Respondent-Driven Sampling: Reaching Male Youth with Drug Use Experience in Thailand

Sawitri Thayansin¹

Abstract

Respondent-Driven Sampling (RDS) was developed and has been widely used for recruiting sample from hard-to-reach populations such as populations at risk of HIV, injection drug users, males who have sex with males, and commercial sex workers. This article reports on the use of RDS technique in a research among Thai male youth aged 15-24 years who ever used drugs. Beginning with seventeen respondents who were used as initial seeds for recruiting more eligible subjects, the research was able to include 749 male youth with drug use experience in the sample. Majority of the recruited youth (41.3%) reported amphetamine as their first drug, followed by those who used *mitragyna speciosa* (20.1%), marijuana (17.2%), and inhalant (13.7%). Distribution of the sample youth recruited by this technique was found to be relatively normal which suggested that RDS was a reliable strategy for recruiting a sample of hard-to-reach populations. It is believed that this sampling technique can also be useful for the study of general hidden populations.

Keywords: Respondent-Driven Sampling, hidden populations, drug users, Thailand

¹ Lecturer, National Institute for Child and Family Development, Mahidol University, 999 Phuttamonthon 4 Road, Salaya, Phuttamonthon, Nakhon Pathom, 73170, E-mail: sawitri_nung@yahoo.com

This paper is based on the research project on Adolescent’s Life Style supported by the grant from the Thailand Research Fund (TRF) to which the author gratefully acknowledges. Special thanks are also given to Prof. Mark VanLandingham of Tulane University, USA, and Lisa Johnston of the University of California for their valuable suggestions.
**Introduction**

Thailand has been fighting with drug problem for decades, and yet the sign of success seems to be far from sight. Despite increasing awareness and much effort to control it, drug use remains widespread. The recent decades have witnessed mushrooming of variety of substances, ranging from traditional locally-produced plants such as marijuana (*ganja*) and mitragyna speciosa (*kratom* leaves) to amphetamine, inhalants, cocaine, ecstasy and other sedatives. Another evidence of seriousness of this problem is indicated by the fact that number of new users of young age is increasing steadily.

Epidemiological information from the drug treatment centers showed that the number of youth patients (aged under 25 years) grew bigger since 1997 (7,317 youth or 18.8 percent of all patients), and the record reached 9,193 youth patients (or 23.2 percent) in 1998 (Office of the Narcotics Control Board, 2003a). In 2003, registered number of patients from 580 treatment centers throughout the country showed that 24,610 patients were adolescents and youth aged below 25. This accounted for 40.5 percent of the total drug patients of 60,758 people (Department of Medical Services and Institute of Health Research, 2004).

Another national source of information on the use of drug was the estimation of population who involved with drug that was conducted in 2001 and 2003. The estimate was based on the population aged between 12-65 years from all regions of the country. In 2003, the survey estimated the following number of people who involved with drugs: 5,439,300 persons ever used any drugs at least once in their lifetime; 533,780 persons used drug within the past 1 year; and 296,380 persons used drug within 30 days before the survey. Among current users (defined as using any drug within 30 days before the survey), 24.0 percent were young people aged 12-24 years. This proportion decreased from about 40 percent in 2001. Amphetamine was the most common drug among them, followed by mitragyna speciosa and marijuana respectively (Office of the Narcotics Control Board, 2003b).

Illicit drug use among Thai adolescents and youth has been the subject of public concern. The concern is based not only on a growing number of young users but more importantly on the consequences of drug use in terms of health cost and social impact as well. Although drug use among young people has been well aware of since the past decades, systematic information at the national level on use of certain drugs such as amphetamine, marijuana, opium, and heroin is limited. There exist scattered, small-scale studies but nearly all of them have focused on in-school youth. The only sources that can give information on the drug use for whole country are the government’s drug treatment centers. But the information from these sources include only the users who voluntarily come to the center for treatment. Many users who did not report themselves for treatment were not included in this database. The number, therefore, does not accurately (or nearly accurately) reflect the actual number of drug users in society; at best it reveals only the tip of the iceberg. To increase the chances of getting sufficient number of cases and reliable information for analysis an innovative strategy is needed for subject recruitment.
The study upon which this paper is based employed a new sampling strategy known as “Respondent-Driven Sampling” (RDS). RDS was first developed by Heckathorn (1997) and recently has been used by many researchers for recruiting a sample of hard-to-reach populations (Heckathorn, 1997; Broadhead et al., 1998; Heckathorn, 2002; Heckathorn, Semaan, Broadhead, & Hughes, 2002; Heckathorn & Rosenstein, 2002). This paper describes how RDS was employed in a Thai setting and how effective it was in recruiting a sample of youth population in a selected area who had drug use experience.

