Hidden populations, online purposive sampling, and external validity: Taking off the blind-fold

Monica J. Barratt

m.barratt@curtin.edu.au

National Drug Research Institute, Faculty of Health Sciences, Curtin University

54-62 Gertrude Street, Fitzroy Victoria 3065, Australia

Jason A. Ferris

j.ferris@uq.edu.au

Institute for Social Science Research, University of Queensland, Australia

Simon Lenton

s.lenton@curtin.edu.au

National Drug Research Institute, Faculty of Health Sciences, Curtin University, Australia

Corresponding author: Monica J. Barratt

Word count inc abstract & refs: 5530 (limit 6000)

Tables in text: 1

Tables in online supplementary: 2

Appendix in online supplementary: 1

1

Hidden populations, online purposive sampling, and external validity: Taking off the blind-fold

Online purposive samples have unknown biases and may not strictly be used to make inferences about wider populations, yet such inferences continue to occur. We compared the demographic and drug use characteristics of Australian ecstasy users from a probability (National Drug Strategy Household Survey, $n = 726$) and purposive sample (online survey conducted as part of a mixed-methods study of online drug discussion, $n = 753$) using non-parametric (bootstrap) and meta-analysis techniques. We found significant differences in demographics and drug use prevalence. Ideally, online purposive samples of hidden populations should be interpreted in conjunction with probability samples and ethnographic fieldwork.

3

Introduction

Due to the illegal and stigmatised nature of drug use, researching drug use and harm requires accessing hidden populations. Some scholars assert that stigmatised practices are best addressed using ethnographic approaches, where the research team participates in the cultural contexts of the target group over time, developing trust and rapport and understanding emic cultural logics (Moore 2005; Peterson et al. 2008; Sifaneck and Neaigus 2001). While ethnographic approaches are vital they do not provide externally valid large-scale quantitative data. For this purpose, the gold standard is the probability sample. By employing inferential statistics, probability sampling allows estimates of population prevalence to a given degree of *confidence*, given particular assumptions are met (Dillman 2007; Kakinami and Conner 2010; Lohr 2010). However, using probability sampling methods in the investigation of hidden populations is extremely limited. Firstly, response rates for general population surveys are decreasing (Groves 2006); moreso, when hidden populations are the target. A response rate of less than 50% seriously decreases confidence in relying on inferential statistics to adjust for non-response bias. Voluntary bias distorts the resultant weighted data. For example, the data from young males who complete a household survey will be weighted more heavily to make up for the lower participation rates of their cohort, yet the young males that respond may have very different characteristics to the young males that do not respond. This kind of systematic bias cannot adequately be adjusted for by demographic weightings (Caetano 2001; Zhao et al. 2009). Secondly, subsamples of a hidden population may be systematically excluded from probability survey sampling frames used in household surveys (i.e. coverage error). Lower-income, transient and young populations are systematically excluded (Zhao et al. 2009). Thirdly, probability sampling methods are expensive (Fricker 2008; Kakinami and Conner 2010) especially when less prevalent behaviours are targeted (Gile and Handcock 2010): if only 1% of the population are regular ecstasy users, then it would be likely to require a probability sample of 20,000 people to

produce a sample of 100 regular ecstasy users, assuming a 50% response rate.

An example of the kinds of problems occurring in 'gold standard' probability surveys can be found in the Australian National Drug Strategy Household Survey, used presently as comparative data. The NDSHS has a low response rate: 49%, 46% and 50% for the years 2007, 2004, and 2001 respectively (Australian Institute of Health and Welfare [AIHW] 2002; 2005; 2008). Although the NDSHS was weighted, this weighting cannot account for specific patterns of drug use that *were not* accessed using household probability sampling—that is, weighting is only based on socio-demographics, not drug use patterns (Caetano 2001; Zhao et al. 2009). It is not known whether the response rate indicates the systematic underreporting of drug use in the NDSHS. Furthermore, the young and transient, and mobile-phone only groups, who are not easily accessed through the household sampling frame are also likely to report different drug use patterns (Livingston et al. 2013).

Online purposive sampling

The changing context of conducting social research has combined with the limitations of probability sampling methods to necessitate new approaches to gathering externally valid data on hidden populations (Groves 2011). Purposive sampling is a method of understanding hidden populations that has a long history in the drugs field (e.g., Braunstein 1993; Peterson et al. 2008; Sifaneck and Neaigus 2001). Purposive sampling relies on the researchers situated knowledge of the field and rapport with members of targeted networks. Purposive sampling methods of hidden populations traditionally produced small samples and were associated only with qualitative research. However, in the last decade, purposive sampling of hidden populations through internet recruitment and survey methods has become increasingly popular among researchers who, at a relatively low cost, use it to engage large samples of people who are otherwise difficult to access, such as illicit drug users (Fricker 2008; Miller and Sønderlund 2010). For example, a sample of 100 regular ecstasy

5

users could be reached online for a mere fraction of the costs of employing probability sampling (e.g., Barratt and Lenton 2010; Miller et al. 2007).[1]

While the large samples achievable through online research methods lend themselves to quantitative analyses, the problem remains that the external validity of results arising from (online) purposive samples is unknown (Couper 2000; Lohr 2010). In much of this research, inferential statistics are applied to purposive sampling with the caveat that the results cannot be inferred or generalised beyond the sample itself (see Williamson 2003, for more detailed discussion). Despite the caveats, these data are indeed often used to infer information about the wider population of drug users in the absence of any other corroborating information.

