STATISTICS IN TRANSITION new series, December 2020
Vol. 21, No. 5, pp. 119–132, DOI 10.21307/stattrans-2020-058
Received – 24.03.2019; accepted – 17.06.2020

Unbiased estimator modeling in unrelated dichotomous randomized response

Adetola Adedamola Adediran[1], Femi Barnabas Adebola[2], Olusegun Sunday Ewemooje[3]

ABSTRACT

The unrelated design has been shown to improve the efficiency of a randomized response method and reduces respondents' suspicion. In the light of this, the paper proposes a new Unrelated Randomized Response Model constructed by incorporating an unrelated question into the alternative unbiased estimator in the dichotomous randomized response model proposed by Ewemooje in 2019. An unbiased estimate and variance of the model are thus obtained. The variance of the proposed model decreases as the proportion of the sensitive attribute π_A and the unrelated attribute π_U increases, in contrast to the earlier Ewemooje model, whose variance increases as the proportion of the sensitive attribute increases. The relative efficiency of the proposed model over the earlier Ewemooje model decreases as π_U increases when $0.1 \leq \pi\_A \leq 0.3$ and increases as π_U increases when $0.35 \leq \pi\_A \leq 0.45$. Application of the proposed model also revealed its efficiency over the direct method in estimating the prevalence of examination malpractices among university students;the direct method gave an estimate of 19.0%, compared to the proposed method's estimate of 23.0%. Hence, the proposed model is more efficient than the direct method and the earlier Ewemooje model as the proportion of people belonging to the sensitive attribute increases.

Key words: dichotomous, relative efficiency, sensitive attribute.

1. Introduction

One of the problems in a survey is non-response; this is referred to as failure of getting the required information from a respondent. Non-response reduces the sample size as some respondents do not give the needed information and thereby making the

[1] Department of Statistics, Federal University of Technology Akure, Nigeria. E-mail: aaadediran@futa.edu.ng. ORCID: https://orcid.org/0000-0003-3176-7872.
[2] Department of Statistics, Federal University of Technology Akure, Nigeria. E-mail: fbadebola@futa.edu.ng. ORCID: https://orcid.org/0000-0001-7790-1331.
[3] Department of Statistics, Federal University of Technology Akure, Nigeria. E-mail: osewemooje@futa.edu.ng. ORCID: https://orcid.org/0000-0003-3236-6018.

accuracy of the estimate to be compromised. Obtaining information about sensitive attributes lead to non-response or false response as participants in the sample may give false response or decide not to give an answer for diverse reasons. In order to reduce error due to this non-response bias, Warner in 1965 developed the Randomized Response Model (RRM) for estimating the proportion of people that belong to a sensitive attribute.

Quite a number of authors have reviewed and expanded the work of Warner, including Horvitz *et al.* (1967) Unrelated Question Design, Greenberg *et al.* (1969) Unrelated Question Design with known distribution, Mangat and Singh (1990) Randomized Response Model (RRM), Hussain-Shabbir (2007) Dichotomous Randomized Response Model (DRRM), Adebola and Adepetun (2011), Tripartite Randomized Response Model (TRRM), Ewemooje (2017) Equal Probabilities of Protection, Adebola *et al.* (2017) Hybrid Tripartite Randomized Response Technique, Ewemooje *et al.* (2019a) Dichotomous Randomized Response Technique, Ewemooje *et al.* (2018) Stratified Hybrid Tripartite Randomized Response Technique. Also, Yu *et al.* (2008) worked on the Crosswise Model (CM) and Triangular Model (TM) while Fox *et al.* (2019) proposed Generalized Linear Mixed Models for Randomized Responses (GLMRR), among others.