An Overview of the Respondent-Driven Sampling

Hard-to-reach and hidden populations, such as drug users, sex workers, males who have sex with males and undocumented immigrants are difficult to reach for collecting sensitive information and to construct a sampling frame of the individual members of the population. Respondent-Driven Sampling (RDS) has been proposed as a method to sample these populations by combining a form of chain-referral or snowball sampling with a strategy that allows the participants who are already included in the study to recruit the larger target population. This technique is based on the belief that the target people are best recruited by their own peers due to greater trust, or familiarity (Heckathorn, 1997). Since the time it was first developed the method has been widely used by many researchers in the studies of hard-to-reach populations (Heckathorn, 1997; Broadhead et al., 1998; Heckathorn & Rosenstein, 2002; Heckathorn et al., 2002). RDS is also useful in sampling the people who do not trust project staff members by relying on peer-driven chain referral strategies. It is predicated on the recognition that peers are better able than project staff members to locate and recruit peers whom they know as members of a hard-to-reach population.

As a chain-referral sampling technique RDS is more effective than the traditional chain-referral sampling such as snowballing for two reasons:

First, RDS could help to reduce the bias associated with masking (under-sampling reclusive respondents). This technique assumes that the best way to effectively access members of stigmatized or hidden populations is through their own peers. Therefore, the friendship network of existing members of the sample is used as the gateway to access those people, create a reliable sample, and reduce masking and the rate of refusal. In other words, RDS relies upon the participants who have already completed an interview to do recruitment or to induce their friends into the study.

Second, RDS employs dual incentives, ‘primary incentive’ and ‘secondary incentive’, which helps reduce refusal rate. Primary incentive is a reward to participants for being interviewed. Secondary incentive is a reward to those who have been interviewed for recruiting their peers into the study. This dual incentive will assure cooperation of participants to help recruit more sample from their peers. Using participants to recruit other subjects who may be their friends has more advantage since the chances of refusal are reduced.
RDS technique is implemented by the following procedures (Heckathorn, 1997):

1. Research staff recruits a limited number of participants who, after being interviewed, can serve as “seeds” for recruiting other participants into the study.

2. Seeds are offered incentives (often in-cash) to recruit their peers into the same interview they have just completed. Specifically, seeds are given recruitment coupons and told that if they pass the coupons on to peers who come for interview, they will be paid (secondary incentive) for each recruited peer.

3. All new recruits are offered the same dual incentives as are those who come before them. Everyone is rewarded both for completing the interview and for recruiting their peers into the research. In this way, incentives create an expanding system of chain-referrals in which each subject recruits more subjects, who recruit still more subjects, and so on, from wave to wave. To ensure that broad array of subjects have an opportunity to be recruited, to prevent the emergence of semi-professional recruiters, and to preclude turf battles over recruitment rights, each subject was limited to three initial coupons. A by-product of this recruitment quota is to increase the number of waves of recruitment required to saturate the population.

4. The traits defining membership in the population must be objectively verifiable, lest respondents react to the recruitment incentives by enlisting persons who are not part of the hidden population.

5. To prevent subject duplication and impersonation, a database is created to record subject’s identification such as identified physical characteristics (scar, tattoo, etc.). The database is useful as some subjects may seek to participate in the study more than once under different identities perhaps as a means to collect reward for recruitment.

6. The sampling process ends either when the targeted community is saturated, or when a minimum target sample size has been reached and the sample composition becomes stable with respect to the traits upon which the research focuses.

