Keywords: social network; sampling method; minority; hidden population; participatory research

Abstrak

Berbeda dengan metode sampling nonprobabilitas lainnya yang mana peneliti secara aktif merekrut calon partisipan, metode *respondent-driven sampling* (RDS) bergantung pada hubungan dan kepercayaan dalam jejaring sosial untuk mengakses populasi tersembunyi atau yang sulit dijangkau melalui proses perekrutan teman sebaya. Ketergantungan dalam relasi tersebut sejalan dengan budaya komunal di Indonesia. Perhitungan ukuran jejaring sosial personal pada metode RDS membuat metode inovatif ini mendekati metode sampling acak di mana hasil temuan pada sampel dapat diproyeksikan kepada populasi target. Dengan mempertimbangkan keunggulan RDS, metode ini dapat diaplikasikan pada penelitian psikologi sosial di Indonesia untuk mengeksplorasi isu-isu sosial kontemporer yang sensitif pada subpopulasi yang tersembunyi atau sulit dijangkau, misalnya, kelompok radikal berbasis agama. Artikel ini bertujuan untuk menjelaskan secara singkat tentang metode RDS; diskusi mengenai pertimbangan etik, serta kekuatan dan kelemahan metode RDS; dan ulasan singkat mengenai potensi penggunaan metode RDS dalam meningkatkan kontribusi riset psikologi sosial di Indonesia, misalnya, dengan mempertajam strategi intervensi sosial. Artikel dilanjutkan dengan pemaparan singkat tentang langkah demi langkah proses dalam studi yang menggunakan metode RDS.

Kata kunci: jejaring sosial; metode sampling; minoritas; populasi tersembunyi; penelitian partisipatoris

Introduction

Respondent-driven sampling (RDS) is a non-probability sampling method, which is relatively new for sampling hidden or hard-to-reach populations (Avery & Rotondi, 2020; WHO, 2013). This sampling method was developed in the late 1990s to assess the prevalence of human immunodeficiency virus (HIV) in the United States among hidden populations, for example, injection drug users (IDUs) (Heckathorn, 1997;

Heckathorn & Cameron, 2017; Johnston, et al., 2016). RDS is often used when probability or random sampling approaches are infeasible, i.e. it is impossible to generate a sampling frame (Avery & Rotondi, 2020; Gile, Beaudry, Handcock, & Ott, 2018). The concept of the RDS method is displayed in Figure 1 (an illustrative study on HIV prevalence among IDUs) where six seeds (non-randomly selected members of the study population to complete data collection and recruit their peers) recruited their peers.

The core of RDS peer-to-peer recruitment relies on social networks that connect members of the population (WHO, 2013). A social network is a connection between people (or groups of people) and others through a particular type of interdependency (*e.g.* kinship, friendship, profession, etc.) (Ott, Gile, Harrison, Johnston, & Hogan, 2017). This interdependency is reflected by the trust through the peer-to-peer recruitment process, hence potential participants are expected to be interested in participating in the study because they are asked by someone they know.

Although a similar social network approach is used in snowball sampling method, RDS is superior due to several fundamental differences and involving complex calculations of social

networks borrowed from sociology, physics, statistics, and mathematics disciplines (Bengtsson, et al., 2012; Gile, et al., 2018; Johnston & Sabin, 2010). Also, six basic assumptions must be met for a successful study with RDS method (Johnston, et al., 2016; WHO, 2013): 1) participants must know one and another as members of the study population; 2) the sample population must be large and dense; 3) all members of the population have multiple chances to participate; 4) participants can provide accurate personal network sizes; 5) peers are recruited randomly from participant's network; and 6) each participant can recruit at least one peer. A summary of the differences between snowball and RDS methods are presented in Table 1 below.

Table 1
A Comparison Between Snowball and RDS Methods

Aspect	Snowball Method	RDS Method
Social network	Relationships between participants are not always reciprocal.	Relationships must be reciprocal.
Bias issue	Participants may refer to an unlimited number of peers. This means those with larger network size may recruit more peers with highly similar traits and vice versa.	Regardless of the size of their social networks, the participant is only able to recruit a limited number of peers, which reduces recruitment bias (<i>i.e.</i> homophily [a tendency for participants to recruit others in their social network who are similar/different to themselves]).
Social network properties	Network traits cannot be measured and information about the size of the social network is unclear.	Network traits can be analysed through a coded coupon and the size of the social network is used in the weighted analysis.
Recruitment method	The participant only refers to their peer's name/information but the researcher must find them, which may be a risk of being rejected due to lack of trust in the researcher.	Participant recruits their peer directly, which may improve the probability of participation through social influence; and information about the peer remains protected if they do not wish to participate.
Incentive	The participant is only rewarded for their participation in the survey/interview.	The participant is rewarded for their participation in the survey/interview (first incentive) as well as for recruiting their peers (second incentive).
Generalization	Findings are not generalizable because it is limited to the sample.	Findings could be extended to the population by calculating the probability of selection within the network.