To test the applicability of the RRM; Jann *et al.* (2012) applied a modified RRM (Crosswise Model by Yu *et al.,* 2008) to elicit information on plagiarism among German and Swiss students. They found out that RRM elicited more socially undesirable answers than direct questioning. Ewemooje *et al.,* (2017) also used Improved Randomized Response Technique for two sensitive attributes (IRRT2) to show that RRM performs better that Direct Method of questioning (DM) by estimating prevalence of induced abortion and multiple sexual partners. Cobo *et al.,* (2016) used RRM to investigate cannabis use by Spanish University students and then compared the result with DM. Their results revealed that RRM increases the response rate for cannabis use and that it is an efficient method. Furthermore, Ewemooje *et al.,* (2019b) measured substance use disorder prevalence using RRM and DM; their findings showed that RRM estimated the disorder better with lower error than DM. Conversely, Hoglinger and Jann (2018) evaluated the variability of several variants of RRM and the crosswise model by comparing the respondents' self-reports on cheating in dice games to actual cheating behaviour; their result showed that the RRM fails to reduce the level of misinterpreting compared to DM and none of the RRMs evaluated outperformed the conventional DM.

Therefore, in this work we consider dichotomous randomized response design in the presence of unrelated questions; the estimator and variance are obtained and compared with the Dichotomous Randomized Response Model by Ewemooje *et al.,* (2019a) using relative efficiency. Also, to verify more-is-better assumption, the proposed method and Direct Method (DM) were applied to the same subpopulation in a survey.

2. Dichotomous Randomized Response Model by Ewemooje *et al.,* (2019a)

In their model, respondents were asked sensitive question directly, if he/she responds "yes" then he/she is not allowed to use the randomized device while if "no", he/she is required to use the randomized device. Two randomized devices were used each consisting of two questions with different selection probabilities. A simple random sample with replacement sampling was adopted in their selection of the sample of size, n with α and β as any two positive real numbers such that $q = \frac{\alpha}{\alpha+\beta}$ is the probability of using the first randomized device and $1 - q = \frac{\beta}{\alpha+\beta}$ is the probability of using the second randomized device.

If all respond truthfully, their population proportion of "yes" answers is given by:

$$P(yes) = \theta_1 = \pi + \frac{\alpha}{\alpha+\beta}(1-P_1)(1-\pi) + \frac{\alpha}{\alpha+\beta}(1-P_2)(1-\pi) \tag{1}$$

where P_1 is the probability of the sensitive attribute in randomized devices R_1 and P_2 is the probability of the sensitive attribute in randomized devices R_2.

This yielded an unbiased estimate of the population proportion as:

$$\hat{\pi} = \frac{\hat{\theta}_1(\alpha+\beta) - P_2\alpha - P_1\beta}{P_1\alpha + P_2\beta} \tag{2}$$

The variance of their estimate was given as

$$V(\hat{\pi}) = \frac{\pi(1-\pi)}{n} + \frac{(1-\pi)(P_2\alpha+P_1\beta)}{n(P_1\alpha+P_2\beta)^2} \tag{3}$$

3. Proposed Model

In sampling a finite population, the simple random sample with replacement was used to obtain the sample size of respondents who respond to sensitive questions using Randomized Response Model. Sensitive question was asked directly from the respondents. If "yes" answer is obtained, he/she does not need to use the randomized device but if he/she answers "no", then he/she uses the randomized device. The two randomized devices R_1 and R_2 consists of two unrelated questions (the sensitive question A in which the interviewer is interested in with probability P, and non-sensitive attribute question B that is unrelated to the sensitive question A with probability, 1-P) each. Say:

Sensitive question: "do you belong to a sensitive attribute A?"

Non-sensitive question: "do you love soccer?"

Two responses were considered for each of the two unrelated questions: "yes" and "no", where α and β are positive real numbers such that $q = \frac{\alpha}{\alpha+\beta}, \alpha \neq \beta$ is the probability of using R_1 and $1 - q = \frac{\beta}{\alpha+\beta}, \alpha \neq \beta$ is the probability of using R_2 with preset probabilities P_1 and P_2 respectively for each of the devices.