In Thailand, as elsewhere, drug use is a personal and sensitive behavior; it is not only associated with strong social stigma but is also subject to legal punishment. For this reason, it is difficult to recruit for interview by conventional means. In the past illegal behavioral data were generally gathered using time location sampling, snowball sampling, or convenience sampling of individuals within intervention programs (Office of the Narcotics Control Board, 2003b; The Department of Medical Services and the Institute of Health Research, 2004; Maneenin & Warasingsook, 2541). This is problematic since the samples from these methods, which potentially miss those who are “hidden” or less accessible, are not representative of the entire population of interest.
To increase the chances of getting sufficient number of cases and reliable information for analysis, the study on which this paper is based adopted the RDS strategy for recruiting participants who were male youth with drug use experience. In the following sections this paper gives a brief account of the target population and the research setting; it then describes the application of RDS to recruit sample, highlights some research findings, and finally identifies the lessons learned and the problems faced in the field research carried out from October to November, 2004 (Thayansin, 2006).

Target Population and the Research Setting

Target population for this study is male youth aged 15-24 years who had experience of using illicit drugs including marijuana, amphetamine, ecstasy, inhalant, opium, heroin, mitragyna speciosa, and cocaine. Youth aged 15-24 are chosen as target of this study for some important reasons. First, youth of these ages are vulnerable to risk behavior of various forms due to their curiosity to learn and experience new things, especially the ones that give them excitement and fun. Second, since youth - particularly those of the younger ages - tend to consider themselves independent from their family control, they are often influenced by their friends who themselves are not matured enough to make independent decision about what they should and should not do. Third, in recent years youth have been the center of national concern regarding drug problem. As reported earlier, available information indicates a rising trend of drug use among Thai youth and the age of new users is getting younger. (See “Introduction” above for more information.) Because of these reasons, the government has placed great significance to measures that lead to understanding drug use behavior among young people and to provide ways to keep them from being the victims of drugs.

Only male youth were included in the sample of this study. Females were not included because the rate of drug use among them was very low. Available statistics showed that of all the adolescents aged 15-24 years with drug use experience only 5 percent were female, the rest (95.0 percent) were males (Office of the Narcotics Control Board, 2001). This being the case, it would require a large sample size to be able to get sufficient number of female users for a meaningful statistical analysis, the requirement that was beyond capacity of this study.

The research was carried out in two districts of Kanchanaburi province located about 130 kilometer West of Bangkok. It covers an area of approximately 19,483 square kilometers which makes it the third largest province in Thailand after Chiang Mai and Nakhon Ratchasima. In 2003 (i.e. around the time the field research was carried out), about 826,169 inhabitants lived in this province; of these, about 80 percent were in rural areas. About 19 percent of the population was of the age of 15-24. Majority of the people engaged in agricultural activities. Most of them were of Thai ancestry with substantial number of Mon and Karen minority groups living in the peripheral hill areas. In economic terms, Kanchanaburi was doing well on a national scale, with annual economic growth of nearly 10 percent. The value
of Gross Provincial Product (GPP) was 50,989 million Baht in 2004, ranking 25th in the national scale. Per capita income in Kanchanaburi was 62,928 Baht per year, ranking 31th in the country. Important industries include sugar, agricultural products and jewelry. Tourism is also a main source of income for some local people as the province has rich tourist resources and a large number of visitors each year. For drug abuse situation, statistics from the Office of the Narcotics Control Board showed that in 2004 the number of patients treated for drug abuse in Kanchanaburi was 442 comparing to the whole country of 35,193 patients.

RDS Strategy and Procedure

In the present study, RDS process began with purposive selection of 17 male youth who ever had drug use experience to participate in the study. According to RDS, these research subjects were initial seeds. Their selection was made with assistance from facilitators, namely, health workers at the local health centers, village health volunteers, and abbots of local temples. All facilitators were informed about the purpose of the study and the selection criteria before identifying eligible youth and inviting them to come for interview. After completing the self-administered questionnaire the seeds received compensation as a primary incentive.

