

# Using Respondent Driven Sampling to Recruit Sexual Minority Women

Kelly R Martin\*, Timothy P Johnson†, Tonda L Hughes‡

Tags: sexual minority women, respondent driven sampling

DOI: [10.29115/SP-2015-0004](https://doi.org/10.29115/SP-2015-0004)

---

## Survey Practice

Vol. 8, Issue 1, 2015

---

**ABSTRACT:** Purpose: To describe: (1) recruitment of a community sample of sexual minority women (SMW) using respondent driven sampling (RDS), (2) and how RDS was modified to reach target recruitment goals, and (3) characteristics of successful recruiters. Methods: We compared the rates of recruitment using RDS and two modifications of RDS that we developed during our recruitment of SMW in Chicago. Key demographics of successful recruiters are described. Results: Using RDS without modifying the standard methodology provided the best rate of recruitment, but moved too slowly to be cost effective. Modifications proved necessary to meet recruitment goals. Participants who successfully recruited others tended to be younger, were African-American, have a high school education or less, and have an income of \$20,000 per year or less. Conclusion: Researchers using RDS to recruit SMW may need to consider modifying RDS methodology. Flexibility and adequate staff resources are important. Using RDS to fill specific cells required to obtain a diverse sample may be difficult given the nature of RDS and the demographic characteristics of successful recruiters.

---

## INTRODUCTION

Sampling rare and hard-to-find populations continues to be a challenge to health and social researchers. This paper describes our experiences with recruiting a sample of lesbian and bisexual women (i.e., sexual minority women [SMW]) as part of the 10-year longitudinal Chicago Health and Life Experiences of Women (CHLEW) study. As probability sampling is not yet feasible in studies of sexual minorities, we employed respondent-driven sampling (RDS), a refinement of chain-referral sampling developed by Heckathorn (1997, 2002) that uses initial participants or “seeds” as recruiters.

We provide an overview of the methods used to sample and recruit study participants, the challenges encountered, and the practical solutions implemented. To meet recruitment goals, we eventually implemented a modified form of RDS.

Three strands of recruitment were utilized, which we referred to as True Seeds,

---

\* **Institution:** University of Illinois at Chicago

† **Institution:** University of Illinois at Chicago

‡ **Institution:** University of Illinois at Chicago

Modified Seeds, and Last Gasp Seeds. Each strand is briefly described and compared to the other strands in terms of recruitment success. Characteristics of successful seeds are also described, and the sample is compared to our existing longitudinal panel which was recruited using convenience sampling. We conclude with recommendations for other researchers who may be interested in employing RDS sampling methodology to recruit hard-to reach populations.

## **BACKGROUND**

The CHLEW study is a National Institutes of Health-funded longitudinal study of SMW's health. Wave 1 data were collected in 2000–2001, and wave 2 in 2003–2004. The third wave of data collection (2010–2012) was completed approximately 6 years later.

The original sample was recruited using a range of convenience sampling strategies. CHLEW study staff made vigorous efforts to maximize CHLEW sample representativeness by including subgroups of lesbians underrepresented in most studies of SMW's health (aged under 25 and over 50, high school education or less, racial/ethnic minority). The study was advertised in local newspapers, on Internet listservs, and on flyers posted in churches and bookstores and distributed to individuals and organizations via formal and informal social events and social networks. Other recruitment sources included clusters of social networks (e.g., formal community-based organizations and informal community social groups) and individual social networks, including those of women who participated in the study. Interested women were invited to call the project office to complete a short telephone screening interview. Eligibility criteria were 18 years or older, English speaking, residence in the Chicagoland area, and lesbian self-identify. The initial screening question asked "Recognizing that sexual identity is only one part of your identity, how do you identify yourself?" Response options were lesbian/gay, bisexual, heterosexual, and transgender. Lesbian women were the initial target population; those who identified as anything else were excluded from the original sample (see Hughes et al. 2006, 2007).

In the second wave of data collection, 384 women from the original sample were re-interviewed, a response rate of 85.9 percent.

In 2009, we received funding for a 10-year follow-up of the original study participants and for recruitment of a new panel to add more racial/ethnic and age diversity and more bisexual women (target sample=350).

## **RECRUITMENT USING RDS**

RDS uses handpicked seeds, who meet study criteria, to initiate sampling chains. Seeds are also selected based on social network size, i.e., having a sufficient number of people in their friendship network who also meet study criteria. Seeds agree to be interviewed and to recruit peers from their social network. Each seed was given three numbered coupons to distribute that

described the study purpose and criteria and provided a telephone number for potential participants to call. In turn, each new participant was given three coupons and invited to recruit others into the study. Participants received \$20 for each eligible woman they recruited with the limit of three coupons serving as a safeguard to overrecruitment of those from a particular social network. Participants were paid the recruitment incentive after their referral was interviewed.