Let π_A be the true proportion of people that belongs to the sensitive attribute and π_u, the proportion of people that belongs to the unrelated non-sensitive attribute. If all respond truthfully as the devices provide protection for respondents, the population proportion of "yes" answers is given by:

$$P(yes) = \theta = \pi_A + \frac{\alpha}{\alpha+\beta}[P_1\pi_A + (1-P_1)\pi_U] + \frac{\beta}{\alpha+\beta}[P_2\pi_A + (1-P_2)\pi_U] \quad (4)$$

where P_1 is the probability of the sensitive attribute in randomized devices R_1 while P_2 is the probability of the sensitive attribute in randomized devices R_2.

Solving equation (4) further yield the estimate of the population proportion of the sensitive attribute

$$\hat{\pi}_A = \frac{\hat{\theta}(\alpha+\beta) - \pi_U((\alpha+\beta) - \alpha P_1 - \beta P_2)}{(\alpha+\beta + \alpha P_1 + \beta P_2)} \quad (5)$$

where $\hat{\theta} = {}^{n_0}/_{n}$, n_0 is the number of respondents that answered "yes" to sensitive question while n is the sample size.

The proposed estimator, $\hat{\pi}_A$, is an unbiased estimator of the population parameter π_A.

3.1. Variance Estimation

The variance of the model is obtained as follows:

$$v(\hat{\pi}_A) = v\left(\frac{\hat{\theta}(\alpha+\beta) - \pi_U((\alpha+\beta) - \alpha P_1 - \beta P_2)}{(\alpha+\beta + \alpha P_1 + \beta P_2)}\right)$$

$$v(\hat{\pi}_A) = \frac{(\alpha+\beta)^2\, v(\hat{\theta})}{(\alpha+\beta + \alpha P_1 + \beta P_2)^2} \quad (6)$$

where $v(\hat{\theta}) = \frac{\theta(1-\theta)}{n}$

recall that $\theta = \left(\frac{\alpha\pi_A + \beta\pi_A + \alpha P_1\pi_A + \beta P_2\pi_A + \alpha(1-P_1)\pi_U + \beta(1-P_2)\pi_U}{\alpha+\beta}\right)$, substituting this in equation (6), the variance of the proposed unbiased estimator is given as:

$$v(\hat{\pi}_A) = \frac{\pi_A\{(\alpha+\beta) - \pi_A(\alpha+\beta + \alpha P_1 + \beta P_2)\}}{n(\alpha+\beta + \alpha P_1 + \beta P_2)} + \frac{\pi_U(\alpha+\beta - \alpha P_1 - \beta P_2)(\alpha+\beta - 2\pi_A(\alpha+\beta + \alpha P_1 + \beta P_2))}{n(\alpha+\beta + \alpha P_1 + \beta P_2)^2} \quad (7)$$

Therefore, the variance of the proposed unbiased estimator can be estimated using:

$$\hat{v}(\hat{\pi}_A) = \frac{\hat{\pi}_A\{(\alpha+\beta) - \hat{\pi}_A(\alpha+\beta + \alpha P_1 + \beta P_2)\}}{(n-1)(\alpha+\beta + \alpha P_1 + \beta P_2)} + \frac{\pi_U(\alpha+\beta - \alpha P_1 - \beta P_2)(\alpha+\beta - 2\hat{\pi}_A(\alpha+\beta + \alpha P_1 + \beta P_2))}{(n-1)(\alpha+\beta + \alpha P_1 + \beta P_2)^2} \quad (8)$$

4. Efficiency Comparison

The proposed model will be more efficient than the conventional one if the condition for the relative efficiency holds:

$$\text{RE} = \frac{\text{variance of conventional model}}{\text{variance of proposed model}} > 1$$

The relative efficiency of the proposed model over the conventional model were gotten for varying sample sizes (n), varying probabilities P_1 and P_2 of using the randomized devices at different values of π_A and π_U.