Note: The differences are summarised from multiple sources (Bengtsson et al., 2012; Gile et al., 2018; Johnston et al., 2016; Johnston & Sabin, 2010; WHO, 2013)

RDS method has been successfully implemented in low- and middle-income as well as high-income nations because it is the least biased among non-probability methods to produce population estimates (Avery & Rotondi, 2020; Carrillo, Rivera, & Braunstein, 2019; Gile et

al., 2018). Most of these were epidemiological studies related to HIV prevalence and prevention, including in Indonesia (Iskandar, et al., 2010; Iskandar, et al., 2012; Morineau, et al., 2012). RDS method has also been used to recruit hidden populations who are less stigmatized than

participants in epidemiological studies. For example, a study on traveling patterns of migrant workers from Myanmar and Cambodia on the Thailand-Cambodia border (Khamsiriwatchara, et al., 2011) and intention to use electronic mental health among Filipino migrant workers in Macao (Hall, Shi, Garabiles, & Chan, 2018). In addition, the RDS method was used for ethnographic-network studies among the indigenous community in Canada, in which the representativeness from the studies was relatively similar to the national census data (Dombrowski, Khan, Moses, Channell, & Misshula, 2013).

Therefore, this article aims to introduce the RDS method to the social psychology community and encourage its use in social psychology research in Indonesia. As explained before, the core of the RDS method is the relational interdependency in the community (Ott, et al., 2017), which aligns with the Indonesian communal culture where loyalty to a group is expected from the members to maintain social harmony (Milla, Hudiyan, Cahyono, & Muluk, 2020). Moreover, this cultural aspect is translated into Indonesian peoples' behaviour of tendency to help others, especially members of their community (Slikkerveer, 2019). Hence, by being recruited by their peers, potential participants may feel an 'obligation' to help them, which is another characteristic of the RDS method that employs the power of social pressure (Heckathorn, 1997; Mosher, Moorthi, Li, & Weeks, 2015).

The following sections will cover the ethical considerations, strengths, and weaknesses of RDS along with its potential use in social psychology research. Following that, a concise and practical step-by-step process to conduct a study with the RDS method will be discussed, from formative assessment to data analysis, which is mainly summarised from an RDS module published by the WHO (2013). However, as a brief overview of RDS methodology, this article will not discuss statistical analyses related to the RDS method. Readers who are interested to know more about RDS methodology can read other publications, including those about the history and development (Heckathorn, 1997, 2007; Heckathorn & Cameron, 2017), statistical analysis (Aronow & Crawford, 2015; Gile, et al., 2018; Khabbazian, Hanlon, Russek, & Rohe, 2017; Lu, Malmros, Liljeros, & Britton, 2013; Salganik & Heckathorn, 2004), evaluation of RDS method implementation (Carrillo, et al., 2019; Gilbert, Haines, Baquero, & Parker, 2018; Hipp, Kohler, & Leumann, 2019; Lattorf, 2018), web-based RDS (WebRDS) as an alternative approach (Bengtsson, et al., 2012; Salamońska & Czeranowska, 2018; Strömdahl, Lu, Bengtsson, Liljeros, & Thorson, 2015), and systematic

reviews of RDS method implementation (Avery & Rotondi, 2020; Johnston, et al., 2016; White, et al., 2015).

Ethical considerations, strength, and weakness of the RDS method

Ethical considerations

RDS method is usually chosen for studies with vulnerable populations that are hard-to-reach or stigmatized. Therefore, it is important to carefully assess potential ethical and safety issues that may affect not only members of the populations but also the researchers and their team members. Below are some ethics recommendations summarised from previous studies (Hipp, et al., 2019; Mosher, et al., 2015; WHO, 2013).

Second incentive

The second incentive is a unique feature of the RDS method to motivate participants to recruit their peers. However, it could be an ethical issue if the total value of the second incentive is excessively high, in which participants might try to earn an income as a recruiter. Hence, the coupons distributed (*i.e.* 3-5 coupons) to each participant as well as the amount that they receive for each successful recruitment should be limited. There is a risk that participant (recruiter) might coerce their peers to be a participant (recruit) if the amount of the second incentive is too high. Additionally, if their peers refused to join or took the coupon but did not come to the study site, recruiters might react harmfully or damage the relationships in the population. Another point of view, excessively high incentives might put the safety of the researcher/ research assistant at risk when potential recruits come to study site but they are not eligible or have not completed the data collection so the second incentive cannot be granted to their recruiters. These recruiters may not be understanding of such reasons and become angry.

Informed consent

Informed consent must be obtained before any data is collected from participants and the researcher must ensure that the participant's involvement is voluntary. Potential participants who come to the study site have the right to know about the study and the potential risks related to this. Their recruiters might not give detailed information or they may have been given inaccurate information. Certain issues need to be emphasized during the peer recruitment process, including potential relationship conflicts, loss of trust, or other harms. Additionally, prior to

taking their first incentive for participating in data collection, participants can be trained on how to inform recruits about the study correctly and how to recruit their peers ethically and appropriately, through role-play or watching a pre-recorded video example.

Confidentiality

The object of study with the RDS method or participants' behaviour may be categorized as something illegal or stigmatized (*e.g.* use of drugs among IDUs, commercial sex among sex workers, use of condoms among non-heterosexual persons, and mental illness diagnosis). Therefore, researchers must ensure data confidentiality and protect participants' privacy and safety. Researchers may consult a colleague with a law background if needed. It is recommended for researchers not to collect any identifiable information about participants (*e.g.* name, address, and ID number). However, if the population is less stigmatized (*i.e.* domestic

workers) and researchers deem that it is safe to collect participants' personal information for research purposes, then the collected personal information must be stored separately from their responses in the study and only the principal investigator can access this information. Confidentiality also applies in the recruitment process by not disclosing information to the participants (the recruiters) about which peers (whom they recruited/recruits) participate or do not participate. Therefore, it is suggested to have at least three coupons for recruitment so that the recruiters will not easily know which of their recruits choose not to participate in the study.

Strengths and weaknesses of the RDS method

Every sampling method has its strengths and weaknesses, including RDS. The summary of the strengths and weaknesses of the RDS method in various aspects is presented in Table 4, as seen below.