Balancing rigour and practicality

We have argued that inferring to a wider population from a purposive sample is problematic. Yet, this practice continues, given the various limitations of probability sampling for hidden populations. Rather than considering all research based on purposive sampling as fully ungeneralisable, we explore the potential to use population sampling as a complement to purposive sampling *in order to improve* generalisability. We formalise what is already happening in the field and explore how it might be done more effectively, while being mindful of the real world constraints within which

---

[1] Respondent Driven Sampling methods have also recently been adapted for online networks (Wejnert and Heckathorn 2008). WebRDS has recently been used successfully to recruit samples of drug users, but only in the context of a wider-reaching study (Bauermeister et al. 2012). We believe WebRDS methodologies hold much promise, but that it would be problematic to use WebRDS to recruit drug users directly due to concerns around anonymity of drug users identified through social media and e-mail. Large samples of generically recruited people would still be needed to access adequate samples of drug users compared with targeted online sampling.

researchers operate. This is an exercise in attempting to balance rigour with the practicalities of doing research, especially with hard-to-reach populations (see Crosby et al. 2010).

It is also important that we explain what we are *not* trying to do. We do not suggest that a purposive sample can be used to estimate the prevalence of behaviours in a wider population. As explained by Lucas (2012), the social world is inherently 'lumpy', and some findings will apply differentially to some subsamples and not to others. We concur with Crosby and colleagues (2010) that the external validity of results gleaned from purposive samples of hard-to-reach populations is relevant, not simply in regards to prevalence estimates, but also in relation to describing the relationships between variables among samples. Although researchers using purposive samples may continue to qualify their results to the 'sample-at-hand' and studiously avoid making generalisations beyond this, we accept that at some point, researchers, or research consumers, often extrapolate findings beyond the sample studied. For that reason we believe it is important to develop methods which 'address the threat' of a 'lumpy social world' (see Lucas 2012) by using representative data to support that collected by purposive methods.

This paper tests how a purposive sample of ecstasy users recruited online and a probability sample of ecstasy users recruited from households differ by demographic and drug use characteristics. Unlike in previous studies (reviewed below), our samples were matched by frequency of use and location of residence. We also use a more appropriate statistical approach to compare these samples by utilising the non-parametric bootstrap method to estimate confidence intervals around the purposive samples (also reviewed below). Furthermore, the purposive sample was part of a larger mixed-methods project informed by virtual and multi-sited ethnography. In the discussion we reflect on the value of situating surveys of purposive samples into a framework of emic knowledge to better understand the sample's representativeness (see Guarte and Barrios 2006).

7

Representativeness of ecstasy user samples

Two comparisons between probability and purposive samples with regular ecstasy users have been undertaken (Miller et al. 2010; Topp et al. 2004) and we see our paper as extending this work.

Topp et al. (2004) compared a purposive sample of regular ecstasy users interviewed face-to-face with two probability samples of regular and recent ecstasy users derived from the NDSHS. When demographic and drug use variables were compared, substantial concordance was found across samples. In an extension of Topp et al.'s research, Miller et al. (2010) compared data from the first published online survey of Australian ecstasy users with both a purposive face-to-face sample and a probability sample (pencil-and-paper and telephone modes), also derived from the NDSHS. The web sample was significantly younger and less likely to report the recent use of other drugs (although more likely to report recent use of GHB) compared with the probability sample. While Miller et al. concluded that "the data provided by these three samples of ecstasy users converged" (p. 444), differences in how the samples were constructed limited interpretation of the results. The web sample was not matched with the other samples by frequency of ecstasy use, and the probability sample was not matched with the other samples by locale. For example, the differences in GHB use likely reflect elevated GHB use in Melbourne compared with other parts of Australia (see Stafford et al. 2005), rather than any differences in sampling type or survey mode. (Please see Table A, online supplementary file, for a tabulated summary of the methods and results of these two studies.)

Statistical approaches with purposive samples

It is typical to find many examples of market research (e.g., Bock and Sargeant 2002; Smith and Albaum 2005) and studies of hidden populations (e.g., Braunstein 1993; Hathaway 2004) where purposive, non-probabilistic sample methodologies have been used. Alongside this growing trend,

8

statistical approaches that attempt to account for bias have been improving. For example, recent publications have begun exploring the utility of 'respondent-driven sampling' (RDS, see Gile and Handcock 2010), a purposive sampling methodology, to recruit samples of MDMA/Ecstasy users (Wang et al. 2005). In situations where samples of a study population are recruited through purposive or convenient designs, such as RDS or the purposive sample used in this study, it is considered inappropriate to draw inferences from the sample to a greater population. Notably, recent research (Heckathorn 2011; Salganik 2006; Salganik and Heckathorn 2004) has demonstrated how RDS can produce unbiased prevalence estimates from the sample and then use bootstrap methods to produce confidence intervals of these estimates. Guarte and Barrios (2006) adopt a similar approach to Salganik et al. to illustrate the benefits of bootstrapping confidence intervals from prevalence estimates drawn from purposive sampling data to estimate population parameters. Bootstrapping is a statistical procedure that provides a way of estimating standard errors and other statistical parameters of interest drawn from the sample data available. Bootstrap confidence intervals treat the purposive sample as the 'population' of interest by taking random samples from that 'population' to estimate confidence limits (Rodgers 1999). According to Adèr, Mellenbergh, and Hand (2008) bootstrap procedures are recommended when the theoretical distribution of a statistic is unknown or complicated. An advisory panel for online public opinion survey established for the Canadian government (Public Works and Government Services Canada 2009) established that bootstrap procedures should be implemented to assess the statistical properties of non-probability data and compared to any available data drawn from probabilistic samples. In essence, bootstrapped results produce unbiased estimates from the available sample; but this does not guarantee that the sample of interest is not inherently biased given the nature of the sampling.