The comparison between the proposed estimator and Ewemooje *et al.* (2019a) estimator at different sample sizes in Table 1 shows that the proposed estimator is approximately ten (10) times more efficient than that due to Ewemooje *et al.* (2019a). As the sample size increases from 50 to 500, the variances due to Ewemooje *et al.* (2019a) estimator reduces from 0.0053 to 0.0005 while the proposed estimator reduces from 0.0005 to 0.0001. Therefore, as the sample sizes increases the variability reduces, this implies consistency of the two models.

Considering a constant sample size at varying probabilities of selecting the randomized device, the variances due to Ewemooje *et al.* (2019a) estimator increases from 0.00131 to 0.00138, the proposed estimator increases from 0.00018 to 0.00022 while the relative efficiency reduces from 7.089 to 6.227 as shown in Table 2.

Table 1. Relative efficiency comparison between the proposed model and Ewemooje *et al.* (2019a) model when π_A = 0.5; π_U= 0.5; P_1= 0.5; P_2= 0.5; α= 25; β= 35 for varying sample sizes (n).

n	π_A	π_U	P_1	P_2	α	β	$v(\hat{\pi})$	$v(\hat{\pi}_A)$	RE
50	0.5	0.5	0.5	0.5	25	35	0.005333	0.000537	9.931034
100	0.5	0.5	0.5	0.5	25	35	0.002667	0.000269	9.931034
150	0.5	0.5	0.5	0.5	25	35	0.001778	0.000179	9.931034
200	0.5	0.5	0.5	0.5	25	35	0.001333	0.000134	9.931034
250	0.5	0.5	0.5	0.5	25	35	0.001067	0.000107	9.931034
300	0.5	0.5	0.5	0.5	25	35	0.000889	0.0000895	9.931034
350	0.5	0.5	0.5	0.5	25	35	0.000762	0.0000767	9.931034
400	0.5	0.5	0.5	0.5	25	35	0.000667	0.0000671	9.931034
450	0.5	0.5	0.5	0.5	25	35	0.000593	0.0000597	9.931034
500	0.5	0.5	0.5	0.5	25	35	0.000533	0.0000537	9.931034

Table 2. Relative efficiency comparison between the proposed model and Ewemooje *et al.* (2019a) model when π_A = 0.5; π_U= 0.5; α= 25; β= 35; n=200 for varying P_1 and P_2

n	π_A	π_U	P_1	P_2	α	β	$v(\hat{\pi})$	$v(\hat{\pi}_A)$	RE
200	0.5	0.5	0.1	0.9	25	35	0.001306	0.000184	7.089034
200	0.5	0.5	0.2	0.8	25	35	0.001312	0.000188	6.966797
200	0.5	0.5	0.3	0.7	25	35	0.001318	0.000193	6.848074
200	0.5	0.5	0.4	0.6	25	35	0.001325	0.000197	6.733115
200	0.5	0.5	0.5	0.5	25	35	0.001333	0.000201	6.622212
200	0.5	0.5	0.6	0.4	25	35	0.001342	0.000206	6.515705
200	0.5	0.5	0.7	0.3	25	35	0.001352	0.000211	6.413994
200	0.5	0.5	0.8	0.2	25	35	0.001363	0.000216	6.317551
200	0.5	0.5	0.9	0.1	25	35	0.001376	0.000221	6.226937