Table 2
Strengths and Weakness of RDS Method

Aspect	Strength	Weakness
Sampling	<ul style="list-style-type: none"> It can obtain a large sample in a relatively short period. Access to the hidden population. Contrary to other non-random sampling methods, selection bias in the RDS method can be minimized and calculated. 	<ul style="list-style-type: none"> Maybe time-consuming or slower than expected if seeds and participants are not productive or successful in recruiting their peers. Not suitable for a population with a small number of members or a big population but without a strong connection between members. Potential challenges in validating membership of potential participants, for example, people may pretend to be population members to get study incentives.
Logistic	<ul style="list-style-type: none"> More practical and time-saving for researchers because they do not need to approach potential participants themselves. 	<ul style="list-style-type: none"> Require research assistants (<i>e.g.</i> at the reception, to collect data, to manage coupons and give incentives), which may need additional budget and resources.
Analysis	<ul style="list-style-type: none"> If the basic assumptions of data collection through the RDS method are met, findings from the sample can be extended to population-level like studies with random sampling methods. 	<ul style="list-style-type: none"> Findings may not be generalizable if data is collected from bottleneck participants. The analysis is (mathematically) complicated and needs special software (however it is available for free).
Costs		<ul style="list-style-type: none"> Need a relatively large budget for the participants' incentives. Need a budget for research assistants' fees and logistical fees (<i>i.e.</i> rent a space for study site that is perceived safe by the participants).
Ethics	<ul style="list-style-type: none"> The population's confidentiality is maintained because their members (insider) are the ones recruiting. 	<ul style="list-style-type: none"> The risk for disrupting social networks in the population if the second incentive is excessively high.

The potential use of the RDS method for social psychology research

Since the RDS method was introduced in 1997, this method has been extensively used and reported in more than 500 articles that have been published until 2013 (Johnston, et al., 2016). Although RDS was originally used for epidemiological and public health studies on HIV, the use of this sampling method has been extended to social sciences with participants whose social networks are also difficult for an outsider to access such as people who are homeless, aging artists, and migrant workers (Hall, et al., 2018; Heckathorn & Cameron, 2017; Khamsiriwatchara, et al., 2011). Additionally, the objects of studies with the RDS method have been varied, for example, about violence experiences, mental health, health care use, and life satisfaction (Avery & Rotondi, 2020).

Reflecting on this advancement, RDS as an innovative and relatively new method could be used for social psychology research. RDS method that relies on social networks in the population can improve the applicability and contribution of social psychology, which is perceived as one of the challenges for the social psychology community in Indonesia (Takwin, 2019). This sampling method can be applied for exploring social issues with hard-to-reach participants, for example, members of radical religious groups, political *buzzers* on social media, survivors of sexual abuse in universities, providers of (illegal) abortion services, collectors of (illegal) antique properties. Additionally, findings of these social psychology studies with the RDS method could be used for developing better policy or regulation, reducing the risk of social conflict, or improving the quality of life of citizens.

For example, deradicalization programs in Indonesia may be expanded by using social psychology studies with the RDS method. Moreover, the communal culture in Indonesia strengthened the social relations and cohesion among radical religious group members (Milla, et al., 2020), which an adequate reason for using RDS. This deradicalization effort is important especially because of the rising number of terrorists associated with radical religious groups in Indonesia (Kurniawan, Mujahid, & Usman, 2018). However, such deradicalization programs organized by the National Agency for Combating Terrorism (*Badan Nasional Penanggulangan Terorisme*) are heavily focused on imprisoned terrorists. This strategy can be extended by targeting members of radical religious groups with an intention of homicide attack (*i.e.* homicide bombing) and decreasing this intention.

The first step of the study would be to recruit people who perpetrated acts of terror and who were imprisoned and completed the deradicalization programs. Then, it would be followed by asking the seeds to recruit their peers, particularly the operational leaders, from their former radical religious group. Previous research on Indonesian terrorist groups found that these operational leaders played a central role in their networks (Milla, et al., 2020). The study use RDS could explore the intention of homicide attacks among their group members (*i.e.* "Would you do a homicide bombing if your mentor/teacher asked you to?"). Findings from this study could be used to sharpen deradicalization program by focusing on which area requires more attention (*i.e.* sub-urban areas), which age group needs to be prioritized (*i.e.* among people from 25 to 35 years old), and which socio-economic backgrounds need to be highlighted (*i.e.* from middle SES and having a bachelor's degree in hard sciences).

In sum, the RDS method can be implemented in social psychology studies to advance strategies for social intervention such as a deradicalization program or to prevent terrorist crimes such as homicide attacks. However, as described in the ethical consideration section and for the RDS study to be successful, participants' privacy and safety must be protected because their behaviours could be categorized as a crime, similar to study among IDUs.

Conducting RDS study: Step-by-step

This section will concisely describe the primary steps to conduct research using the RDS method adapted from the WHO (2013), including 1) formative assessment; 2) preparation; 3) data collection; and 4) data analysis.

Formative assessment

The first step is to gather essential information to consider whether the RDS method is feasible and to serve as a foundation in designing the implementation of RDS. Formative assessment can be conducted through multiple methods that depend on what information is needed and the availability of resources. Stakeholders or key informants (*e.g.* local community leaders and local non-governmental organizations) should also be involved in this preliminary stage, or as members of the advisory group (Hipp, et al., 2019; Lattof, 2018). In addition, potential seeds from the population may also be involved in this stage, which is why the RDS method is also known as participant-driven recruitment (PDR) and community-based

participatory research (CBPR) (Tiffany, 2006). A summary of formative assessment and

preparation to implement Respondent-Driven Sampling (RDS) is presented in Table 3.