9

Aim of current study

This paper aims to extend the work of Topp et al. (2004) and Miller et al. (2010) by comparing samples of both regular and occasional ecstasy users recruited online with randomly selected probability samples from the NDSHS. We are interested in testing the external validity of online purposive samples, given that their use is increasingly common among hard-to-reach groups.

Our analysis differs from these previous papers in the following ways:

1. Unlike Miller et al., our samples are otherwise matched by frequency of use and location of residence.

2. Appropriate non-parametric statistical tests are used with the purposive sample, including bootstrapping (Miller et al. used parametric statistical analysis with a purposive sample, whereas Topp et al. used descriptive statistics).

3. Our online purposive sample is grounded in multi-sited and virtual ethnography.

Methods

Data sources

Online purposive sample

A web survey, conducted in 2007 as part of the first author's PhD, explored the use of online forums by people who use psychostimulants and hallucinogens ('party drugs'). The survey was part of a mixed-methods project guided by virtual and multi-sited ethnographic methods, involving participation and observation of internet forums where drugs were discussed (Barratt and Lenton 2010). Multi-sited ethnography is spatially decentred, tracing networks of people rather than constructing place-based boundaries around fieldwork sites (Falzon 2009). Virtual ethnography occurs in online spaces and is, by the nature of online spaces and networks, also multi-sited (Hine

10

2008). Most of the 837 respondents (74%) reported finding out about the study through a 'thread in online forum', 19% reported being 'referred via email/through internet', 6% 'saw the link on a social networking site', and 2% were 'referred by word-of-mouth (offline)'.

No reimbursements or other incentives were offered to survey participants. Offering reimbursement or prizes provides motivation for multiple responding (Bowen et al. 2008) and also involves taking identifying information from participants. Instead, we emphasised the anonymous nature of participation in our study, which did not collect Internet Protocol (IP) addresses or email addresses, which could potentially be used to identify individuals. While we cannot identify respondents who may have taken the survey more than once from the same computer, the lack of financial incentive and the 15-minute time commitment reduced the motivation for respondents to deliberately respond more than once. The data were also inspected for duplicate entries (same data submitted at a similar timepoint) and none were found. Although we did not collect IP addresses within the dataset, we did use Google Analytics to conduct traffic analysis on the URL used to promote the survey. These data indicate that 95% of the website visits over the study period were unique visitors suggesting that multiple responding from the same computer could not have grossly affected the data.

The final sample consisted of 837 Australian residents aged 16 and over who reported recent (last 12 month) use of party drugs and recent (last 6 month) participation in online drug discussion. For this paper, two subsamples were drawn: occasional ecstasy users ($n = 328$) who reported ecstasy use from 1 to 5 times in the last 6 months, and regular ecstasy users ($n = 425$) who reported ecstasy use monthly or more often in the last 6 months.

Probability sample

The 2007 NDSHS was the ninth iteration of a series of surveys asking Australians about their knowledge of and attitudes towards drugs and their history of drug consumption and related

behaviours (AIHW 2008). It used a multistage, stratified area random sample design with households selected from within individual census collection districts within each geographic region of Australia. Two survey modes were used: 'Drop and Collect' and 'Computer Assisted Telephone Interviewing (CATI)'. A total of 23,356 people successfully completed the survey, mostly via the Drop and Collect mode ($n = 19,818$; 85%). The overall response rate was 49%. Design weights provided by the AIHW were used to adjust for imbalances arising in the design and execution of the sampling. For comparative analysis with the online sample data two subsamples were drawn from those aged between 16 and 60 years: occasional ecstasy users ($n = 546$) who reported ecstasy use 'every few months' or 'once or twice a year' in the last 12 months, and regular ecstasy users ($n = 180$) who reported ecstasy use 'about once a month', 'once a week or more' or 'every day' in the last 12 months.

Measures

Most survey items were comparable between the two samples; however, there were some differences. Recent drug use in the NDSHS referred to the last 12 months, whereas the online survey measured use in the last 6 months. We acknowledge that this difference biases our comparisons towards underreporting drug use prevalence in the online sample due to the shorter duration defined as recent use; however, as this was counter to the likely difference between the two samples, this was less of a concern than if the bias were reversed. There were also differences in the way drug types were described: the online survey asked specifically about three different types of hallucinogens, a category that was aggregated in the NDSHS, and frequency of drinking alcohol had additional response categories in the NDSHS which were collapsed into 'weekly or more often' to match the online survey item. The items used to derive recent and regular illicit drug use were sufficiently similar to be considered comparable (see Appendix) and any differences have been taken into consideration in our interpretation of the results. Readers can view the NDSHS

12

questionnaire in the *First Results* report (AIHW 2008) and the online survey is available by contacting the first author.

Analysis

Means for continuous variables or percentages for categorical variables are presented for each sample. Stata 11.1 was used to estimate confidence intervals. Design weights, and weighted numbers and percentages were reported for the NDSHS data. Non-parametric bootstrap confidence intervals were estimated around comparable estimates and proportions from the purposive sample. Subsampling was repeated 250 times to generate each estimate. While bootstrap confidence intervals are better suited to estimating confidence intervals using the convenience sample as the population, standard linearised confidence intervals (not shown) also produced very similar results. To compare the results from the two samples directly, analysis was undertaken using techniques typically applied in meta-analysis studies (Sterne 2009) using an alpha level of .05. This study was approved by the Curtin University and Turning Point Alcohol and Drug Centre ethics committees.

