Title: Application of randomized response techniques for investigating cannabis use by Spanish university students

Short title: RRT for investigating cannabis use

Beatriz Cobo[1], Mª Mar Rueda[1], Francisca López -Torrecillas[2]

Email address of the authors

beacr@ugr.es

mrueda@ugr.es

fcalopez@ugr.es

Correspondence: mrueda@ugr.es, , telephone : +34953241735, fax: +34958243267

[1]Departamento de Estadística e Investigación Operativa and Instituto de Matemáticas (IEMath-GR), Facultad de Ciencias. Campus Universitario de Fuentenueva s/n, Universidad de Granada, 18071 – Granada, España.

[2]Departamento de Personalidad, Evaluación y Tratamiento Psicológico, Facultad de Psicología. Campus Universitario de Cartuja s/n, Universidad de Granada. 18071 – Granada, España

Abstract

Cannabis is the most widely used illicit drug in developed countries, and has a significant impact on mental and physical health in the general population. Although the evaluation of levels of substance use is difficult, a method such as the randomized response technique (RRT), which includes both a personal component and an assurance of confidentiality, provides a combination which can achieve a considerable degree of accuracy. Various RRT surveys have been conducted to measure the prevalence of drug use, but to date no studies have been made of the effectiveness of this approach in surveys with respect to quantitative variables related to drug use.

This paper describes a probabilistic, stratified sample of 1146 university students asking sensitive quantitative questions about cannabis use in Spanish universities, conducted using the RRT.

On comparing the results of the direct question (DQ) survey and those of the randomized response (RR) survey, we find that the number of cannabis cigarettes consumed during the past year (DQ: 3, RR: 17 aproximately), and the number of days when consumption took place (DQ: 1, RR: 7) are much higher with RRT.

The advantages of RRT, reported previously and corroborated in our study, make it a useful method for investigating cannabis use.

Keywords: Cannabis consumption, randomized response techniques, complex surveys.

1. Introduction

Cannabis is the illicit drug that is most commonly used by young adults in Spain. On average, it is consumed by nearly 17% of the EU population aged 15-34 years (EMCDDA, 2015; PNSD, 2013). The prevalence of past 30-day use from 1999 to 2013 for all groups and both sexes was 22.5%. Although young males have historically had a higher prevalence of marijuana use, current results indicate that male-female differences in marijuana use are decreasing (Johnson et al. 2015). Cannabis is often used for its mental and physical effects, such as heightened mood and relaxation, and it has been cited in the medical literature as a potential secondary treatment agent for severe pain, muscle spasticity, anorexia, nausea, sleep disturbances and numerous other conditions (Lamarine, 2012). However, the Spanish National Plan on Drugs (PNSD, 2013) has called for a change in the approach taken to understanding this phenomenon, especially as regards how young people, influenced by subcultural networks, become regular cannabis users. Patterns of consumption should be analysed so that appropriate intervention and prevention programmes can be designed.

Healthcare and social problems related to the use of cannabis have led researchers to investigate screening procedures aimed at detecting persons at risk. Two such procedures, which are now commonly used, are the Cannabis Abuse Screening Test (CAST; Cuenca-Royo et al., 2012), in which response options are based on a 5-point Likert scale (*never, rarely, occasionally, quite often* and *very often*) and the cannabis Severity of Dependence Scale (SDS; Cuenca-Royo et al., 2012), with response options based on a 4-point Likert scale (*never, rarely, often* and *always*). These screening instruments are capable of detecting (probable) cannabis dependence or problematic use and have been used in Spain in surveys for the National Plan on Drugs in schools and among the general population (Cuenca-Royo et al., 2012). ~~CAST obtains high internal consistency (*α=0.73*-0.82), sensitivity (83%-93%) and specificity (66%-85%). At a cutoff score of 2-7, the area under the ROC curve ranges from 0.82-0.93 (Bastiani et al., 2013; Cuenca-Royo et al., 2012; Gyepesi et al., 2014; Legleye et al., 2013). SDS also presents high internal consistency (α=0.74-0.83), sensitivity (59%-86%) and specificity (83%-90%), and in this case at a cutoff score of 2-4 the area under the ROC curve ranges from 0.67-0.88 (Bastiani et al., 2013; Cuenca-Royo et al., 2012; Hides et al., 2007).~~ The application of short screening scales to assess dependence and other problems related to the use of cannabis presents a time and cost-saving means of estimating the overall prevalence of cannabis use and of related negative consequences (Bastiani et al., 2013, Gyepesi et al, 2014, Hides et al., 2007, Legleye et al., 2013). These instruments can also help identify persons at risk, as an initial approach before using more extensive diagnostic instruments. Nevertheless, there is a need to formally evaluate the validity of the data gathered (Piontek et al., 2008). Studies by Harrison (1997) and Ramo et al. (2012) have evaluated the reliability and validity of anonymous studies of cannabis use, but these reports do not examine the other side of validity, namely the fact that respondents may lie, when faced with a question that they find embarrassing, or refuse to answer, or choose a response that prevents them from having to continue and, clearly, this situation may arise in questionnaires related to the use of illegal drugs. Other potential threats to survey accuracy are nonresponse and reporting error (Tourangeau and Yan, 2007).