Table 3
Information Needed and Methods to be Used in Formative Assessment

Information	Methods			
	Literature review ¹	In-depth interviews	Focus group discussions	Observation
Past or current conditions in general to identify the knowledge gap	X			
Potential seeds and key informants	X		X	X
Any challenges that may rise		X	X	
Options for data collection site and potential geographical/logistical difficulties			X	X
Social networks in the population, including its size, complexity, and sub-populations		X	X	X
Recommendation for data collection location, incentive levels/types, and other logistical aspects (<i>i.e.</i> coupon design)	X	X	X	
Appropriateness and acceptance of the RDS method and research in general		X	X	
Culturally-sensitive instruments	X		X	
Social space in the community for gathering, including the time they gather		X		X
Daily interactions between members in the population, including their languages and behaviors				X

Note: ¹ = *e.g.* policy / regulation documents, news, previous reports in academic journals or from government / non-governmental archives.

Preparation for RDS implementation

Sample size

Although RDS is part of non-probability sampling methods, the sample size needs to be calculated *a-priori* to extend the estimated findings to the target population. There are two formulas used to calculate the required sample size for studies using the RDS method. The first is to estimate the prevalence of certain characteristics based on several pre-specified conditions. Second is to test the changes of an estimate over time (*i.e.* whether condom use at the last sexual intercourse has increased from the survey at time 1 to time 2

among IDUs, age 18 years or older, and living in city Y). Only the first formula will be discussed in this article because it is more relatable for cross-sectional social psychology research whereas the second formula is mostly used for longitudinal-experimental research to detect a proportion change. The formula is presented below.

As an example, a researcher wants to estimate the prevalence of depression among non-partnered homosexual men in the past year, over the age of 20 and living in Jakarta city. Previous reports on mental health suggested that the estimated proportion of working-age-males with depression in Jakarta is up to 35% ($P = 0.35$).

$$n = D \frac{Z_{1-\alpha}^2 * P(1 - P)}{d^2}$$

n = Sample size required

D = Design effect, conventionally set at 2 (moderate effect)

$Z_{1-\alpha}$ = The z score for the desired confidence level, usually set at 1.96 (95% confidence)

P = Expected proportion of people with the condition / characteristic

d = Precision, conventionally set at 5%

Therefore, the final sample size needed is 699.17 or rounded up to 700 non-partnered homosexual men as calculated below:

The D (the estimated ratio of non-random sampling variance over simple random sampling variance) is conventionally set at 2, which means to collect a sample double than what simple random sampling would need. However, a design effect of 3 or 4 is recommended if the resources (e.g. funding and research assistant) and time permits because it means more participants will be recruited (Johnston, Chen, Silva-Santisteban, & Raymond, 2013). If there is no previous information about P in the target population, a researcher may use P from

other places with similar geographic or population characteristics. For instance, the national prevalence of depression might be used if the mental health national survey did not breakdown its results into a province, age, or sex categories. However, if there is absolutely no previous information regarding prevalence, P can be set at 0.5 (50%) (Khamsiriwatchara, et al., 2011), which results in 769 participants as the minimum sample size. It is important to remember that this formula calculation is based on the need for precise prevalence estimation in a target population, which does not include the statistical power.

$$n = 2 \frac{1.96^2 * 0.35(1 - 0.35)}{0.05^2} = 699.17$$

Coupon design

One of the most notable aspects that differentiate RDS from the snowball method is the use of coupons, which limits the recruitment of peers and allows for tracking the recruitment process. Furthermore, the use of coupons is essential for data analysis in studies with the RDS method. The basic elements printed on the coupon should cover: 1) unique RDS identification code; 2) study site location; 3) contact number/email; 4) days and hours of operation; 5) expiry date. Coupons should be printed on thick paper so that it is not easily damaged and should be designed carefully to avoid imitation (Hipp, et al., 2019). The size of the coupon may follow business card size or the same size as local banknotes divided into two parts: a referral part (to be given to peer/potential participant) and a payment part (to be kept by the participant/recruiter to collect their second incentive). Below is an example of the RDS coupon design (Figure 2). In addition, each coupon will contain a unique code, which can be a series of numbers or combinations of numbers and letters, to track the recruitment process (*i.e.* who recruited who) without disclosing participant's

personal information. For example, from Figure 1, code A-03-02 means the individual who brings this coupon (HIV-, female, Ethnicity B) is the second person (02) recruited by Participant 3 (03) of Seed A.

Incentives

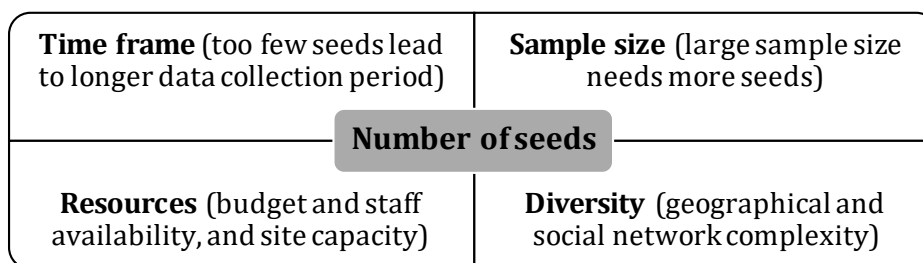
Participants in studies with the RDS method will be rewarded for their time to participate in the survey/interview (first incentive) and effort for recruiting their peers (second incentive). The most common type of incentive is money. However, some alternatives that can be considered are grocery vouchers, credit to SIMcard, and gifts (*eg.* umbrella or t-shirts). Additionally, the option to donate incentives to the community/organization of participant's choice can be added to attract participants from middle and high socioeconomic status (SES) who are not interested in financial incentives (Bengtsson, et al., 2012). Researchers should also consider covering transport expenses, in addition, to reward for participating if the study site is located far from where the participant resides (Lattof, 2018).