Each initial seed was then asked to recruit others three youth with similar drug experience who might be their friends to participate in the study. All seeds were given a brief training on the recruitment process, especially with regard to who should be recruited and how. Each was then given three coupons and was advised to give these coupons to the three other youth whom they recruited. Each coupon was printed with a serial number and other information to enable the field staff link coupons to the recruiter who gave it and to determine when the recruiters should be paid incentive. Moreover, coupons contained the study name, study location, time that the storefront opened, and a brief explanation of the study. The seeds were given another (secondary) reward if they recruited other participants exactly as instructed, not otherwise.

When new recruits arrived at the field office, they were questioned by the screener to ensure eligibility. A recruitee with no drug experience was excluded from the study.

All eligible recruits were asked for their verbal consent and, if agreed, they were asked to complete the questionnaire. To maximize openness of the answer respondents were assured that all information they gave together with their personal information would be kept confidential, they were asked not to give their names and addresses in the questionnaire. To ensure participants’ comfortability, special place with sufficient privacy was arranged for participants while they were completing the questionnaire.

After completion of the questionnaires by these recruits, the first wave in the RDS process was completed and the initial seeds who had recruited them were paid an additional (secondary) incentive or reward for recruiting their peers. The new wave of recruitment then began with those who had just finished the questionnaires being asked to act as new recruiters to bring in
three recruits each for interview. In this way, number of recruiters increased in subsequent waves resulting in increasing number of recruits until the desired number of cases for the study was reached.

To ensure that opportunity for being recruited is not limited to certain groups of youth only and to control overrepresentation of members of some groups, each recruiter was allowed to recruit up to three peers from their networks and each referral chain of recruitment was allowed to extend not beyond the fifth wave. The process of recruitment at each site continued in this way until no new participant was recruited. In some sites (for example, small communities), recruitment was terminated at the end of the second or third wave. On average, recruitment at each site was kept open for at least one week after the first participants (seeds) were interviewed. Figure 1 shows structure of the actual recruitment at a site where a single subject (seed) generated almost 100 recruits from the total of 5 waves of recruitment in this study.

Figure 1: An example of a five-wave recruitment network generated by RDS, beginning with a single seed aged 21 years.

Note: 1. Recruitment started with this single seed resulted in a total of 89 cases at the end of the fifth wave.
2. The number 21y, 19y, etc, in the diagram indicates a case and his age.
Results

General Characteristics and Drug Use Experiences

During the fieldwork period of 41 work-days, from October to November 2004, the RDS recruitment process yielded a sample of 749 male youth who ever used drugs at least once in their lifetime. Of these, 48.3 percent (362 participants) were current users who used drugs within the past 30 days before interview while 51.7 percent (387 participants) reported that their last drug use was more than 30 days prior to the interview.

Majority of the respondents (60.6 percent) were in the age-group of 15-19; the mean age of the entire sample was 19 years. About 40 percent (295 cases) were of the age of 20-24 years. The smaller proportion of youth in older age (20-24 years) was due in part to the size of actual population in this age-group and, in part, to the fact that during the survey months many young people in this age-group left their communities for various reasons including work and education. There was also a difference with regard to schooling status; the youth who were not students were of greater proportion (57.5 percent) comparing to those who were in-school (42.3%). Almost half of the in-school youth were studying at the high school level. Majority of those who were out-of-school (51.9%) completed middle school level; only 0.1 percent of this group had no formal schooling.

As mentioned earlier, all youth included in this study (749 cases) were those who ever used drug of any kind in their life. For the largest majority, drug use began between the ages of 15-19 regardless of the type of drug. The age at first use, however, varies greatly, ranging from the age of 6 to 23 years (Figure 2) with an average age of first use of 16.2 (Table 1). It can be seen from Figure 2 that drug debut begins quite early; for some youth this happens around age 10 and below. The use during early adolescent years (ages 12-15) rises sharply and reaches the peak at age 15 where about 22 percent of youth had the first try. After age 15 the proportion of first use falls steadily until age 20. And by age 23 the proportion involved in drug use is minimal. (Note, however, that this distribution is likely to be the effect of data truncation.) The finding seems to suggest a significant policy implication. Since drug use for most of the sample youth begins early in adolescent years, it calls for the program activities to pay special attention to adolescents of younger age as a target of intervention. If male youth in young adolescent years can be kept away from drug, it is likely that most of them will not turn to it in their twenties and possibly in the rest of their life.
Figure 2: Proportion of respondents who ever used drug, by age at first use (N = 745)