Results

[Insert Table 1 about here]

There were statistically significant differences in age, sex, rurality, and education between occasional ecstasy users recruited online versus those who participated in the NDSHS (see Table 1). Occasional ecstasy users recruited online were younger (mean difference [M_{diff}] = 3.9 years, 95% CI 2.8–4.9, p<.001) and a greater proportion were male (M_{diff} = 7.5%, CI 0.3–14.5, p=.040), living in capital cities (M_{diff} = 9.6%, CI 3.3–16.0, p=.003) and had completed secondary education (M_{diff} = 16.9%, CI 10.7–23.2, p<.001). Statistically significance differences between the regular using samples were less pronounced, but in the same direction, with the regular-using online sample being younger (M_{diff} = 1.9, CI 0.6–3.2, p=.003), more likely to report living in a capital city (M_{diff} =

13

12.7%, CI 1.6–23.8, p=.025) and having completed secondary education (M_{diff} = 11.7%, CI 1.1–22.3, p=.030) than the matched NDSHS sample. The proportion reporting completion of a post-secondary-school qualification did not differ between either samples.

With respect to 'other' drugs there were a number of differences between the occasional ecstasy users in the online purposive sample and the equivalent probability sample (see supplementary Table B). Occasional ecstasy users in the online purposive sample were more likely to report recent use of meth/amphetamine (M_{diff} = 20.4%, CI 13.2–27.6, p<.001), hallucinogens (M_{diff} = 24.8%, CI 18.8–30.7, p<.001), ketamine (M_{diff} = 11.8%, CI 7.7–15.9, p<.001), GHB (M_{diff} = 5.7%, CI 2.9–8.5, p<.001) and benzodiazepines for non-medical use (M_{diff} = 17.7, CI 12.1–23.3, p<.001) than the matched NDSHS sample. Occasional ecstasy users from the online sample were also more likely to report regular use of hallucinogens (M_{diff} = 6.4%, CI 3.5–9.2, p<.001) and benzodiazepines (M_{diff} = 9.2%, CI 5.0–13.4, p<.001). In the regular-using samples, recent use of heroin (M_{diff} = 4.5%, CI 0.4–8.7, p=.032) and cocaine (M_{diff} = 11.7%, CI 0.4–23.0, p=.043), and having ever injected a drug (M_{diff} = 7.2%, CI 0.2–14.1, p=.043), were significantly greater in the NDSHS sample, whereas recent use of hallucinogens was significantly greater in the online sample (M_{diff} = 11.0%, CI 0.9–21.2, p=.033). Occasional ecstasy users in the NDSHS sample were more likely to report drinking alcohol monthly or more often (M_{diff} = 6.8%, CI 2.1–11.4, p=.004) and weekly or more often (M_{diff} = 14.5%, CI 7.4–21.5, p<.001) than the online purposive samples. Similarly, regular ecstasy users in the NDSHS sample were more likely to report drinking alcohol monthly or more often (M_{diff} = 8.5%, CI 4.0–13.1, p<.001) and weekly or more often (M_{diff} = 18.0%, CI 10.6–25.4, p<.001) than the online purposive samples.

Discussion

Purposive sampling of drug users through online methods increasingly occurs. Even though we make caveats that the findings should not be generalised, findings are often generalised regardless,

14

perhaps due to a lack of alternative 'better' evidence. One way to test the representativeness of purposive samples is to compare their characteristics with carefully matched probability samples, where such samples exist. We have performed this comparison here. When matched by time, locale and frequency of ecstasy use, online purposive and NDSHS probability samples differed considerably, with the online samples generally younger, better educated, and more likely to report concurrent use of a range of stimulant and hallucinogenic drugs (i.e. 'party drugs'). Coverage error and voluntary bias operate across both types of sampling. When used together, it may be possible to increase overall coverage of hard-to-access populations like ecstasy users, who are more likely to be the young and mobile groups least reached by household survey methods.

Our results regarding drug use estimates deviate from those reported by Miller et al. (2010): their online-purposive sample reported less recent drug use than their probability sample. Our results challenge Miller and colleagues' suggestion that the internet "may be an effective way to recruit more dedicated, less polydrug-using ERD [ecstasy] users than other methods" (2010, p. 444). As previously mentioned, Miller's results may be explained by the two samples being mismatched on frequency of ecstasy use and locale. In the present study, online recruitment, specifically through online forums where drugs are discussed, provided access to samples of occasional ecstasy users who were *more* likely to report the use of other 'party drugs' and regular ecstasy users who were *just as likely* to report other 'party drug' use.

Value of ethnographic grounding

The online purposive sample reported in this paper formed part of a larger project that also included online interviews and online participant observation, informed by virtual and multi-sited ethnography (see Barratt 2012; Barratt and Lenton 2010). We identify three reasons why an ethnographic grounding is useful for online purposive surveys. First, developing an understanding of the sites and networks within which participants interact and interacting with them to gain their

15

trust and respect is a vital part of the online recruitment process (Barratt and Lenton 2010; Potter and Chatwin 2011). While this step is critical for traditional 'face-to-face' ethnography, it is particularly important in virtual ethnographies where trust can only be built through online interaction. Second, participant observation of online sites generates additional research sites. In this case 40% of 40 forums were found through observing already-known forums. Thus, participant observation enabled the inclusion of a wider variety of online communities in the study. Third, we are better able to interpret the survey findings. For example, we know that the forums that were most successful in recruiting participants to the online survey hosted online discussions about the non-medical use of prescription drugs and the use of novel hallucinogens. Thus, it is not surprising that our purposive sample reported significantly more use of these substances than the matched household survey sample.