Table 3 shows that for varying π_A and π_U, P_1= 0.3; P_2= 0.7, the variance of the Ewemooje *et al.* (2019a) model increases at all values of π_A while the variance of the proposed model increases as π_U increases when $0.1 \leq \pi_A \leq 0.3$ and decreases as π_U increases when $0.35 \leq \pi_A \leq 0.45$. The relative efficiency of the proposed model over Ewemooje *et al.* (2019a) reduces as π_U increases when $0.1 \leq \pi_A \leq 0.3$ and increases as π_U increases when $0.35 \leq \pi_A \leq 0.45$. However, as the sensitive character, π_A increases, the relative efficiency increases with the values ranging from 1.0135 to 21.4409. The relative efficiency (RE) is greater than 1 for $\pi_A = 0.1$ when $0.1 \leq \pi_U \leq 0.4$, RE greater than 1 for $\pi_A = 0.15$ when $0.1 \leq \pi_U \leq 0.7$ and RE greater than 1 when $0.2 \leq \pi_A \leq 0.45$ at all values of π_U. This shows that the proposed model is more efficient than the Ewemooje *et al.* (2019a) model as the proportion of people belonging to the sensitive attribute increases.

In Table 4, the probability of selecting the sensitive attribute was increased to 0.4 i.e. P_1= 0.4 while P_2= 0.6. The relative efficiency of the proposed model over Ewemooje *et al.* (2019a) also reduces as π_U increases when $0.1 \leq \pi_A \leq 0.3$ and increases as π_U increases when $0.35 \leq \pi_A \leq 0.45$. The relative efficiencies range between 1. 0284 and 18.8538. This shows that there is increase in efficiency as P_1 increases.

Table 3. Relative efficiency comparison between the proposed model and Ewemooje *et al.* (2019a) model when $P_1 = 0.3$; $P_2 = 0.7$; $\alpha = 25$; $\beta = 35$; n=200 for varying π_A and π_U.

π_A	π_U	$v(\hat{\pi})$	$v(\hat{\pi}_A)$	RE	π_A	π_U	$v(\hat{\pi})$	$v(\hat{\pi}_A)$	RE
	0.1	0.000573	0.000345	1.662303		0.1	0.001146	0.000536	2.137366
	0.2	0.000573	0.000413	1.387377		0.2	0.001146	0.000543	2.108094
	0.3	0.000573	0.000481	1.191302		0.3	0.001146	0.000551	2.080862
	0.4	0.000573	0.000549	1.044416		0.4	0.001146	0.000557	2.055543
0.1	0.5	0.000573	0.000616	0.930275	0.3	0.5	0.001146	0.000564	2.032025
	0.6	0.000573	0.000683	0.839031		0.6	0.001146	0.00057	2.010205