The aim of the randomized response (RR) technique is to decrease social desirability bias, thus guaranteeing confidentiality, improving respondent cooperation and procuring reliable estimates. This technique obtains stronger estimates of sensitive characteristics, compared to direct questioning (DQ), by reducing respondents' motivation to falsely report their attitudes.

The RR was introduced by Warner in 1965. The procedure is as follows, to estimate for a community the proportion of people bearing a stigmatizing characteristic (denoted by the symbol A), like addiction to marijuana consumption, a sampled person is offered a box of a considerable number of identical cards with a proportion p ($0<p<1, p \neq 0.5$) of them marked A and the rest marked A^c. The person on request is required to draw a "random" card and respond by answering "Yes" for a "match" between the card type and the person's own real characteristic or a "No" for a "nonmatch" before returning the card to the box.

The randomization is performed by the interviewee, and the interviewer is not permitted to observe the outcome of the randomization. The interviewee responds to the question selected by the randomization device, and the interviewer knows only the response provided. The respondent's privacy or anonymity is fully protected because no one but the respondent knows which question was answered. But it is possible statistically to derive a plausible estimate, on analyzing the bunch of randomized responses thus collectively gathered, for the required proportion bearing A. It is hoped that the privacy of the person responding is securely protected.

It is assumed that respondents are more willing to provide honest answers with this technique because their answers do not reveal any information about themselves.

Some studies have addressed situations in which the response to a sensitive question results in a quantitative variable. Thus, Greenberg et al. (1971) extended RR to this case, rather than a simple *Yes* or *No*. Other important work in this respect includes Bar-Lev et al. (2004), Bouza (2009), Diana and Perri (2010, 2012), Eichhorn and Hayre (1983), Saha (2007), Gjestvang and Singh (2006, 2007) and Odumade and Singh (2010). These authors concur that the RRT is an effective means of obtaining estimates with a relatively high degree of reliability. However, most studies concern only simple random sampling, while most of the surveys conducted in practice are complex, involving stratification, clustering and unequal probability in the sample selection.

The RR technique was developed in an attempt to improve the quality of self-reported survey research, but it is not very often applied in the educational or psychological context (Dietz et al., 2013; Goodstadt and Gruson, 1975; Pitsch et al., 2007; Striegel et al., 2009; Weissman et al., 1986). Specifically, Goodstadt and Gruson (1975) compared 854 students' responses concerning drug use, derived from either traditional direct questioning or an indirect, more anonymous method of inquiry, the randomized response procedure. The results obtained in this study showed that the randomized response procedure produced significantly fewer response refusals and significantly higher drug use estimates, thus supporting the hypothesized greater sensitivity and validity of the randomized response procedure. The results further suggested that previous estimates derived from standard questionnaire forms may have underestimated the incidence of drug use. Weissman et al. (1986) examined whether telephone interviewing could be a viable alternative to field interviewing as a method for eliciting drug use information. In this study, a variation of Warner's randomized response technique (RRT) was employed, and the telephone responses obtained with the RRT were compared with those obtained through direct questioning. It was found that in 75% of cases the RRT produced a stronger estimate. Pitsch et al. (2007) used the RRT to examine whether the use of performance-assisting doping was prevalent in certain professional sports. The question posed was, "Have you ever used banned substances or methods in order to enhance your performance?". The authors established a lower interval limit of 25.8% and an upper limit of 48.1%, and reported that at the lower end of the scale 20.4% of athletes admitted to using illegal drugs or methods, while at the upper end, this figure rose to 38.7%. Of the athletes in the study sample, 51.9% had been "honest non-dopers" throughout their career, and for the current year, the corresponding figure was 61.3%. Striegel et al. (2009) estimated the prevalence of doping and illicit drug abuse among athletes. In this study, the subjects were either asked to complete an anonymous standardized questionnaire (SQ; n=1394) or were interviewed using the RRT (n=480). According to this analysis, doping tests produce 0.81% positive test results, but the RRT shows that the real prevalence is 6.8%. In another study, Dietz et al. (2013) calculated the prevalence of students who take drugs in order to improve their cognitive performance, and reported that 20% used cognitive-enhancing drugs. This prevalence varied by sex (male 23.7%, female 17.0%), field of study (highest in students studying sports-related fields, 25.4%), and semester (first semester 24.3%, beyond first semester 16.7%). The authors concluded that the RRT revealed a high 12-month prevalence of cognitive-enhancing drug use by university students and suggested that other direct survey techniques may underestimate the use of these drugs, a fact which should be taken into consideration by universities and in the development of drug prevention programmes. Kerkvliet (1994) combines randomized response techniques with logit models. The academic performance of college students, their personal habits and socioeconomic characteristics are used to estimate a logit model for predicting wheter or not they have consumed cocaine.