There is no rule to decide how many seeds are needed—this can range from 3 to 15 (or more). As a recommendation, usually up to eight seeds are needed for a sample size of 400. However, there are some considerations in deciding the number of seeds, including 1) seeds that are too few may increase the time needed to reach sample size; 2) seeds that are too many

may result in short waves, which becomes a risk of not reaching equilibrium (a condition where the final wave is not biased by the purposive selection of seeds); 3) the larger the sample size, the more seeds that may be required. Some important factors that may affect in deciding seed numbers are summarized in Figure 3.

Figure 3

Factors to be Considered for Seed Numbers



Study site

Researchers need to consider some factors when deciding the study site, including (WHO, 2013): 1) accessibility (easy to be accessed); 2) confidentiality (participant's privacy can be maintained); 3) safety (both participants and researchers feel safe); and 4) comfortability (*e.g.* clean, space is adequate for research activity, and have a waiting area for participants). A study with the RDS method is commonly conducted at a fixed study site such as inside university. However, researchers need to consider an alternative study site if the university location is difficult to be accessed by participants and data collection in the university may also create a social barrier (Salamońska & Czeranowska, 2018). Some alternatives for study sites are (WHO, 2013): public facility (*i.e.* public health centre), non-governmental organization office, commercial building rented house, and mobile sites with van or minibus. This selection of the most appropriate study site should be discussed with stakeholders during the formative assessment stage (Lattof, 2018).

Data collection

The data collection process in the RDS method is summarized in Table 3. In general, the RDS data collection process is the same as other survey studies. However, some initial steps need to be performed concerning assessing the participant's eligibility and membership in the study population. Also, as the uniqueness of the

RDS method, seeds, and participants who completed the data collection will be asked to recruit their peers from the same population. This peer-recruitment process is managed and monitored through the coded coupon as explained in the previous section. With a limited coupon given, participants are expected to approach and recruit their peers who are eligible and would come to the study site and participate; and this peer-recruitment will be repeated in multiple-waves until the targeted sample size is achieved.

Controlling sample growth

It is difficult to predict the growth of the seed and recruitment process and how long it will take to meet the required sample size in the RDS method. For example, a survey study among IDUs with an initial eight seeds took six weeks to reach a sample size of 400, while another study starting with 20 seeds took eight weeks to reach the same sample size (WHO, 2013). Therefore, researchers need to closely monitor recruitment pace and the number of participants who participate in the study. If necessary, the researcher may interfere with the recruitment process. For example, if the sample exponentially grows in the first three waves, the researcher may slow it down by gradually reducing the recruitment coupon (*i.e.* from three to two coupons for each participant in fourth and fifth waves, then to only one coupon for participants in the sixth wave). Conversely, if recruitment is slower than expected, the researcher may recruit some new seeds.

Table 4
Steps of Data Collection

Step	Description
Validating coupon and eligibility	<ul style="list-style-type: none"> • Checking coupons (unique code, originality, and expiry date) • Confirming eligibility (<i>i.e.</i> age) • Informing about the research and asking for consent
Verifying membership in the survey population	<p>This step ensures that the person does not disguise themselves as population members. For example, if the study is among people with a homosexual orientation, the researcher can ask the person about:</p> <ul style="list-style-type: none"> • The frequency of sexual activity with a same-sex person • The role in sexual activity with a same-sex person
Recording social network size	<p>Asking participants about the total members of the surveyed population that they know personally. "Know" in this context means that participant and their peer have reciprocal relationships (participant recognizes their peer, and their peer also recognizes the participant). Asking this question is to prevent the participant from recruiting a random person outside the targeted population, or someone that they do not know (<i>i.e.</i> friend of a friend).</p> <p>The question about social network size/degree described above can be combined with some other characteristics (<i>e.g.</i> how many males vs females, how many from Ethnicity A vs B vs C, and how many are under 20 years old) or specific timeframe (<i>e.g.</i> that they have met them within the last week / last month / previous three months).</p>
Inquiring about recruiters	<p>Asking about the information below to assess the relationship between participants and their recruiters:</p> <ul style="list-style-type: none"> • Participant's relationship with the recruiter (<i>e.g.</i> a close friend, relative, family member, etc.) • Degree of their closeness (<i>e.g.</i> very close, somewhat close, or not very close) • Intensity/frequency of their encounter (<i>e.g.</i> every day, once a week, once a month, etc.) • Duration of their relationship (<i>e.g.</i> less than a month, less than a year, more than a year, etc.)
Collecting data	Asking the participant to fill out questionnaires or participate in an interview
Explaining the recruitment process and giving incentive	<p>Training the seed/participant on how to recruit their peers, including:</p> <ul style="list-style-type: none"> • Through roleplay (by seed/participant) on explaining the aim and benefit of the study, and data collection process • Reminding about the expiry dates <p>Providing information about collecting the second incentive:</p> <ul style="list-style-type: none"> • Reminding to keep the payment coupon to claim the second incentive • The participant will receive the second incentive only if their peer brings the recruitment coupon, is eligible, and completes the data collection • Informing that the second incentive can only be collected during a specific time/day (to avoid the crowd in the study site due to overlap with participants who are currently participating in data collection)

Ending RDS recruitment

Once the sample size is reached, the study site can be closed. However, the researcher needs to check whether all participants (recruiters) have claimed their second incentives. Some participants may not have collected their second incentives, especially in the second last wave, which could be because of geographical barriers to reach the study site. Also, some potential participants who are recruited in the second last wave may come on the last day after the coupon's expiry date.