### **TARGET GROUPS**

Recruitment of the new study panel was designed to oversample Black, Latina, and younger lesbians (age 18–25). An open-ended screening question was used regarding sexual identity. Those who did not give 1 of 5 answers (lesbian, mostly lesbian, bisexual, heterosexual, transgender) were then asked: “Knowing that no label can fully capture your identity, if you had to choose one of the following terms to describe yourself, which one comes the closest? The five choices were then read to the potential participant. Those who did not identify as lesbian, mostly lesbian or bisexual were ineligible to participate.

Recruitment criteria were changed several times to adjust for over- or underrecruitment of specific groups in regard to age and race. Changing recruitment criteria affected the recruitment chains because participants who distributed coupons to social contacts who were willing to participate but who did not meet the current study recruitment needs were put on a waiting list and ultimately not interviewed.

Recruitment goals were exceeded – with some significant modifications and caveats to the RDS model. Below we briefly describe each strand and the rationale for modifying the RDS methodology.

### **TRUE RDS**

Possible seeds were suggested by several community leaders for our targeted groups. These True Seeds came from various parts of Chicago, as Heckathorn and Magnani (2004) suggest that seeds in proximity to one another will tend to recruit the same type of participants. The majority of seeds had a large social network. The question to determine the size of a seed’s possible social network was, “Of the individuals you know by name, how many would you say are White, African-American or Latina women, age 18 or older, who are lesbian or bisexual women and live in the Chicago area?” (Twenty one seeds were identified; 3–5 seeds in each targeted group.) As recruitment ebbed and flowed and recruitment chains appeared to end, we added more seeds, as Heckathorn and Magnani (2004) suggest that not all seeds need to be recruited at the beginning of the recruitment period.

In the True Seed strand, 525 coupons were given to 175 people for a return of 175 participants – a recruitment rate of 33.3 percent, or one for every 3 coupons (Table 1).

**Table 1** RDS Recruitment of sample by type of referral seed.

	Percent of seeds who recruited participants	Average number per seed	Average number for those who recruited at least one	Percent of total sample successfully recruited <sup>+</sup>	Number of coupons distributed	Rate of return	Longest recruitment chain/ Average number of waves in chains
True seeds (n=28)	46.4%	1.0	2.1	42.6% (n=175)	525	33.3%	10/3.6
Modified seeds (n=247)*	13.8%	0.2	1.6	55.1% (n=226)	1,263	17.9%	8/2.1
Last gasp seeds (n=12)	41.7%	0.7	1.6	2.2% (n=9)	56	16.1%	2/1.3
Total seeds	18.1%	0.3	1.7	NA <sup>++</sup>	1,844	20.1%	10/2.4

\*Seeds in the modified strand are all those given recruitment coupons.

<sup>+</sup>Our definition of successfully recruited are those who met study criteria – whether or not they were later dropped, not interviewed or put on our waiting list.

<sup>++</sup>Not Applicable in that the seeds recruited 100% of the non-seed participants.

## **STRATEGIES TO ACILITATE RECRUITMENT**

As part of our validation process, we called participants who had no coupons redeemed after 10 days and asked “Did you receive three recruitment coupons?” and “Have you given recruitment coupons to three women that you know?”

Most participants volunteered how many coupons they had given out or explained why they had not given out any. Common explanations included waiting for a certain social event to take place, not having seen any friends lately, and not knowing anyone who fit the criteria. Staff encouraged participants to give out their coupons or remind the people to whom they had given coupons to call the study office. When the study ended, we completed 190 successful calls: 44 participants subsequently had a successful referral – sometimes within a few hours or a day of the validation call. We mailed \$20 money orders to participants the next business day after their referral was interviewed, hoping that a prompt reward would have a positive impact on recruitment behavior, though we had no way of directly measuring its effect.

Assessment of recruitment after 6 months revealed that recruitment was moving much too slowly: we had sampled 21 seeds and obtained only 54 additional participants.

## **MODIFIED RDS RECRUITMENT**

Because we were also conducting the 10-year follow-up of our original sample, we decided to ask these women to assist with recruitment. In essence, these women became Modified Seeds. We gave coupons only to women who still lived in the Chicagoland area ( $n=247$ ). Of these, 14 percent ( $n=34$ ) successfully recruited one or more new participants.

In the Modified Seed strand, 1,263 coupons were given to 421 people for a return of 226 participants: a rate of 18.1 percent or one for every 5.6 coupons (Table 1).