Surveys based on the RRT are widely used when the questions are sensitive, and especially when the variable of interest is a qualitative one. Techniques also exist for quantitative variables, but these are not used as commonly. In our study, conducted in Spain (where RR techniques are not generally used for studies of drug consumption), we took into account quantitative variables in order to make the scope of the study as complete as possible.

2. Methods

To investigate cannabis use in the Spanish universities, we conducted a survey of university students.

Participants and Sampling Method

The target population for this survey included students at the University of Granada and the University of Murcia. Subjects were selected using probabilistic sampling stratified by university. Respondents were randomly selected to use the RR technique (subsample 1) and to be asked directly about illicit drug use (subsample 2). We determinate the sample size to estimate the population mean in stratified sampling with a coefficient of variation of 0.25. We used a pilot sample to estimate the unit relvariances.

1146 student participants voluntarily responded to a questionnaire. All questionnaires were administered during the class time break. All students were invited to participate in a study and provided informed consent by signed.

Procedure and Measure

The questionnaire is the same in two subsamples. This questionnarie began with some academic questions followed by a set of basic demographic questions and then a block containing the sensitive questions, referring to drug use (taken from the questionnaire proposed by Miller and Rollnick, 2015).

The following sensitive questions were asked:

- P1: How many cannabis cigarettes did you consume last year?
- P2: Over the past 90 days, how many days did you consume cannabis?

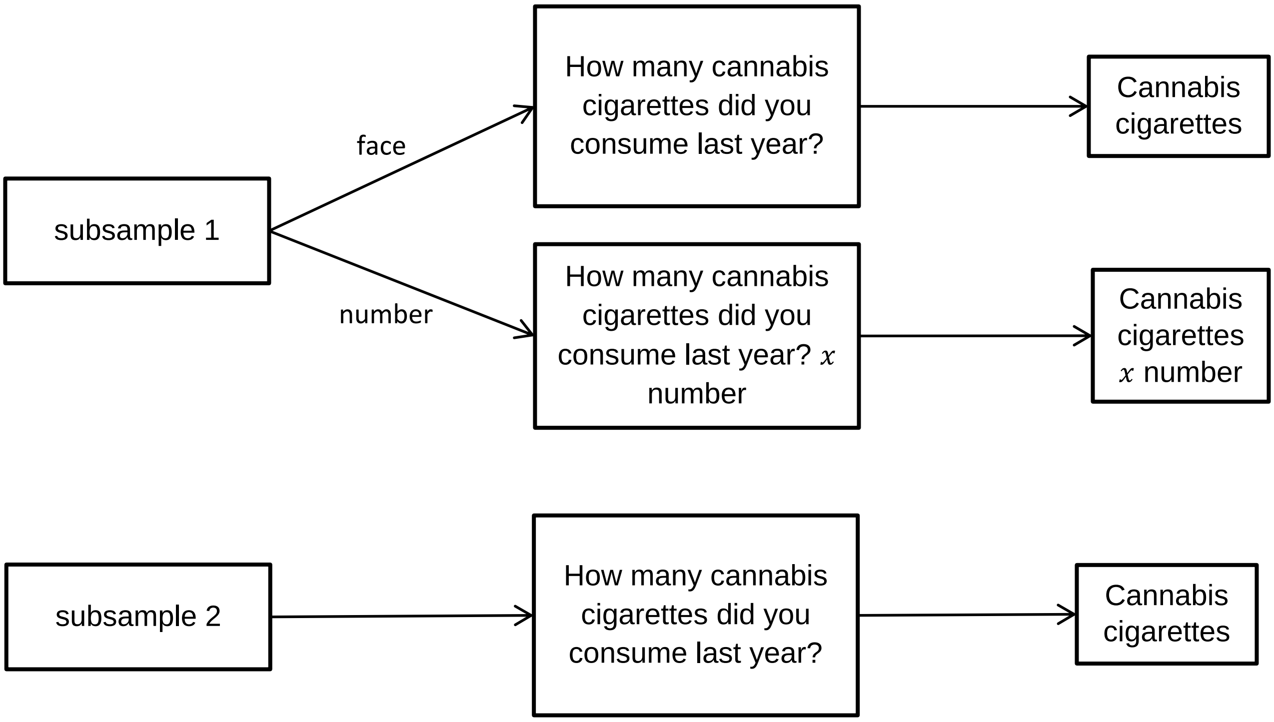
~~We estimated they used when asked directly how many cannabis cigarettes they thought they got from a gram of cannabis: four, or 0.25 grams per cannabis cigarette (Desrosiers et al. 2014).~~