Therefore, the study site should remain open for at least a week after the coupon expires, after the final wave has ended because the sample size has been achieved. A clear and noticeable announcement containing contact details (*i.e.* "The study has ended. Please contact [name] [phone/email address] for further information.") should be placed at the study site (*i.e.* entrance) when it is closed so participants/potential participants are still able to contact the researcher.

Data analysis

This article will only describe the basic concept of RDS analysis (homophily and equilibrium), which will assist readers to understand what essential data are needed in the data collection to conduct a rigorous RDS analysis. A detailed explanation of statistical calculation and analysis are beyond the scope of this article. In addition, there are software programs to analyze RDS data available to download for free. For instance, the RDS Analysis Tool (Volz, et al., 2012) that is complemented with a user manual (Spiller, Cameron, & Heckathorn, 2012).

RDS method is superior to other non-random sampling methods because of the basic assumption that participants will recruit their peers randomly from their social network, which is proximate to the random sampling methods principle. Therefore, for proper RDS analysis, the researcher must collect three essential information about the participants' social network, including 1) social network size of each participant (see Table 3); 2) participant's unique identification number (from the coupon code); and 3) recruiter's unique identification number (from the coupon code). The size of the participant's social network is used to apply statistical weights to mitigate bias due to different network sizes. The unique identification number of participants is used to track the recruitment process (*i.e.* who recruited whom) and calculate the transition probabilities (*i.e.* the probability of a female to recruit a female versus a male).

Homophily

Some sampling biases may potentially occur in the RDS method as different recruitments may result in over-sampling or under-sampling. Over-sampling may occur due to volunteerism or when one particular group/sub-population is more cooperative than others. Conversely, under-sampling may occur as a result of masking, in which one particular group/sub-population is more hidden or intentionally isolate themselves from the researcher or peers from different sub-populations. For example, if women with depression recruit more people (especially other women) on average than men do, while men with depression are more hidden and less willing to join the study, this means women may be over-sampled while men are under-sampled. Nevertheless, diverse seeds, limited coupon numbers (quota), and double incentive aspects in the RDS method should minimize these biases.

To ensure that the collected data is not over-sampled or under-sampled, the analysis software calculates homophily or the index of clustering, which is a tendency for participants to

recruit others in their social network who are similar (or dissimilar) to themselves. Homophily score (H) ranges from -1 (completely heterophilous) to $+1$ (completely homophilous), where a higher score indicates that participant recruits more often within their group (Spiller, 2012). For example, $H = .70$ of a group of women from Ethnicity C, indicates that they are more often (70%) recruited from the same ethnicity and less often (30%) recruited randomly or from other ethnicity groups. However, it is important to note that the definition and interpretation of homophily in this article is not the definitive version as there are some inconsistencies on this aspect among studies using RDS method (Crawford, Aronow, Zeng, & Li, 2017; Gile, et al., 2018; Zeng, Li, & Crawford, 2019).

Equilibrium

An excellent data collection with the RDS method is represented by attaining equilibrium (calculated by the software), which is an indication that the final sample is not biased by the purposive selection of initial seeds. In other words, equilibrium is the state when sample variation will not occur (or in extremely small percentage) even though more waves are produced. This is also referred to as convergence and stabilization. As an example, the estimate of depression is likely to change as more participants are recruited into the sample. However, the study will reach an equilibrium when the ratio of people with depression to people without depression (or other characteristics such as sex and ethnicity) remains stable despite the addition of new participants in the study. This equilibrium state may be comparable to data saturation condition in qualitative research, where additional data from new participants do not result in any new categories/themes. Equilibrium can be achieved by having recruitment chains (waves) that are long enough, hence the bias presented from the initial selection of seeds can be reduced.

Conclusion

The basic principle of the RDS method is the utilization of reciprocal relationships between members in populations to access hidden or hard-to-reach populations through peer-to-peer recruitment. This innovative non-probability sampling method is originally used for epidemiological studies. However, the RDS method has recently been applied in social sciences as well due to its potential for generalizing findings as in random sampling methods. Due to its superiority compared to other non-random sampling methods, RDS should be considered for social psychology research in Indonesia. Moreover, the

inter-dependency and social pressure characteristics of RDS aligned with the communality culture in Indonesia. RDS method can assist to explore current sensitive social issues among hidden or hard-to-reach populations, and findings from these studies can hopefully improve the contribution of social psychology research in Indonesia, for examples, by sharpening social policies or programs, reducing the risk of social conflict, or improving quality of life of Indonesian people. The concise description of primary steps to conduct a study with RDS method in this article is only an overview to inform about the uniqueness of RDS method (e.g. coupon system and *a-priori* sample size) compared with other non-probability sampling methods, and readers who are interested to implement RDS method into their research should access the recommended references for more detail guideline.

Acknowledgments

AL is supported by a Postdoctoral Fellowship (Class A) – Macao Talent Program from University of Macau. AL thanks Rachmania P. Wardhani for the proofreading of the manuscript draft. The authors thank the editor, co-editors, and reviewer for their valuable feedback.