Limitations

This study has some limitations. Like Miller et al. (2010) and Topp et al. (2004), survey items between samples were not identical. Specifically, the alcohol question differences were likely to have biased the results towards more recent and regular alcohol use in the household survey given the more acute time frame, and the composite item measuring three different types of hallucinogens was likely to have biased the results towards more recent and regular hallucinogen use in the online survey. We believe that the differences in question wording cannot explain the other estimates that differed significantly which were assessed using single comparable items in both surveys. Wider windows of recall are likely to increase measurement error (Dillman 2007) and this fact informed the use of a 6-month window in the online survey; however, the difference in time frames between samples is a limitation of this study. Identical questions and increased sample sizes—more specifically greater observations in population studies—would strengthen future studies of this type.

Future directions

A general limitation of these kinds of studies is confounding of mode of data collection with sampling technique. Future research would benefit by implementing online versions of household drug use prevalence surveys aimed at improving access to younger subgroups of recreational drug users drawn from a probability sampling frame. Using identical items, order of presentation and context would allow for comparisons between internet purposive, internet probability and standard probability sampling modes. Teasing out how these samples differ will further our understanding of how best to interpret both bias in purposive samples and prevalence estimates of probability samples. Such work is also important because it informs interpretations of results arising from internet research methods that are increasingly utilised for drug research.

Conclusion

Hidden populations cannot be effectively and meaningfully analysed in generalised population-based surveys as they are often out of scope or hard to reach. Obtaining generalisable results from such methods would be financially unviable. Consequently, convenience samples often are employed. But if these are increasingly rejected by many in the research community, researchers of hard-to-reach populations are stuck between a rock and a hard place. Yet, if purposive sample studies can be combined with comparable probability samples and ethnographic fieldwork, we can have more confidence in understanding and evaluating the external validity of purposive samples, and interpret the representativeness of the resultant findings. Probability samples can be enhanced by purposive online samples gathered and informed through participant observation of online discussions. We hope this paper will prompt readers to take off the blind-fold, admit that (online) purposive sampling of hidden populations happens and will continue to happen, and focus on developing better ways of harnessing these data in order to make more accurate inferences about hidden and stigmatised practices use in the wider population.

17

References

Adèr, H. J., G. J. Mellenbergh, and D. J. Hand. 2008. *Advising on research methods: A consultant's companion*. Huizen, The Netherlands: Johannes van Kessel Publishing.

Australian Institute of Health and Welfare [AIHW]. 2002. "2001 National Drug Strategy Household Survey: First results." Author, Canberra.

—. 2005. "2004 National Drug Strategy Household Survey. First results." Author, Canberra.

—. 2008. "2007 National Drug Strategy Household Survey. First results." Author, Canberra.

Barratt, M. J. 2012. "The efficacy of interviewing young drug users through online chat." *Drug and Alcohol Review* 31:566–572.

Barratt, M. J. and S. Lenton. 2010. "Beyond recruitment? Participatory online research with people who use drugs." *International Journal of Internet Research Ethics* 3:69-86.

Bauermeister, J. A., M. A. Zimmerman, M. M. Johns, P. Glowacki, S. Stoddard, and E. Volz. 2012. "Innovative recruitment using online networks: lessons learned from an online study of alcohol and other drug use utilizing a web-based, respondent-driven sampling (webRDS) strategy." *Journal of Studies on Alcohol and Drugs* 73:834-8.

Bock, T. and J. Sargeant. 2002. "Small sample market research." *International Journal of Market Research* 44:235-244.

Bowen, A., C. Daniel, M. Williams, and G. Baird. 2008. "Identifying multiple submissions in Internet research: Preserving data integrity." *AIDS and Behavior* 12:964-973.

Braunstein, M. S. 1993. "Sampling a hidden population: noninstitutionalized drug users." *AIDS Education and Prevention* 5:131-40.

Caetano, R. 2001. "Non-response in alcohol and drug surveys: a research topic in need of further attention." *Addiction* 96:1541-1545.

Couper, M. P. 2000. "Web surveys: A review of issues and approaches." *Public Opinion Quarterly* 64:464-94.

Crosby, R. A., L. F. Salazar, R. J. DiClemente, and D. L. Lang. 2010. "Balancing rigor against the inherent limitations of investigating hard-to-reach populations." *Health Education Research* 25:1-5.

Dillman, D. A. 2007. *Mail and Internet surveys. The tailored design method*. Hoboken, NJ: Wiley.

Falzon, M.-A. 2009. "Introduction." pp. 1-23 in *Multi-sited ethnography: Theory, praxis and locality in contemporary research*, edited by M.-A. Falzon. Surrey, England: Ashgate.

Fricker, R. D. 2008. "Sampling methods for web and e-mail surveys." pp. 195-216 in *The Sage Handbook of Online Research Methods*, edited by N. Fielding, R. M. Lee, and G. Blank. London: Sage.

Gile, K. J. and M. S. Handcock. 2010. "Respondent-driven sampling: An assessment of current methodology." *Sociological Methodology* 40:285-327.

Groves, R. M. 2006. "Nonresponse rates and nonresponse bias in household surveys." *Public Opinion Quarterly* 70:646-675.

—. 2011. "Three Eras of Survey Research." *Public Opinion Quarterly* 75:861-871.

Guarte, J. M. and E. B. Barrios. 2006. "Estimation Under Purposive Sampling." *Communications in Statistics - Simulation and Computation* 35:277-284.