In subsample 1 (using RR technique), for the block of sensitive questions, the interviewer explained how the survey was being conducted, and gave an example of its use. The responses were randomized using a generalization of the model proposed by Bar-Lev et al. (2004) for simple random sampling and later extended by Arcos et al. (2015) for use with complex samples.

The randomizing device used was the app “Baraja Española” (“Baraja Española” is a deck composed of 40 cards, divided into four families or suits, each numbered 1 to 7 and three figures) (Play Store 2015), which had previously been installed on the student’s phone (see figure 1 and 2 in the Annex 1). The application is very simple to use: for each sensitive question the user touches the screen and a card is shown. If it is a face card, the sensitive question (P1, P2) should be answered; otherwise, the real number should be given, multiplied by the number shown on the card

Figure 3 shows the procedure of response for the two subsamples.

Figure 3



The interviewer explained that this technique preserved the students’ anonymity with the aim not to provoke mistrust of them~~. and all students completed the full questionnaire~~.

~~On the contrary, in subsample 2 (using direct question), not all the respondents completed the survey. For the sensitive questions, the nonresponse rate varied from 11-14%. In addition, 4% of the responses made were invalid (very extreme or non-numeric values).~~

The data collection and field work was conducted by the research group FQM365 of the Andalusian Research Plan. The interviews were carried out during 2015, in Spain. Data were obtained from 654 students using the RRT and from 492 using DQ.

Response rates

Table 1 shows the nonresponse rates for the questions for the full sample and for the sample separated by gender considering direct response and randomized response.

	DQ	RR
P1: units	0.10975610	0.05810398
P2: days	0.12601626	0.01834862
men		
P1: units	0.30331754	0.08088235
P2: days	0.32227488	0.03676471
women		
P1: units	0.24199288	0.04188482
P2: days	0.256227758	0.005235602

Table 1: Nonresponse rates

The nonresponse rate for the question was lower in the RR than in the DQ condition. These differences are statistically significatives (p-value<0.001). However, the nonresponse rate between men and women is similar and there is not significant statistically (p-value>0.05).

Statistical analysis

Inference is used in survey sampling to estimate the parameters of interest. The Horvitz-Thompson estimator (Singh, 2003) was used to estimate the mean values for the direct questions.

We use the unified method of estimating population surveys characteristic in RR proposed by Arnab (1994). For each unit in the sample the RR induces a random response (denoted scrambled response) and we can obtain a certain transformation of these scambled responses that is an unbiased estimation of the population mean (see Arnab, 1994 or Arcos et al. 2015 for a detailed description.)

Software

Standard software packages for complex surveys cannot be used directly when the sample is obtained using RR techniques. Although analyses with standard statistical software, with certain modifications in the randomized variables, can yield correct point estimates of population parameters they could still yield incorrect results for the standard errors estimated.

R-packages have been developed for estimation with RR surveys, such as the RRreg package (Heck and Moshagen, 2014) and the rr package (Blair et al., 2015) but the methods implemented in these packages assume simple random sampling. Therefore, we used the package RRTCS (Rueda et al., 2015), which is the only one that incorporates estimation procedures for handling RR data obtained from complex surveys.

In this package, the function `BarLev()` (Annex 2) implements the BarLev model.

Results

The sociodemographic distribution of the samples is shown in Table 2.

	DQ		RR	
	Frequency	***Percentage***	***Frequency***	***Percentage***
Total	492	100%	654	100%
Male	211	42.8861%	272	41.5902%
Female	281	57.1138%	382	58.4098%

Table 2 Sociodemographic distribution of sample

The study was conducted for all students and also separating respondents by gender.

In DQ, the survey had a population of 492 individuals, of whom 42.89% were men and 57.11% were women. In RR, the study population was composed of 654 students, with 41.59% men and 58.41% women.

The point estimates of the sensitive variables and the corresponding 95% confidence intervals for each technique (DQ and RR) are summarized in Table 3.