References

- Aronow, P. M., & Crawford, F. W. (2015). Nonparametric identification for respondent-driven sampling. *Statistics & Probability Letters*, *106*, 100-102. <https://doi.org/10.1016/j.spl.2015.07.003>
- Avery, L., & Rotondi, M. (2020). More comprehensive reporting of methods in studies using respondent driven sampling is required: A systematic review of the uptake of the STROBE-RDS guidelines. *Journal of Clinical Epidemiology*, *117*, 68-77. <https://doi.org/10.1016/j.jclinepi.2019.09.024>
- Bengtsson, L., Lu, X., Nguyen, Q. C., Camitz, M., Hoang, N. L., Nguyen, T. A., Liljeros, F., & Thorson, A. (2012). Implementation of web-based respondent-driven sampling among men who have sex with men in Vietnam. *PLoS One*, *7*, e49417. <https://doi.org/10.1371/journal.pone.0049417>
- Carrillo, S. A., Rivera, A. V., & Braunstein, S. L. (2019). Implementing respondent-driven sampling to recruit women who exchange sex in New York City: Factors associated with recruitment and lessons learned. *AIDS and Behavior*, (online first), 1-12. <https://doi.org/10.1007/s10461-019-02485-w>
- Crawford, F. W., Aronow, P. M., Zeng, L., & Li, J. (2017). Identification of homophily and preferential recruitment in respondent-driven sampling. *American Journal of Epidemiology*, *187*, 153-160. <https://doi.org/10.1093/aje/kwx208>
- Dombrowski, K., Khan, B., Moses, J., Channell, E., & Misshula, E. (2013). Assessing respondent driven sampling for network studies in ethnographic contexts. *Advances in Anthropology*, *3*, 1-9. <https://doi.org/10.4236/aa.2013.31001>
- Gilbert, P. A., Haines, H., Baquero, B., & Parker, E. A. (2018). Respondent-driven sampling to recruit Latinos in a midwest micropolitan area: Lessons learned and recommendations for translational work. *Journal of Clinical and Translational Science*, *2*, 245-248. <https://doi.org/10.1017/cts.2018.322>
- Gile, K. J., Beaudry, I. S., Handcock, M. S., & Ott, M. Q. (2018). Methods for inference from respondent-driven sampling data. *Annual Review of Statistics and Its Application*, *5*, 65-93. <https://doi.org/10.1146/annurev-statistics-031017-100704>
- Hall, B. J., Shi, W., Garabiles, M. R., & Chan, E. W. W. (2018). Correlates of expected eMental health uptake among Filipino domestic workers in China. *Global Mental Health*, *5*, e33-46. <https://doi.org/10.1017/gmh.2018.25>
- Heckathorn, D. D. (1997). Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems*, *44*, 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*, 151-208. <https://doi.org/10.1111/j.1467-9531.2007.00188.x>
- Heckathorn, D. D., & Cameron, C. J. (2017). Network sampling: From snowball and multiplicity to respondent-driven sampling. *Annual Review of Sociology*, *43*, 101-119. <https://doi.org/10.1146/annurev-soc-060116-053556>
- Hipp, L., Kohler, U., & Leumann, S. (2019). How to implement respondent-driven sampling in practice: Insights from surveying 24-hour migrant home care workers. *Survey Methods: Insights from the Field*, 1-13. <https://doi.org/10.13094/SMIF-2019-00009>
- Iskandar, S., Basar, D., Hidayat, T., Siregar, I. M. P., Pinxten, L., van Crevel, R., Van der Ven, A.