	0.7	0.000573	0.00075	0.764424		0.7	0.001146	0.000576	1.989992
	0.8	0.000573	0.000816	0.702286		0.8	0.001146	0.000581	1.971302
	0.9	0.000573	0.000882	0.649735		0.9	0.001146	0.000586	1.954063
	1	0.000573	0.000948	0.604711		1	0.001146	0.000591	1.938206
	0.1	0.000754	0.00043	1.752585		0.1	0.001226	0.000521	2.352242
	0.2	0.000754	0.000483	1.559987		0.2	0.001226	0.000514	2.387847
	0.3	0.000754	0.000536	1.406395		0.3	0.001226	0.000505	2.426134
	0.4	0.000754	0.000588	1.281057		0.4	0.001226	0.000497	2.46731
0.15	0.5	0.000754	0.00064	1.176839	0.35	0.5	0.001226	0.000488	2.511608
	0.6	0.000754	0.000692	1.088822		0.6	0.001226	0.000479	2.559291
	0.7	0.000754	0.000744	1.013506		0.7	0.001226	0.00047	2.610657
	0.8	0.000754	0.000795	0.948329		0.8	0.001226	0.00046	2.666043
	0.9	0.000754	0.000846	0.891377		0.9	0.001226	0.00045	2.725833
	1	0.000754	0.000896	0.841189		1	0.001226	0.000439	2.790465
	0.1	0.000909	0.00049	1.854419		0.1	0.001282	0.000482	2.661543
	0.2	0.000909	0.000528	1.721451		0.2	0.001282	0.000459	2.79495
	0.3	0.000909	0.000566	1.607214		0.3	0.001282	0.000435	2.944671
	0.4	0.000909	0.000603	1.508023		0.4	0.001282	0.000412	3.113841
0.2	0.5	0.000909	0.00064	1.421098	0.4	0.5	0.001282	0.000388	3.306451
	0.6	0.000909	0.000676	1.344305		0.6	0.001282	0.000363	3.52767
	0.7	0.000909	0.000713	1.275977		0.7	0.001282	0.000339	3.784304
	0.8	0.000909	0.000749	1.214796		0.8	0.001282	0.000314	4.085509
	0.9	0.000909	0.000784	1.159703		0.9	0.001282	0.000288	4.443899
	1	0.000909	0.000819	1.109838		1	0.001282	0.000263	4.877343
	0.1	0.00104	0.000526	1.978355		0.1	0.001313	0.000417	3.147851
	0.2	0.00104	0.000548	1.896602		0.2	0.001313	0.000379	3.465365
	0.3	0.00104	0.000571	1.822394		0.3	0.001313	0.00034	3.857866
	0.4	0.00104	0.000593	1.754752		0.4	0.001313	0.000301	4.355411
0.25	0.5	0.00104	0.000614	1.692864	0.45	0.5	0.001313	0.000262	5.006603
	0.6	0.00104	0.000636	1.636043		0.6	0.001313	0.000223	5.895497
	0.7	0.00104	0.000657	1.583709		0.7	0.001313	0.000183	7.181136
	0.8	0.00104	0.000677	1.53537		0.8	0.001313	0.000143	9.205181
	0.9	0.00104	0.000698	1.490601		0.9	0.001313	0.000102	12.85955
	1	0.00104	0.000718	1.449035		1	0.001313	0.0000612	21.44086