Hathaway, A. D. 2004. "Cannabis careers reconsidered: transitions and trajectories of committed long-term users." *Contemporary Drug Problems* 31:401-423.

Heckathorn, D. D. 2011. "Snowball versus respondent-driven sampling." *Sociological Methodology* 41:355-366.

Hine, C. 2008. "Virtual ethnography: Modes, varieties, affordances." pp. 257-270 in *The Sage handbook of online research methods*, edited by N. Fielding, R. M. Lee, and G. Blank. London, England: Sage.

Kakinami, L. and K. R. Conner. 2010. "Sampling strategies for addiction research." pp. 27-42 in *Addiction Research Methods*, edited by P. G. Miller, J. Strang, and P. M. Miller. Chichester, UK: Wiley-Blackwell.

Livingston, M., P. Dietze, J. Ferris, D. Pennay, L. Hayes, and S. Lenton. 2013. "Surveying alcohol and other drug use through telephone sampling: a comparison of landline and mobile phone samples." *BMC Medical Research Methods*. Accepted for publication.

Lohr, S. L. 2010. *Sampling: Design and analysis*. Pacific Grove, CA: Brooks/Cole.

Lucas, S. R. 2012. "Beyond the existence proof: ontological conditions, epistemological implications, and in-depth interview research." *Quality and Quantity*, *Early View*. doi: 10.1007/s11135-012-9775-3

Miller, P. G., J. Johnston, M. Dunn, C. L. Fry, and L. Degenhardt. 2010. "Comparing probability and non-probability sampling methods in ecstasy research: Implications for the Internet as a research tool." *Substance Use and Misuse* 45:437-450.

Miller, P. G., J. Johnston, P. R. McElwee, and R. Noble. 2007. "A pilot study using the internet to study patterns of party drug use: Processes, findings and limitations." *Drug and Alcohol Review* 26:169-174.

Miller, P. G. and A. L. Sønderlund. 2010. "Using the internet to research hidden populations of illicit drug users: A review." *Addiction* 105:1557-1567.

Moore, D. 2005. "Key moments in the ethnography of drug-related harm: reality checks for policy-makers?" pp. 433-442 in *Preventing harmful substance use. The evidence base for policy and practice*, edited by T. Stockwell, P. J. Gruenewald, J. W. Toumbourou, and W. Loxley. New York: John Wiley and Sons.

Peterson, J. A., H. Schacht Reisinger, R. P. Schwartz, S. G. Mitchell, S. M. Kelly, B. S. Brown, and M. H. Agar. 2008. "Targeted sampling in drug abuse research: a review and case study." *Field Methods* 20:155-170.

Potter, G. R. and C. Chatwin. 2011. "Researching cannabis markets online: some lessons from the virtual field." pp. 69-84 in *Markets, methods and messages. Dynamics in European drug research*, edited by J. Fountain, V. A. Frank, and D. J. Korf. Lengerich: Pabst Science Publishers.

Public Works and Government Services Canada. 2009. "The Advisory Panel on Online Public Opinion Survey Quality." Public Works and Government Services Canada, Ontario.

Rodgers, J. L. 1999. "The bootstrap, the jackknife, and the randomization test: a sampling taxonomy." *Multivariate Behavioral Research* 34:441-456.

Salganik, M. 2006. "Variance estimation, design effects, and sample size calculations for respondent-driven sampling." *Journal of Urban Health* 83:98-112.

Salganik, M. J. and D. D. Heckathorn. 2004. "Sampling and estimation in hidden populations using respondent-driven sampling." *Sociological Methodology* 34:193-240.

Sifaneck, S. J. and A. Neaigus. 2001. "The ethnographic accessing, sampling and screening of hidden populations: heroin sniffers in New York City." *Addiction Research and Theory* 9:519-543.

Smith, S. M. and G. S. Albaum. 2005. *Fundamentals of marketing research*. London: Sage.

19

Stafford, J., L. Degenhardt, M. Agaliotis, F. Chanteloup, J. Fischer, A. Matthews, J. Newman, P. Proudfoot, M. Stoove, and J. Weekley. 2005. "Australian trends in ecstasy and related drug markets 2004. Findings from the Party Drugs Initiative (PDI)." National Drug and Alcohol Research Centre, Sydney.

Sterne, J. A. C. 2009. *Meta-analysis in Stata: An updated collection from the Stata Journal*. College Station, TX: Stata Press.

Topp, L., B. Barker, and L. Degenhardt. 2004. "The external validity of results derived from ecstasy users recruited using purposive sampling strategies." *Drug and Alcohol Dependence* 73:33-40.

Topp, L., C. Breen, S. Kaye, and S. Darke. 2002. "NSW Party Drug Trends 2001. Findings of the Illicit Drug Reporting System (IDRS) Party Drugs Module." National Drug and Alcohol Research Centre, University of New South Wales, Sydney.

Wang, J., R. G. Carlson, R. S. Falck, H. A. Siegal, A. Rahman, and L. Li. 2005. "Respondent-driven sampling to recruit MDMA users: a methodological assessment." *Drug and Alcohol Dependence* 78:147-57.

Wejnert, C. and D. D. Heckathorn. 2008. " Web-based network sampling: efficiency and efficacy of respondent-driven sampling for online research." *Sociological Methods and Research* 37:105-134.

Williamson, G. R. 2003. "Misrepresenting random sampling? A systematic review of research papers in the Journal of Advanced Nursing." *Journal of Advanced Nursing* 44:278-288.