With regard to type of drug in the first use, Table 1 shows that majority of the sample youth (41.3%) began with amphetamine, followed by those who tried mitragyna speciosa or *kratom* leaves (20.1%), marijuana (17.2%), and inhalant (13.7%), respectively, in their first experience of drug. Youth who began with harder drug such as ecstasy, opium or heroin and cocaine altogether account for about 8 percent of all. The fact that amphetamine dominated the first drug use for the largest proportion of the sample reflected its relatively easy access around the time of the survey comparing to other drugs.
Table 1: Distribution of the sample youth who ever used any illicit drugs, by age at first use and type of drugs.

<table>
<thead>
<tr>
<th>Type of drugs</th>
<th>Total (%)</th>
<th>Age at first use</th>
<th>Percent (%)</th>
<th>Age at first use (year)</th>
<th>Age at first use (year)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Before age 15</td>
<td>Age 15-19</td>
<td>Age 20-24</td>
<td>Lowest</td>
</tr>
<tr>
<td>Any illicit drugs</td>
<td>100.0</td>
<td>17.3</td>
<td>75.3</td>
<td>7.4</td>
<td>6</td>
</tr>
<tr>
<td>Type of drug</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Amphetamine (Yaba)</td>
<td>41.3</td>
<td>17.2</td>
<td>75.6</td>
<td>7.1</td>
<td>11</td>
</tr>
<tr>
<td>2. Mitragyna speciosa (Kratom)</td>
<td>20.1</td>
<td>20.8</td>
<td>69.8</td>
<td>9.4</td>
<td>6</td>
</tr>
<tr>
<td>3. Marijuana (Ganja)</td>
<td>17.2</td>
<td>6.3</td>
<td>87.5</td>
<td>6.3</td>
<td>10</td>
</tr>
<tr>
<td>4. Inhalant</td>
<td>13.7</td>
<td>24.5</td>
<td>71.6</td>
<td>3.9</td>
<td>7</td>
</tr>
<tr>
<td>5. Ecstacy</td>
<td>3.4</td>
<td>8.0</td>
<td>80.0</td>
<td>12.0</td>
<td>9</td>
</tr>
<tr>
<td>6. Opium/Heroin</td>
<td>2.4</td>
<td>27.8</td>
<td>61.1</td>
<td>11.1</td>
<td>13</td>
</tr>
<tr>
<td>7. Cocaine</td>
<td>1.9</td>
<td>28.6</td>
<td>57.1</td>
<td>14.3</td>
<td>13</td>
</tr>
</tbody>
</table>

Low Homophily and Equilibrium Distribution of Participants

Although RDS is effective in reaching the target sample from hidden population, it is a non-probability sampling technique. This means that it is subject to sampling bias to some extent. For example, it may be possible that individuals who are recruited share similar attributes with the ones who recruit them. This might happen because there is a tendency that the recruiter selects only the persons like him/her (homophily). If this is the case, bias in the chain of recruitment will be difficult to avoid; and this will result in samples bias which becomes a limitation in an inferential analysis.

To this concern, Heckathorn (1997, 2002) argued that the choice of initial subjects does not introduce a bias into the sample. He demonstrated that as the RDS sample grows in size from wave to wave, the alleged bias is progressively reduced and the composition of the sample ceases to change (or reach a stable or equilibrium composition) if the sampling process is allowed to proceed through sufficient number of waves. Equilibrium is an important concept, which refers to the stage where sample characteristics cease to fluctuate and bias from the seeds chosen by project staff disappears when the number of waves is large. Generally, the stable or equilibrium sample composition will be reached after the fourth or fifth wave. It does not matter whether initial seeds were drawn from the same group or from different groups; the ultimate composition of the sample will be the same. Heckathorn, therefore, concludes that RDS can be considered a statistical sampling method.
Composition of the sample in the present study which employed RDS confirms what was pointed out by Heckathorn above, and this is shown in Figures 3A and 3B. Overall, the two figures show that distribution of the sample youth derived from initial seeds aged 15-19 are very much the same as those derived from the seeds who are 20-24 years old.