Table 4. Relative efficiency comparison between the proposed model and Ewemooje *et al.* (2019a) model when $P_1 = 0.4$; $P_2 = 0.6$; $\alpha = 25$; $\beta = 35$; n=200 for varying π_A and π_U

π_A	π_U	$v(\hat{\pi})$	$v(\hat{\pi}_A)$	RE	π_A	π_U	$v(\hat{\pi})$	$v(\hat{\pi}_A)$	RE
0.1	0.1	0.000586	0.000353	1.660952	0.3	0.1	0.001156	0.000548	2.107674
	0.2	0.000586	0.000425	1.377199		0.2	0.001156	0.000557	2.073896
	0.3	0.000586	0.000498	1.177079		0.3	0.001156	0.000566	2.042447
	0.4	0.000586	0.00057	1.02837		0.4	0.001156	0.000574	2.013165
	0.5	0.000586	0.000641	0.913521		0.5	0.001156	0.000582	1.985906
	0.6	0.000586	0.000713	0.822151		0.6	0.001156	0.000589	1.960538
	0.7	0.000586	0.000783	0.747731		0.7	0.001156	0.000597	1.936947
	0.8	0.000586	0.000854	0.685946		0.8	0.001156	0.000603	1.915028
	0.9	0.000586	0.000924	0.633833		0.9	0.001156	0.00061	1.894686
	1	0.000586	0.000994	0.589287		1	0.001156	0.000616	1.875838
0.15	0.1	0.000766	0.000439	1.74396	0.35	0.1	0.001236	0.000535	2.310815
	0.2	0.000766	0.000496	1.544414		0.2	0.001236	0.000528	2.341486
	0.3	0.000766	0.000552	1.386723		0.3	0.001236	0.00052	2.37458
	0.4	0.000766	0.000608	1.258975		0.4	0.001236	0.000513	2.410267
	0.5	0.000766	0.000664	1.153387		0.5	0.001236	0.000505	2.448743
	0.6	0.000766	0.000719	1.064657		0.6	0.001236	0.000496	2.490223
	0.7	0.000766	0.000774	0.989051		0.7	0.001236	0.000487	2.534952
	0.8	0.000766	0.000829	0.923862		0.8	0.001236	0.000478	2.583208
	0.9	0.000766	0.000883	0.867078		0.9	0.001236	0.000469	2.635303
	1	0.000766	0.000937	0.817176		1	0.001236	0.000459	2.691595
0.2	0.1	0.000921	0.0005	1.839616	0.4	0.1	0.001291	0.000496	2.601386
	0.2	0.000921	0.000541	1.700959		0.2	0.001291	0.000473	2.727499
	0.3	0.000921	0.000582	1.582692		0.3	0.001291	0.00045	2.868693
	0.4	0.000921	0.000622	1.480635		0.4	0.001291	0.000426	3.027789
	0.5	0.000921	0.000662	1.391678		0.5	0.001291	0.000402	3.208359
	0.6	0.000921	0.000701	1.313461		0.6	0.001291	0.000378	3.414995
	0.7	0.000921	0.00074	1.244157		0.7	0.001291	0.000353	3.653698
	0.8	0.000921	0.000779	1.182331		0.8	0.001291	0.000328	3.932472
	0.9	0.000921	0.000817	1.126842		0.9	0.001291	0.000303	4.262223
	1	0.000921	0.000855	1.076768		1	0.001291	0.000277	4.658218
0.25	0.1	0.001051	0.000537	1.956943	0.45	0.1	0.00132	0.000432	3.053186
	0.2	0.001051	0.000562	1.870325		0.2	0.00132	0.000394	3.354703
	0.3	0.001051	0.000586	1.79212		0.3	0.00132	0.000354	3.725976
	0.4	0.001051	0.00061	1.72118		0.4	0.00132	0.000315	4.194316
	0.5	0.001051	0.000634	1.656556		0.5	0.00132	0.000275	4.80343
	0.6	0.001051	0.000658	1.59746		0.6	0.00132	0.000235	5.627905
	0.7	0.001051	0.000681	1.543228		0.7	0.00132	0.000194	6.80632
	0.8	0.001051	0.000704	1.4933		0.8	0.00132	0.000153	8.628625
	0.9	0.001051	0.000726	1.447199		0.9	0.00132	0.000112	11.82045
	1	0.001051	0.000748	1.404517		1	0.00132	0.00007	18.85377

Table 5 shows that as $P_1 = P_2 = 0.5$, the variance of the Ewemooje *et al.* (2019a) model increases at all values of π_A from 0.00040 to 0.00133 while the variance of the proposed model decreases as π_A increases with values ranging from 0.00108 to 0.00008.

The relative efficiency of the proposed model over Ewemooje *et al.* (2019a) also reduces as π_U increases when $0.1 \leq \pi_A \leq 0.3$ and increases as π_U increases when $0.35 \leq \pi_A \leq 0.45$. However, as the sensitive character π_A increases, an appreciable increase is noticed in the values of the relative efficiency ranging from 1.0144 to 16.5954.

Table 5. Relative efficiency comparison between the proposed model and Ewemooje *et al.* (2019a) model when $P_1 = 0.5$; $P_2 = 0.5$; $\alpha = 25$; $\beta = 35$; n=200 for varying π_A and π_U