Study technique	DQ (n=492)				RR (n=654)			
	Estimation	Standard deviation	Confidence Interval (95%)		Estimation	Standard deviation	Confidence Interval (95%)	
			Lower bound	*Upper bound*			*Lower bound*	*Upper bound*
P1: units	3.1142	0.5592	2.0181	4.2103	17.0011	3.6790	9.7903	24.2119
P2: days	0.6837	0.1498	0.3902	0.9773	7.0179	0.9367	5.1819	8.8538
				men				
P1: units	6.3452	1.2522	3.8910	8.7995	24.4972	7.3536	10.0843	38.9100
P2: days	1.2805	0.3220	0.6494	1.9115	8.7713	1.6352	5.5664	11.9763
				women				
P1: units	0.2479	0.1090	0.0341	0.4616	11.6636	3.4850	4.8331	18.4941
P2: days	0.1304	0.06605	0.0010	0.2599	5.7693	1.0999	3.6136	7.9250

Table 3. Estimation of the patterns of cannabis consumption

- *P1: How many cannabis cigarettes did you consume last year?*
- *P2: Over the past 90 days, how many days did you consume cannabis?*

P1: By DQ, the mean number of cannabis cigarettes consumed in the previous year was approximately 3, but according to RR, 17 units were consumed.

P2: By DQ, the students had consumed cannabis on approximately 1 of the previous 90 days, and on 7 according to RR.

The estimate of the number of cannabis cigarettes consumed and the estimate of the number of days when consumption took place for the RR group were significantly higher thant the estimates for the DQ group (p-values <0.001).

For all questions, the standard deviation was higher for the RRT than for DQ.This result is as we expect because surveys conducted with RRT require large sample sizes.

If we consider the results by gender, we get more units of cannabis consumed and more number of days consuming in men than women. This difference is statistically significant by DQ (p-value= $3.8*10^{-5}$ and 0.002 respectively) but this difference is not statisitically significant by RR (p-value 0.105 and 0.108 respectively) because the RR estimates have higher variances.

3. Discussion

~~For clinical and research purposes, the evaluation of substance use is often difficult. Clearly, simply asking people about their drug or alcohol use may not always yield accurate information, because of mistrust, and drug screening raises a series of ethical issues. The use of a method such as randomized response, however, provides a technique that includes both a personal component and an assurance of confidentiality, a combination which potentially fosters accuracy. Standard RR methods are used primarily in surveys which require a binary response to a sensitive question, and seek to estimate the prevalence of people presenting a given (sensitive) characteristic. There is considerable evidence that RR obtains more accurate estimates of the prevalence of socially undesirable behaviour than is the case when sensitive questions are asked directly. In most cases, the use of RR has resulted in an increased reporting of sensitive behaviour, in comparison to the reporting of the same behaviour in response to a direct question. Some studies have addressed situations in which the response to a sensitive question results in a quantitative variable. Thus, Greenberg et al. (1971) extended RR to this case, rather than a simple *Yes* or *No*. In this study, the respondent was asked to select, by means of a randomization device, one of two questions; the sensitive one or an unrelated question, the answers to which were of about the same order of magnitude. Other important work in this respect includes Bar-Lev et al. (2004), Bouza (2009), Diana and Perri (2010, 2012), Eichhorn and Hayre (1983), Saha (2007), Gjestvang and Singh (2006, 2007) and Odumade and Singh (2010). These authors concur that the RRT is an effective means of obtaining estimates with a relatively high degree of reliability. However, most studies concern only simple random sampling, while most of the surveys conducted in practice are complex, involving stratification, clustering and unequal probability in the sample selection. Data from complex survey designs require special consideration with regard to the estimation of finite population parameters and the corresponding variance estimation procedures, as a consequence of significant departures from the simple random sampling assumption. In such a complex survey design, unbiased variance estimation is not easy to calculate because of clustering and due to the involvement of second-order inclusion probabilities, which are generally complex.~~

In this paper, we present a survey related to the use of cannabis, in which a RRT is used to determine population means that are valid for any sampling design. On comparing the results of the direct survey and those of the randomized response survey we find that the number of cannabis cigarettes consumed during the past year (DQ: 3, RR: 17 aproximately), and the number of days when consumption took place (DQ: 1, RR: 7 approximately) are much higher with RRT in these universities.

These results are in line with those reported by Dietz et al. (2013), Goodstadt and Gruson (1975), Pitsch et al. (2007) and Striegel et al. (2009). All of these authors conclude that the RRT is an effective means of obtaining estimates with a relatively high degree of reliability. In the case of doping among professional athletes, this approach could be a promising means of evaluating the effectiveness of anti-doping programmes. The RRT has also highlighted the existence of high values for the 12-month prevalence of cognitive-enhancing drug use among university students, which suggests that other direct survey techniques underestimate this kind of drug use.