About five months after we began using our original panel as seeds, one of our referral chains began to grow rapidly and outstrip other chains in length and number. A seed from the original sample who lived in transitional housing referred other women at this same address. We enrolled and interviewed 16 women from the same address before deciding not to accept additional residents. The most compelling reason for this decision was that these women appeared to have similar life experiences and represented what Heckathorn (2002) refers to as homophily (i.e., a group with similar characteristics). In this case, both poverty and propinquity made these participants too obviously similar. An additional reason to end this chain was our concern that women living in transitional housing would likely prove difficult to locate for potential follow-up interviews.

## **LAST GASP SEEDS**

As we neared our recruitment goal of 350 women, we decided to recruit 20–25 additional participants, primarily Latina lesbian and bisexual women. Because recruitment chains appeared to have ended, we advertised through a mass email from a community agency serving SMW and by posting our recruitment criteria on two Chicago area websites. The 12 women who called and met recruitment criteria became seeds. They had not been personally recruited and did not necessarily have large social networks. Thus, they were different from True Seeds. They were also different from Modified Seeds in that they had no familiarity or history with the study. Five (41.7 percent) of these seeds referred additional participants.

For this Last Gasp strand, 56 coupons were given to 17 participants with a return of 9 participants: a rate of 16.1 percent or one for every 6.25 coupons. The recruitment chains from these seeds ended almost immediately.

## **WHO RECRUITED?**

Compared to the 379 participants who were not successful in recruiting other eligible participants, the 234 women who were successful tended to be younger (35.8 vs. 41.8 years;  $t$ -test=9.4  $df$ =581,  $p$ <0.001), have a slightly larger social network (25.5 vs. 22.0 persons;  $t$ -test=1,  $df$ =501, ns), more likely to be African-American (50.4 percent vs. 29.8 percent), and less likely to be white (21.8 percent vs. 43.8 percent;  $\chi^2$ =37.1,  $df$ =3,  $p$ <0.001). In addition, a larger percentage reported having a high school education or less (71.5 percent vs. 45.4 percent;  $\chi^2$ =49.1,  $df$ =5,  $p$ <0.001) and over half (53.8 percent vs. 23.6;  $\chi^2$ =55.9,  $df$ =1,  $p$ <0.001) reported household income of less than \$20,000 per year (Table 2).

**Table 2** Demographics characteristics of those who successfully referred other participants.

	Wave 3 sample (n=727)	Those given coupons (n=613)	Successfully referred others (n=234)	Did not successfully refer others (n=379)	T-test or Chi-square values for those who referred vs. those who did not
	(n) %	(n) %	(n) %	(n) %	
Race/ethnicity					
White, non-Hispanic	(272) 37.4	(217) 35.4	(51) 21.8	(166) 43.8	$\chi^2=37.1$ , df=3, $p<0.001$
Black, non-Hispanic	(260) 35.8	(231) 37.7	(118) 50.4	(113) 29.8	
Hispanic	(168) 23.1	(149) 24.3	(57) 24.4	(92) 24.3	
Other (bi/multi-racial)	(27) 3.7	(16) 2.6	(8) 3.4	(8) 2.1	
Education level					
High school or less	(146) 20.1	(135) 22.1	(77) 33.0	(58) 15.3	$\chi^2=49.1$ , df=5, $p<0.001$
Some college	(224) 30.8	(204) 33.3	(90) 38.5	(114) 30.1	
College degree or higher	(355) 48.9	(272) 44.4	(66) 28.2	(206) 54.4	
Annual household income is less than \$20,000	(221) 31.8	(205) 33.5	(120) 53.8	(85) 23.6	$\chi^2=55.9$ , df=1, $p<0.001$
	M (SD)	M (SD)	M (SD)	M (SD)	
Mean age (standard deviation)	40.0 (14.0)	39.5 (14.4)	35.8 (13.5)	41.8 (14.5)	t-test=9.4 df=581, $p<0.001$
Average number in social network of African-American, Hispanic or White lesbian or bisexual women age 18 or older who live in Chicago area (SD)	22.1 (37.6)	23.5 (38.5)	22.0 (32.1)	25.5 (45.0)	t-test=1, df=501 $p<0.1$ (NS)

Note: One hundred and fourteen participants were not given coupons – either because they were part of the longitudinal cohort who no longer lived in the Chicago area or because recruitment was nearly complete.