- J.A.M., & De Jong, C. A. J. (2010). High risk behavior for HIV transmission among former injecting drug users: A survey from Indonesia. *BMC Public Health*, *10*, 472. <https://doi.org/10.1186/1471-2458-10-472>
- Iskandar, S., de Jong, C. A., Hidayat, T., Siregar, I. M., Achmad, T. H., van Crevel, R., & van der Ven, A. (2012). Successful testing and treating of HIV/AIDS in Indonesia depends on the addiction treatment modality. *Journal of Multidisciplinary Healthcare*, *5*, 329-336. <https://doi.org/10.2147/JMDH.S37625>
- Johnston, L. G., Chen, Y.-H., Silva-Santisteban, A., & Raymond, H. F. (2013). An empirical examination of respondent driven sampling design effects among HIV risk groups from studies conducted around the world. *AIDS and Behavior*, *17*, 2202-2210. <https://doi.org/10.1007/s10461-012-0394-8>
- Johnston, L. G., Hakim, A. J., Dittrich, S., Burnett, J., Kim, E., & White, R. G. (2016). A systematic review of published respondent-driven sampling surveys collecting behavioral and biologic data. *AIDS and Behavior*, *20*, 1754-1776. <https://doi.org/10.1007/s10461-016-1346-5>
- Johnston, L. G., & Sabin, K. (2010). Sampling hard-to-reach populations with respondent driven sampling. *Methodological Innovations Online*, *5*, 38-48. <https://doi.org/10.4256/mio.2010.0017>
- Khabbazian, M., Hanlon, B., Russek, Z., & Rohe, K. (2017). Novel sampling design for respondent-driven sampling. *Electron. J. Statist.*, *11*, 4769-4812. <https://doi.org/10.1214/17-EJS1358>
- Khamsiriwatchara, A., Wangroongsarb, P., Thwing, J., Eliades, J., Satimai, W., Delacollette, C., & Kaewkungwal, J. (2011). Respondent-driven sampling on the Thailand-Cambodia border. I. Can malaria cases be contained in mobile migrant workers? *Malaria Journal*, *10*, 120. <https://doi.org/10.1186/1475-2875-10-120>
- Kurniawan, W., Mujahid, D. R., & Usman, Y. A. (2018). Apakah sastra mampu menurunkan cognitive closure pada narapidana terorisme di Indonesia? *Jurnal Psikologi Sosial*, *16*, 114-124. <https://doi.org/10.7454/jps.2018.11>
- Lattof, S. R. (2018). Collecting data from migrants in Ghana: Lessons learned using respondent-driven sampling. *Demographic Research*, *38*, 1017-1058. <https://doi.org/10.4054/DemRes.2018.38.36>
- Lu, X., Malmros, J., Liljeros, F., & Britton, T. (2013). Respondent-driven sampling on directed networks. *Electron. J. Statist.*, *7*, 292-322. <https://doi.org/10.1214/13-EJS772>
- Milla, M. N., Hudiyana, J., Cahyono, W., & Muluk, H. (2020). Is the role of ideologists central in terrorist networks? A social network analysis of Indonesian terrorist groups. *Frontiers in Psychology*, *11*. <https://doi.org/10.3389/fpsyg.2020.00333>
- Morineau, G., Bollen, L. J. M., Syafitri, R. I., Nurjannah, N., Mustikawati, D. E., & Magnani, R. (2012). HIV prevalence and risk behaviours among injecting drug users in six Indonesian cities implications for future HIV prevention programs. *Harm Reduction Journal*, *9*, 37-43. <https://doi.org/10.1186/1477-7517-9-37>
- Mosher, H. I., Moorthi, G., Li, J., & Weeks, M. R. (2015). A qualitative analysis of peer recruitment pressures in respondent driven sampling: Are risks above the ethical limit? *International Journal of Drug Policy*, *26*, 832-842. <https://doi.org/10.1016/j.drugpo.2015.05.027>
- Ott, M. Q., Gile, K. J., Harrison, M. T., Johnston, L. G., & Hogan, J. W. (2017). 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)*, 1-34. <https://doi.org/10.1111/rssc.12353>
- Salamońska, J., & Czeranowska, O. (2018). How to research multiple migrants? Introducing web-based respondent-driven sampling survey. In *search of a theory of multiple migration: A quantitative and qualitative study of Polish migrants after 1989*, 1-36. Centre of Migration Research. <http://www.migracje.uw.edu.pl/wp-content/uploads/2018/09/WP110168.pdf>
- Salganik, M. J., & Heckathorn, D. D. (2004). 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>
- Schonlau, M., & Liebau, E. (2012). Respondent-driven sampling. *The Stata Journal*, *12*, 72-93. <https://doi.org/10.1177/1536867X1201200106>
- Slikkerveer, L. J. (2019). Gotong royong: An indigenous institution of communality and mutual assistance in Indonesia. In L. J. Slikkerveer, G. Baourakis, & K. Saefullah (Eds.), *Integrated community-managed development: Strategizing indigenous knowledge and institutions for poverty reduction and sustainable community development in Indonesia* (pp. 307-320).

- Springer International Publishing
https://doi.org/10.1007/978-3-030-05423-6_14
- Spiller, M. W., Cameron, C., & Heckathorn, D. D. (2012). *RDS Analysis Tool 7.1 User Manual*. Respondent Driven Sampling http://www.respondentdrivensampling.org/reports/RDSAT_7.1-Manual_2012-11-25.pdf
- Strömdahl, S., Lu, X., Bengtsson, L., Liljeros, F., & Thorson, A. (2015). Implementation of web-based respondent driven sampling among men who have sex with men in Sweden. *PLoS One*, *10*, e0138599. <https://doi.org/10.1371/journal.pone.0138599>
- Takwin, B. (2019). Pesan dari editor-in-chief: Riset psikologi sosial yang dibutuhkan Indonesia. *Jurnal Psikologi Sosial*, *17*, 59-64. <https://doi.org/10.7454/jps.2019.9>
- Tiffany, J. (2006). Respondent-driven sampling in participatory research contexts: Participant-driven recruitment. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, *83*, i113-124. <https://doi.org/10.1007/s11524-006-9107-9>
- Volz, E., Wejnert, C., Cameron, C., Spiller, M., Barash, V., Degani, I., & Heckathorn, D. D. (2012). Respondent-Driven Sampling Analysis Tool (RDSAT) Version 7.1. Cornell University. Diakses melalui <http://www.respondentdrivensampling.org/>
- White, R. G., Hakim, A. J., Salganik, M. J., Spiller, M. W., Johnston, L. G., Kerr, L., Kendall, C., Drake, A., Wilson, D., Orroth, K., Egger, M., & Hladik, W. (2015). Strengthening the reporting of observational studies in epidemiology for respondent-driven sampling studies: "STROBE-RDS" statement. *Journal of Clinical Epidemiology*, *68*(12), 1463-1471. <https://doi.org/10.1016/j.jclinepi.2015.04.002>
- World Health Organization, Regional Office for the Eastern Mediterranean. (2013). *Introduction to HIV/AIDS and sexually transmitted infection surveillance: Module 4: Introduction to respondent-driven sampling*. <https://apps.who.int/iris/handle/10665/116864>
- Zeng, L., Li, J., & Crawford, F. W. (2019). Empirical evidence of recruitment bias in a network study of people who inject drugs. *The American Journal of Drug and Alcohol Abuse*, *45*, 460-469. <https://doi.org/10.1080/00952990.2019.1584203>