The results obtained suggest that estimates derived from standard questionnaire forms underestimate the incidence of drug use by university students. We believe that the advantages of randomized response revealed in this study and elsewhere make it a useful method to investigate sensitive behaviour related to alcohol and drug use. It must be stressed, however, that randomized response has wide confidence intervals. The randomization procedure introduces additional random error into the data and increases the standard errors of the parameters estimated: thus, larger sample sizes are needed in order to increase the statistical power. Another important issue in RRT is the choice of an appropriate randomizing device, which should be implemented in such a way as to make the confidentiality protection offered very clear to the respondent. The randomizing devices most commonly employed to date have been serial numbers on a banknote, the flip of a coin, a spinner, playing cards, numbers selected from the phone book or the respondent’s month of birth. However, the new technologies currently available offer alternatives that are more attractive to users, such as mobile phones. Thanks to smartphones, we have access to many interesting applications that can help in the randomization of telephone and personal surveys (Rueda et al., 2016), especially among young people.

Acknowledgements*:* This study was partially supported by Ministerio de Educación, Cultura y Deporte (grant MTM2015-63609-R and FPU grant program, Spain) and by Consejería de Economía, Innovación, Ciencia y Empleo (grant SEJ2954, Junta de Andalucía, Spain).

This study was partially supported by Ministerio de Educación, Cultura y Deporte (FPU grant program, Spain), by Ministerio de Economía y Competitividad (MINECO) and Fondo Europeo de Desarrollo Regional (FEDER) (grant MTM2015-63609-R) and by Consejería de Economía, Innovación, Ciencia y Empleo (grant SEJ2954, Junta de Andalucía, Spain).

Declaration of interests' statement:

The authors have no competing interests.

References

Arcos, A., Rueda, M., Singh, S. A., 2015. Generalized approach to randomized response for quantitative variables. *Qual Quant*, **49**, 1239-1256. doi:10.1007/s11135-014-0046-3

Arnab, R. (1994). Non-negative variance estimator in randomized response surveys. *Comm. Stat. Theo. Math.* **23**, 1743-1752.

Bar-Lev S.K., Bobovitch, E., Boukai, B., 2004. A note on randomized response models for quantitative data. *Metrika*, **60**, 255-260. *doi*: 10.1007/s001840300308.

Bastiani, L., Siciliano, V., Curzio, O., Luppi, C., Gori, M., Grassi, M., Molinaro, S., 2013. Optimal scaling of the CAST and of SDS Scale in a national sample of adolescents. *Addict Behav*. **38**, 2060-2067. doi: 10.1016/j.addbeh.2012.12.016.

~~Bhavsar, V., 2015. Environmental factors, including cannabis, are strongly related to the age of onset and morbidity of schizophrenia. *Evid Based Ment Health*. **18**, 84-93. doi: 10.1136/eb-2014-102040. 25979516.~~

Blair, G., Imai, K., Zhou, Y.Y., 2015. Package 'rr': Statistical methods for the randomized response technique. http://cran.r-project.org/web/packages/rr.

Bouza, C., 2009. Ranked set sampling and randomized response procedures for estimating the mean of a sensitive quantitative character. *Metrika.* **70**, 267-277. doi: 10.1007/s00184-008-0191-6

Cuenca-Royo, A.M., Sánchez-Niubó, A., Forero, C.G., Torrens, M., Suelves, J.M., Domingo-Salvany, A., 2012. Psychometric properties of the CAST and SDS scales in young adult cannabis users. *Addict Behav*. **37,** 709-715. doi: 10.1016/j.addbeh.2012.02.012.

~~Desrosiers, N. A., Himes, S. K., Scheidweiler, K. B., Concheiro-Guisan, M., Gorelick, D. A., & Huestis, M. A. (2014). Phase I and II cannabinoid disposition in blood and plasma of occasional and frequent smokers following controlled smoked cannabis. *Clinical Chemistry*, **60**(4),631-43.Retrieved from~~

~~http://search.proquest.com/docview/1529962294?accountid=14542~~

Diana, G., Perri, P.F., 2010. New scrambled response models for estimating the mean of a sensitive quantitative character. *J Appl Stat.* **37,** 1875–1890. doi: 10.1080/02664760903186031.

Diana, G., Perri, P.F., 2012. A calibration-based approach to sensitive data: a simulation study. *J Appl Stat*. **39**, 53-65. http://dx.*doi*.org/10.1080/02664763.2011.578615

Dietz, P., Striegel, H., Franke, A.G., Lieb, K., Simon, P., Ulrich, R., 2013. Randomized response estimates for the 12-month prevalence of cognitive-enhancing drug use in university students. *Pharmacotherapy.* **33,** 44-50. doi: 10.1002/phar.1166.