Because of deliberate oversampling and specific age and race criteria, it is not possible to ascertain if our new panel approximates a probability sample. Rather than comparing the pool of referrals to the seeds as is usual in RDS, it may be more instructive to compare the new panel to our longitudinal panel. Studies of SMW have historically overrepresented White women, younger women and women with a higher level of education. Our longitudinal sample – recruited through convenience sampling, but with concerted efforts to reach women of color, older women and women with lower levels of education – was fairly representative of the general female population in Chicago in terms of race, but overrepresented those with a higher education (Chicago Fact Finder 2002). Our new panel has a better distribution of education than the longitudinal panel and is more representative of the general Cook County female population in this regard. Nearly one-third of the new panel has a high school education or less. However, nearly half (47 percent) of the women in our new panel report a yearly income of less than \$20,000 per year while 24.6 percent of all households in Chicago report a yearly income of less than \$20,000 per year (U.S. Census Bureau American Community Survey 2011). This may be due to several factors: nearly one-third of the new panel are younger women who may still be in school or who have lower-paying jobs; many represent female-headed households which tend to have lower incomes; and people with lower incomes may be more motivated to participate in social science surveys, particularly when monetary incentives are offered for both participating and referring others.

Using a modified version of RDS we were successful in recruiting women with lower socio-economic status, one of the primary differences between the new sample and our original convenience sample (Table 3).



**Table 3** Demographic characteristics of longitudinal panel and new panel.

	Longitudinal panel n (%)	New panel n (%)	Cook county %	Chicago %
Race/Ethnicity				
White, non-Hispanic	211 (47)	90 (24)	47	35
Black, non-Hispanic	125 (28)	164 (44)	25	34
Hispanic	88 (20)	113 (30)	20	24
Other (bi/multi-racial)	25 (05)	6 (2)**	8	7
Education level				
Less than high school	13 (3)	54 (15)	16	19
High school	48 (11)	64 (17)	25	24
Some college	135 (31)	149 (40)	25	24
Bachelor's degree	118 (26)	60 (16)	21*	21*
Graduate/professional	133 (30)	44 (12)	12*	12*
Annual household income				
Less than \$20,000	118 (26)	174 (47)	20	25
\$20,000–39,999	113 (25)	73 (20)	21	22
\$40,000–74,999	121 (27)	70 (19)	26	25
\$75,000 or more	95 (21)	30 (8)	33	29

Note: All Cook County and Chicago data are from the American Community Survey 1-year Estimates (U.S. Census Bureau 2011), for women age 18 and older. Annual household income is not directly comparable as it includes both genders and people age 16 and older.

\*Data only available for age 25 and older.

\*\* Multiracial identification needed to include White, Black, or Hispanic.

## CONCLUSIONS AND LESSONS LEARNED

Our recruitment goals were exceeded using a modified version of RDS in which we continually added new seeds. A total of 1,839 coupons were distributed to 613 women – a return of 410 possible participants (22 percent), or one participant for every 4.5 coupons. More than one-third (38 percent;  $n=238$ ) of those given coupons recruited participants – an average of 1.57 recruits per successful recruiter.

Although the True Seed strand – who were younger and had a larger social network – had the highest average number of referrals, we would not have reached recruitment goals using only this strand. More seeds could have been added to this strand over the 2 years of the study, but using a large number of handpicked seeds defeats the purpose of RDS.

We were fortunate to be able to employ our original sample as additional seeds. Although their social network size was unknown, we expected a better rate of return from this cohort given their long-term investment in the study. A possible explanation for the low rate of return is that, in general, participants recruited within  $\pm 5$  years of their age and because of our specific age, race, and sexual identity criteria older White lesbians in the longitudinal sample tended not to recruit.

The Last Gasp strand was least effective in recruiting. A self-selected group who

responded to web postings and/or the mass email, the number of women in their social networks who met our study criteria was unknown. The relatively poor response from this group may be because these seeds were neither personally recruited – nor had they previously participated in CHLEW. Referral chains from this group died a very quick death, suggesting that researchers should carefully consider whether to recruit seeds through advertising.

Although we gave out a substantial number of coupons, we were unable to assess how many coupons were actually distributed. Our fairly low response rate makes it difficult to believe that the population had been saturated. A recent Gallup Special Report (Gates and Newport 2012) found that 3.6 percent of adult women in the United States identify as lesbian, gay, bisexual, or transgender. The 2010 Census (U.S. Census Bureau 2010) estimates that 1,080,737 adult women live in Chicago; therefore, approximately 38,900 women would be expected to identify as SM. Changing our criteria, and thus our recruitment coupons, several times also hindered the recruitment process. Our criteria were complicated because of the several sexual identity categories and age limitations. A potential participant would have to read the coupon fairly carefully to ascertain eligibility.

We conclude by summarizing key lessons learned in conducting this project using RDS sampling methodologies.