Zhao, J., T. I. M. Stockwell, and S. Macdonald. 2009. "Non-response bias in alcohol and drug population surveys." *Drug and Alcohol Review* 28:648-657.

Table 1

Demographic characteristics of occasional and regular ecstasy users for the online purposive sample matched with the 2007 NDSHS probability samples

Demographic variables	Occasional ecstasy users (less than monthly)				Regular ecstasy users (monthly or more often)			
	Online purposive (N = 328)		NDSHS probability (N = 546) [a]		Online purposive (N = 425)		NDSHS probability (N = 180) [a]	
	n	% (95% CI Boot)	n	% (95% CI Linear)	n	% (95% CI Boot)	n	% (95% CI Linear)
Age (valid n)	328		546		425		180	
mean	24.2	(23.6–24.8)	28.0	(27.2–28.9)***	22.9	(22.3–23.5)	24.8	(23.6–26.0)**
SD (range)	5.72	(16–51)	7.69	(16–55)	6.25	(16–51)	5.76	(16–55)
Sex (valid n)	326		546		423		180	
Male	221	67.8 (62.3–73.3)	330	60.3 (55.3–65.4)*	313	74.0 (69.9–78.0)	120	66.5 (57.5–75.5)
Rurality classification (valid n)	326		546		415		180	
Capital city	266	81.6 (77.3–85.8)	393	71.9 (67.1–76.7)**	327	78.8 (75.1–82.5)	119	66.1 (55.5–76.7)*
Year of school completed [b] (valid n)	306		498		386		166	
Year 12 (secondary school)	264	86.3 (82.6–89.9)	345	69.3 (64.4–74.2)***	307	79.5 (75.6–83.5)	112	67.8 (57.9–77.7)*
Highest qualification [b] (valid n)	283		504		360		167	
Any qualification [c]	186	65.7 (60.3–71.1)	305	60.6 (55.3–66.0)	215	59.7 (54.8–64.6)	102	60.9 (50.7–71.1)
University qualification	98	34.6 (29.2–40.1)	142	28.1 (23.2–33.1)	93	25.8 (21.3–30.4)	33	19.5 (10.9–28.1)

[a] All household survey estimates are weighted. [b] Excludes those 'still at school'. [c] Includes trade or technical certificate or diploma, undergraduate and postgraduate qualifications.

* p<0.05; ** p<0.01; *** p<0.001

1

Table A

Selected methodological, demographic and drug use characteristics of ecstasy user samples described in Topp et al. and Miller et al.

Sample characteristics	Topp et al. 2004			Miller et al. 2010		
	NDSHS recent [a]	NDSHS regular	Purposive	Probability	Face-to-face non-probability	Internet non-probability
	N = 199	N = 48	N = 163	N = 178	N = 100	N = 381
Methodological characteristics						
Sample type	Probability	Probability	Purposive	Probability	Purposive	Purposive
Survey mode	Pencil-and-paper, Telephone, Face-to-face	Pencil-and-paper, Telephone, Face-to-face	Face-to-face	Pencil-and-paper, Telephone	Face-to-face	Web
Recruitment	Random selection	Random selection	Street press [b]	Random selection	Websites/forums Street press [b]	Websites/forums Street press [b]
Frequency of ecstasy use	Last 12 months	At least monthly	At least monthly	At least monthly	At least monthly	Last 6 months
Age range in years	17–45	17–45 [c]	17–45 [d]	16–52	16–44	16–54
Locale	Sydney	Sydney	Sydney	Australia	Melbourne	Melbourne
Year of data collection	2001	2001	2001	2004	2004	2004
Demographic characteristics						
Mean Age (SD)	25.5 (6.1)	24.4 (5.4)	24.7 (6.2)	25.0 (0.5) [e]	23.5 (5.2)	23.4 (5.4)
Male (%)	59	62	58	63.5	58.0	66.1
Drugs recently used (%)						
Alcohol	99	97	98	96.4	94.0	71.1
Cannabis	84	89	82	87.3	78.0	61.7
Meth/amphetamine	64	89	87	87.1	75.0	76.9

1

Cocaine	39	59	57	46.2	48.0	26.2
Hallucinogens	18	35	23	27.8	39.0	18.4
Ketamine	NA	NA	NA	23.1	45.0	35.4
GHB	NA	NA	NA	7.8	27.0	19.2
Heroin	2	5	6	3.9	9.0	2.6
Ever injected	NA	NA	NA	24.0	15.0	10.6

[a] Sample names as given by study authors. [b] Newspapers aimed at younger audiences distributed for free, especially around the city and inner city suburbs. Also radio, flyers/posters, interviewer contacts, snowballing. [c] Not stated but sample is a subset of larger household sample. [d] From Topp et al. 2002. [e] Standard error.

Table B

Drug use characteristics of occasional and regular ecstasy users for the online purposive sample matched with the 2007 NDSHS probability samples