~~Donoghue, K., Doody, G.A., Murray, R.M., Jones, P.B., Morgan, C., Dazzan, P., Hart, J., Mazzoncini, R., Maccabe, J.H., 2014. Cannabis use, gender and age of onset of schizophrenia: data from the ÆSOP study. *Psychiatry Res*. **30**, 528-532. doi: 10.1016/j.psychres.2013.12.038.~~

Eichhorn, B.H., Hayre, L.S., 1983. Scrambled randomized response methods for obtaining sensitive quantitative data. *J Stat Plan Inference*. **7,** 306-316. *doi:*10.1016/0378-3758(83)90002-2

EMCDDA. 2015. European Monitoring Centre for Drugs and Drug Addiction Annual Report 2015. *The State of the Drug Problem in Europe*. Luxembourg, Publications Office of the European Union. http://www.emcdda.europa.eu/publications/edr/trends-developments/2015

~~Fernandez-Artamendi, S., Fernández-Hermida, J.R., Muñiz-Fernández, J., Secades-Villa, R., García-Fernández, G., 2012. Screening of cannabis-related problems among youth: the CPQ-A-S and CAST questionnaires. *Subst Abuse Treat Prev Policy*. **2**, 13-22. doi: 10.1186/1747-597X-7-13.~~

Gjestvang, C.R., Singh, S., 2006. A new randomized response model. *J. Royal Statist. Soc Ser*. **68**,523-530. doi: 10.1111/*j*.1467-9868.2006.00554.x

Gjestvang, C.R., Singh, S., 2007. Forced quantitative randomized response model: a new device. *Metrika*. **66**, 243-257. doi: 10.1007/s00184-006-0108-1;

Goodstadt, M.S., Gruson, V., 1975. The randomized response technique: a test on drug use. *JASA. **70,*** 814- 818. http://www.jstor.org/stable/2285441.

~~Greenberg, B.G., Abul-Ela, A.-L., Simmons, W.R., Horvitz, D.G., 1969. The unrelated question RR model: Theoretical framework. *J. Amer. Statist. Assoc*. **64**, 520-539.~~

Greenberg, B.G., Kuebler, R.R., Abernathy, J.R., Horvitz, D.G., 1971. Application of the randomized response technique in obtaining quantitative data. *J Am Stat Assoc*. **66**, 243–250.

Gyepesi, A, Urbán, R., Farkas, J., Kraus, L., Piontek, D., Paksi, B., Horváth, G., Magi, A., Eisinger, A., Pilling, J., Kökönyei, G., Kun, B., Demetrovics, Z., 2014. Psychometric properties of the Cannabis Abuse Screening Test in Hungarian samples of adolescents and young adults. *Eur Addict Res*. **20,** 119-128. doi: 10.1159/000353238.

Harrison, L. 1997. The validity of self-reported drug use in survey research: an overview and critique of research methods. *NIDA Res Monogr*. **167**,17-36.

Heck, D.W., Moshagen, M., 2014. Package 'RReg': Correlation and Regression Analyses for Randomized Response Data. https://cran.r-project.org/web/packages/RRreg.

Hides, L., Dawe, S., Young, R.M., Kavanagh, D.J., 2007. The reliability and validity of the Severity of Dependence Scale for detecting cannabis dependence in psychosis. *Addiction*. **102**, 35-40. doi: 10.1111/j.1360-0443.2006.01669.x

~~Horvitz, D.G., Shah, B.V., Simmons, W.R., 1967. The unrelated question RR model. *Proc. Social Statist. Sec. ASA*. 65-72.~~

Johnson, R.M., Fairman, B., Gilreath, T., Xuan, Z., Rothman, E.F., Parnham, T., Furr-Holden, C.D. 2015. Past 15-year trends in adolescent marijuana use: Differences by race/ethnicity and sex. *Drug Alcohol Depend*. **1**,155-158. doi: 10.1016/j.drugalcdep.2015.08.025.

Kerkvliet, J., 1994. Estimating a logit model with randomized data: The case of cocaine use. *Australian Journal of Statistics* **36**, 9-20.

Lamarine, R.J., 2012. Marijuana: modern medical chimaera. *J Drug Educ*. **42**, 1-11. doi*:* 10.2190/DE.42.1.a.

Legleye, S., Piontek, D., Kraus, L., Morand, E., Falissard, B.A., 2013. Validation of the Cannabis Abuse Screening Test (CAST) using a latent class analysis of the DSM-IV among adolescents. *Int J Methods Psychiatr Res*. **22**,16-26. doi: 10.1002/mpr.1378.