**Allow ample time for recruitment.** Recruitment will likely move more quickly if researchers are not concerned with filling certain cells or with future follow-up of participants and are more willing to let chains (e.g., homeless women) play out to their natural end.

**Staff time.** Sufficient staff resources are needed to conduct follow-up of participants who are given recruitment coupons. Our 190 successful follow-up calls took an average of two attempts over several weeks.

**Flexibility.** Researchers must decide how much they are willing to modify the RDS methodology to reach recruitment goals.

**Timeliness of scheduling.** Scheduling and completing interviews as quickly as possible is crucial to facilitate coupon distribution. Referrals tended to call in 2–3 weeks after the person who referred them was interviewed. Whether this occurred because the referral delayed calling in, or because the participant delayed distributing coupons, is unknown.

**Motivation/personality of participants.** Motivation to participate and to recruit is clearly an important factor. Monetary incentives were a motivation for some, which likely accounts for overrepresentation of women with low incomes in our new panel. Even participants who were motivated to recruit likely faced time constraints, family issues, or health concerns. In addition, temperament may influence the ability to successfully recruit study

participants. For example, participants may be well-intentioned but procrastinate and thus delay scheduling their interview or distributing their coupons. Some participants may also have more social capital or be more persuasive in encouraging others to participate. A larger social network may give a participant a more diverse pool of people to recruit but does not necessarily correspond with timeliness in distributing coupons or persuading others to participate.

**Be directive about dissemination of coupons.** Social psychologists have noted that people are more likely to follow through with a task if a plan for doing so is explicitly stated. It may be helpful for the interviewer to ask specific questions about coupon dissemination. Examples are “Do you know who you will give your coupons to?” “When will you see this person next?” “What will help you remember to take the coupons with you?”

Although conceptually elegant, we found RDS challenging to implement in practice for the recruitment of SMW. It was successful, however, in enabling us to recruit several particularly difficult-to-reach population subgroups. Hence, we consider this a successful experience and recommend consideration of RDS sampling strategies when attempting to contact SMW.

#### **ACKNOWLEDGEMENTS**

This research was supported by National Institute on Alcohol Abuse and Alcoholism grants K01 AA00266 and R01 AA13328 (to Tonda L. Hughes). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Alcohol Abuse and Alcoholism or the National Institutes of Health. This paper was previously presented at the annual MAPOR conference in November of 2012 and we thank MAPOR for that opportunity. The authors would like to thank the women of Chicago who participated in the CHLEW study and especially to thank Robyn Nisi, who was instrumental in the implementation of respondent driven sampling in recruiting our new panel.

## REFERENCES

- Chicago Fact Finder. 2002. Institute for Latino Studies, University of Notre Dame. Notre Dame, IN. <http://www3.nd.edu/~chifacts/>.
- Gates, G.J., and F. Newport. 2012. "Special Report: 3.4% of U.S. Adults Identify as LGBT." Gallup Politics, Washington, DC. <http://www.gallup.com/poll/158066/special-report-adults-identify-lgbt.aspx>.
- Heckathorn, D.D. 1997. "Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations." *Social Problems* 44 (2): 174–99.
- . 2002. "Respondent-Driven Sampling II: Deriving Valid Population Estimates from Chain-Referral Samples of Hidden Populations." *Social Problems* 49 (1): 11–34.
- Heckathorn, D.D., and R. Magnani. 2004. "Snowball and Respondent-Driven Sampling." [http://tiger.uic.edu/~yoosik/classes/Soc509/Heckathorn\\_2004.pdf](http://tiger.uic.edu/~yoosik/classes/Soc509/Heckathorn_2004.pdf).
- Hughes, T.L., T.P. Johnson, S.C. Wilsnack, and L.A. Szalacha. 2007. "Childhood Predictors of Alcohol Abuse and Psychological Distress in Adult Lesbians." *Child Abuse & Neglect* 31 (7): 769–89.
- Hughes, T.L., S.C. Wilsnack, L.A. Szalacha, T.P. Johnson, W.B. Bostwick, R. Seymour, F. Aranda, P. Benson, and K.E. Kinnison. 2006. "Age and Racial/Ethnic Differences in Drinking and Drinking-Related Problems in a Community Sample of Lesbians." *Journal of Studies on Alcohol* 67 (4): 579–90.
- U.S. Census Bureau. 2010. The U.S. Census Bureau, Washington. 2010. <http://www.census.gov/2010census/>.
- . 2011. American Community Survey 1-Year Estimates-Washington, DC: The U.S. Census Bureau. 2011. <http://factfinder2.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t2011>.