In Figure 3A which shows the result of initial seeds aged between 15 and 19 years, we got the youth sample of both younger and older age-groups. The distribution of youth in these two age-groups gradually converge until it becomes stable around the fifth wave where the proportions of those in the younger and older age-groups are about 60 percent and 40 percent respectively. In a similar fashion, when we started with the seeds from the age-group of 20-24 (Fig. 3B) the end results are the same; that is, we got the sample of 60 percent of younger youth and 40 percent of older youth.

Note that wave 0 refers to the seed, wave 1 refers to the seeds’ recruits, wave 2 refers to the recruits’ recruits, and so forth. Also note that stable distribution of recruits after the fourth or the fifth wave (i.e. about 60 percent of younger and 40 percent of older youth) as is the case in this study is reflecting the age structure of actual youth population in the research areas.

RDS is sometimes criticized for being based on social network of the recruiters and those who are potential recruits. In principle as well as in practice, seeds recruit participants from his social network. This excludes other people - those who do not have social network with the seed - from being recruited into the study. This could happen, but, considering the fact that most youth in a community do associate with others, it is believed that effect of this “selectivity” can be a minor problem in the RDS process. The study upon which this paper is based was aware of this and tried to minimize the problem by selecting recruiters from different parts of the study sites so that the sample youth were brought in from different parts representing diverse groups within the study area.

Because participants in RDS technique are given incentive, some people may try to show up for interview more than once. If this happens, it can violate the sampling criteria (and thus data quality suffers) which allow only one interview for each selected individual. To avoid this duplication, software for RDS called IRIS Plus (IRIS = Integrated Risk Information System) was used to track the coupons and payment for each respondent. The recorded information in IRIS Plus such as respondent’s code, coupon number, and the number of the coupons distributed enables the project staffs to link coupons together and reject duplicate coupon numbers. Without IRIS Plus in this study, the project staffs were advised to recognize respondents by face and visible physical characteristics such as scars, birthmarks, tattoos, etc. In addition, field staff also used other techniques such as screening questions to ensure that youth who showed up for interview were really new cases. In the entire field data collection of this research the field staff found only two such cases who appeared twice for interview. They were politely advised that each participant could be interviewed only once.
**Figure 3:** Age distribution of recruited cases starting with 15-19 year-old seeds compared to age distribution of the cases started with 20-24 year-old seeds (in percentage), RDS technique.

3A: Starting with seeds aged between 15-19 years

3B: Starting with seeds aged between 20-24 years
Lessons Learned

For reasons discussed above and with the cautions that we took while conducting fieldwork, it is believed that the sample derived from application of RDS technique in the present study is of acceptable quality and thus suitable for inferential statistical analysis without serious problems. Below are some of the lessons learned from applying this innovative technique.

RDS can be useful for sampling of populations who do not trust sponsoring agencies and for reaching persons who shun public venues. The non-government organizations’ staff in recruiting hidden population helped them dare appear and feel comfortable about the interviewing process.

It is very important that selected group of initial seeds must be of adequate number to ensure that recruitment chains are long enough to mitigate biases while reaching the pre-determined sample size. There is no appropriate number of seeds. For this study, the RDS began with seventeen seeds while the past studies conducted from six to twenty seeds. However, if few seeds die out, new seeds may have to be added later on in the study.

Although RDS use the dual incentive to increase recruitment, the incentive must be appropriate, not too large or too small, to motivate economically and socially diverse members of the target population. If the incentive is too large, it will increase the chances that someone pretends to be a member of the population of interest.

To ensure participants’ comfortability, the research storefront to which participants come for interview should be easy to reach, quiet, and private. Survey should be administered in community locations recommended by participants as accessible such as a corner in the temple ground, youth center, or a corner of other community places.

References