~~Martínez-Loredo, V., Fernández-Hermida, J.R., Fernández-Artamendi, S., Carballo, J.L., García-Cueto, E., García-Rodríguez, O., 2015. The association of both self-reported and behavioral impulsivity with the annual prevalence of substance use among early adolescents. *Subst Abuse Treat Prev Policy*. **10**, 23-31. doi: 10.1186/s13011-015-0019-0.~~

Miller, W.R., Rollnick, S., 2015. *Motivational Interviewing: To prepare for change of addictive behaviors.* 3rd ed. The Guilford Press, London.

Odumade, O., Singh, S., 2010. An alternative to the Bar-Lev, Bobovitch and Boukai Randomized Response Model. *Socio. Meth. Res*. **39**, 206-221. doi: 10.1177/0049124110378094.

Pitsch, W., Emrich, E., Klein, M., 2007. Doping in elite sports in Germany: results of a www survey. *EJSS*. **4,** 89-102. doi:10.1080/15298868.2015.1074101.

Piontek, D., Kraus, L., Klempova, D., 2008. Short scales to assess cannabis-related problems: a review of psychometric properties. *Subst Abuse Treat Prev Policy*. **2**, 25- 35. doi: 10.1186/1747-597X-3-25.

Play Store, 2015. Download the application "Baraja Española". https://play.google.com/store/apps/details?id=cartas.barajaes

~~Play Story, 2015. Download the application "Randomizers". https://play.google.com/store/apps/details?id=org.random.randomapp~~

PNSD. 2013. Plan Nacional Sobre Drogas. *Consumo problemático de cannabis en estudiantes españoles de 14-18 años: validación de escalas. Ministerio de Sanidad y Política Social.*

http://www.pnsd.msssi.gob.es/profesionales/publicaciones/catalogo/bibliotecaDigital/publicaciones/pdf/ConsProblematico_cannabis.pdf

Ramo, D.E., Liu, H., Prochaska, J.J., 2012. Reliability and validity of young adults' anonymous online reports of marijuana use and thoughts about use. *Psychol Addict Behav*. **26**, 801-811. doi: 10.1037/a0026201.

Rueda, M., Cobo, B., Arcos, A., 2015. Package 'RRTCS': Randomized Response Techniques for Complex Surveys. http://cran.r-project.org/web/packages/RRTCS/

Rueda, M., Cobo, B., Arcos, A., Arnab, R., 2016. *Software for randomized response techniques, in Handbook of Statistics: Data Gathering, Analysis and Protection of Privacy through Randomized Response Techniques - Qualitative And Quantitative Human Traits.* Elsevier. In press.

Saha, A.A., 2007. Simple randomized response technique in complex surveys. *Metron LXV*. **1**, 59-66. doi:10.1080/02664763.2013.846305

Singh, S., 2003. *Advanced Sampling Theory with Applications. How Michael 'selected' Amy*. Springer, Netherlands.

Striegel, H., Ulrich, R., Simon, P., 2009. Randomized response estimates for doping and illicit drug use in elite athletes. *Drug Alcohol Depend*. **15**, 230-232. doi: 10.1016/j.drugalcdep.2009.07.026.

Tourangeau, R., Yan, T., 2007. Sensitive questions in surveys. *Psy Bul*. **133,** 859-863. doi: 10.1037/0033-2909.133.5.859.

Warner, S.L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *JASA*, **60**, 63-69.

Weissman, A.N., Steer, R.A., Lipton, D.S., 1986. Estimating illicit drug use through telephone interviews and the randomized response technique. *Drug Alcohol Depend*. **18**, 225-233.

White, J., Walton, D., Walker, N., 2015. Exploring comorbid use of marijuana, tobacco, and alcohol among 14 to 15-year-olds: findings from a national survey on adolescent substance use. *BMC Public Health*. **10**, 233-242. doi: 10.1186/s12889-015-1585-9.

ANNEX 1: "Baraja Española" app

Figure 2

Figure 3

ANNEX 2: Description of use of BarLev function in R

```
BarLev(z,p,mu,sigma,pi,type=c("total","mean"),cl,N=NULL,pij=NULL)
```

`z` vector of the observed variable; its length is equal to n (the sample size)

`p` probability of direct response

`mu` mean of the scramble variable S

`sigma` standard deviation of the scramble variable S

`pi` vector of the first-order inclusion probabilities

`type` the estimator type: total or mean

`cl` confidence level

`N` size of the population. By default it is NULL

`pij` matrix of the second-order inclusion probabilities. By default it is NULL

Harrison L. The validity of self-reported drug use in survey research: An overview and critique of research methods. In: The validity of self-reported drug use: improving the accuracy of survey estimates. NIDA Research Monograph No. 167. Washington, D.C.: U.S. Government Printing Office.