Drug use variables	Occasional ecstasy users (less than monthly)				Regular ecstasy users (monthly or more often)			
	Online purposive (N = 328)		NDSHS probability (N = 546) [a]		Online purposive (N = 425)		NDSHS probability (N = 180) [a]	
	n	% (95% CI Boot)	n	% (95% CI Linear)	n	% (95% CI Boot)	n	% (95% CI Linear)
Meth/amphetamine [b, c] (valid n)	321		541		422		179	
Used in the last 6/12 months	183	57.0 (51.6–62.4)	198	36.6 (31.5–41.6)***	290	68.7 (64.1–73.3)	123	68.5 (59.0–78.0)
Monthly or more often	45	14.0 (10.2–17.8)	49	9.0 (6.0–12.1)	164	38.9 (34.1–43.6)	69	38.5 (28.4–48.5)
Cannabis (valid n)	323		546		420		180	
Used in the last 6/12 months	212	65.6 (60.6–70.7)	371	68.0 (63.0–73.0)	327	77.9 (74.0–81.7)	149	83.0 (75.3–90.7)
Monthly or more often	103	31.9 (26.6–37.2)	195	35.8 (30.7–40.9)	197	46.9 (42.0–51.8)	103	57.3 (46.5–68.1)
Heroin (valid n)	309		545		403		180	
Used in the last 6/12 months	8	2.6 (0.8–4.4)	11	2.1 (0.6–3.6)	4	1.0 (0.0–1.9)	10	5.5 (1.4–9.6)*
Monthly or more often	5	1.6 (0.3–3.0)	5	1.0 (0.0–2.2)	1	0.2 (0.0–0.6)	6	3.1 (0.1–6.2)
Cocaine (valid n)	316		546		412		180	
Used in the last 6/12 months	76	24.1 (19.5–28.6)	152	27.8 (23.1–32.5)	148	35.9 (31.3–40.6)	86	47.6 (37.1–58.1)*
Monthly or more often	11	3.5 (1.5–5.5)	16	3.0 (1.2–4.8)	44	10.7 (7.8–13.6)	27	15.3 (8.0–22.5)
Hallucinogens [d] (valid n)	312		546		414		180	
Used in the last 6/12 months	102	32.7 (27.3–38.1)	43	7.9 (5.1–10.7)***	154	37.2 (32.6–41.8)	47	26.2 (17.3–35.1)*
Monthly or more often	21	6.7 (4.1–9.4)	2	0.4 (0.0–0.8)***	52	12.6 (9.3–15.8)	18	10.0 (3.6–16.3)
Ketamine (valid n)	314		543		412		179	
Used in the last 6/12 months	44	14.0 (10.3–17.7)	12	2.2 (0.8–3.6)***	81	19.7 (16.0–23.3)	21	11.8 (4.8–18.9)
Monthly or more often	8	2.5 (0.7–4.4)	0	0	25	6.1 (3.9–8.3)	7	3.9 (0.0–8.1)

3

GHB (valid n)	313		542		410		178	
Used in the last 6/12 months	20	6.4 (3.8–9.0)	4	0.7 (0.0–1.3)***	42	10.2 (7.4–13.0)	16	9.1 (2.3–15.9)
Monthly or more often	10	3.2 (1.3–5.1)	0	0	18	4.4 (2.3–6.4)	8	4.5 (0.0–10.0)
Benzodiazepines [c] (valid n)	314		544		412		180	
Used in the last 6/12 months	90	28.7 (23.7–33.7)	60	11.0 (7.8–14.2)***	96	23.3 (19.3–27.3)	46	25.4 (16.1–34.7)
Monthly or more often	42	13.4 (9.7–17.1)	23	4.2 (2.0–6.3)***	46	11.2 (8.3–14.1)	21	11.4 (4.5–18.4)
Alcohol [e] (valid n)	327		544		425		180	
Monthly or more often [f]	284	86.9 (83.1–90.6)	510	93.6 (90.9–96.4)**	371	87.3 (84.1–90.5)	173	95.8 (92.6–99.1)***
Weekly or more often	206	63.0 (57.6–68.4)	421	77.5 (72.9–82.0)***	296	69.6 (65.2–74.1)	158	87.7 (81.8–93.6)***
Injecting (valid n)	325		546		416		180	
Ever injected	42	12.9 (9.5–16.4)	79	14.5 (11.0–18.0)	30	7.2 (4.9–9.5)	26	14.4 (7.8–21.0)*
Injected in the last 6/12 months	19	5.8 (3.3–8.4)	34	6.2 (3.8–8.7)	14	3.4 (1.7–5.1)	12	6.7 (2.4–11.0)

Note. See Appendix for survey items from which estimates of drug use were derived. * p<0.05; ** p<0.01; *** p<0.001

[a] Household survey estimates are weighted. [b] Defined as meth/amphetamine in the household survey; non-medical use of pharmaceutical stimulants were explicitly included in the online survey. [c] For non-medical reasons only. [d] Defined as hallucinogens in the household survey; LSD, mushrooms and other psychoactive plants were measured separately in the online survey and then combined into a composite hallucinogens variable. [e] Four alcohol response categories from the household data were collapsed into 'weekly or more often' to match the online forum data. [f] Includes weekly or more often.

Appendix: Survey items from which estimates of recent and regular drug use were derived

Online purposive survey

Have you ever used *[drug name]* for non-medical reasons? If yes, how often in the past 6 months?
(a) no, never
(b) yes, but not in the past 6 months
(c) 1 to 5 times in the past 6 months
(d) 1 to 3 times per month in the past 6 months
(e) weekly or more often in the past 6 months

Ever used = b–e
Used in the last 6 months = c–e
Used monthly or more often = d–e
Used weekly or more often = e

Probability survey (NDSHS)

1. Have you ever used *[drug name]*? Yes/No
2. *If yes to 1,* Have you used *[drug name]* in the last 12 months? Yes/No
3. *If yes to 2,* In the last 12 months, how often did you use *[drug name]*?
(a) Every day
(b) Once a week or more
(c) About once a month
(d) Every few months
(e) Once or twice a year
Ever used = Yes to 1, no missing data on subsequent questions
Used in the last 12 months = Yes to 2, no missing data on subsequent question
Used monthly or more often = 3a–c
Used weekly or more often = 3a–b