π_A	π_U	v(π^)	v(π^_A)	RE	π_A	π_U	v(π^)	v(π^_A)	RE
	0.1	0.0006	0.000361	1.662391		0.1	0.001167	0.000561	2.079894
	0.2	0.0006	0.000438	1.3694		0.2	0.001167	0.000571	2.041478
	0.3	0.0006	0.000515	1.165049		0.3	0.001167	0.000582	2.005731
	0.4	0.0006	0.000591	1.014402		0.4	0.001167	0.000591	1.972448
0.1	0.5	0.0006	0.000668	0.898752	0.3	0.5	0.001167	0.000601	1.941448
	0.6	0.0006	0.000743	0.807175		0.6	0.001167	0.00061	1.912568
	0.7	0.0006	0.000819	0.732866		0.7	0.001167	0.000619	1.885663
	0.8	0.0006	0.000894	0.671363		0.8	0.001167	0.000627	1.860602
	0.9	0.0006	0.000968	0.619621		0.9	0.001167	0.000635	1.83727
	1	0.0006	0.001043	0.575488		1	0.001167	0.000643	1.815562
	0.1	0.000779	0.000448	1.737559		0.1	0.001246	0.000548	2.271653
	0.2	0.000779	0.000509	1.530835		0.2	0.001246	0.000542	2.297251
	0.3	0.000779	0.000569	1.36896		0.3	0.001246	0.000536	2.325039
	0.4	0.000779	0.000629	1.238775		0.4	0.001246	0.000529	2.355155
0.15	0.5	0.000779	0.000688	1.131809	0.35	0.5	0.001246	0.000522	2.387755
	0.6	0.000779	0.000748	1.042363		0.6	0.001246	0.000514	2.423015
	0.7	0.000779	0.000806	0.966464		0.7	0.001246	0.000506	2.46113
	0.8	0.000779	0.000865	0.901253		0.8	0.001246	0.000498	2.502325
	0.9	0.000779	0.000923	0.844625		0.9	0.001246	0.000489	2.546848
	1	0.000779	0.00098	0.794993		1	0.001246	0.00048	2.594986
	0.1	0.000933	0.000511	1.826749		0.1	0.0013	0.000511	2.5444
	0.2	0.000933	0.000555	1.682243		0.2	0.0013	0.000488	2.663126
	0.3	0.000933	0.000598	1.559889		0.3	0.0013	0.000465	2.795699
	0.4	0.000933	0.000641	1.454965		0.4	0.0013	0.000441	2.944631
0.2	0.5	0.000933	0.000684	1.364005	0.4	0.5	0.0013	0.000418	3.113082
	0.6	0.000933	0.000727	1.284404		0.6	0.0013	0.000393	3.305085
	0.7	0.000933	0.000769	1.214165		0.7	0.0013	0.000369	3.525866
	0.8	0.000933	0.00081	1.151737		0.8	0.0013	0.000344	3.782328
	0.9	0.000933	0.000852	1.09589		0.9	0.0013	0.000318	4.08377
	1	0.000933	0.000893	1.045643		1	0.0013	0.000293	4.443038
	0.1	0.001063	0.000548	1.937363		0.1	0.001329	0.000448	2.964072
	0.2	0.001063	0.000576	1.845746		0.2	0.001329	0.000409	3.249943
	0.3	0.001063	0.000603	1.763485		0.3	0.001329	0.000369	3.600451
	0.4	0.001063	0.000629	1.689239		0.4	0.001329	0.000329	4.040248
0.25	0.5	0.001063	0.000655	1.621908	0.45	0.5	0.001329	0.000288	4.608347
	0.6	0.001063	0.000681	1.560588		0.6	0.001329	0.000248	5.37037
	0.7	0.001063	0.000706	1.504523		0.7	0.001329	0.000206	6.445891
	0.8	0.001063	0.000731	1.453083		0.8	0.001329	0.000165	8.078222
	0.9	0.001063	0.000756	1.405733		0.9	0.001329	0.000123	10.85034
	1	0.001063	0.00078	1.362018		1	0.001329	0.00008	16.59538

As the probability of selecting the sensitive attribute was increased to $P_1 = 0.6$ and $P_2 = 0.4$. The relative efficiency of the proposed model over Ewemooje *et al.* (2019a) increases with each value of π_U as π_A increases when $0.1 \leq \pi_A \leq 0.3$ and decreases when $0.35 \leq \pi_A \leq 0.45$. The variance of the Ewemooje *et al.* (2019a) model increases at all values of π_A from 0.00041 to 0.00134, the variance of the proposed model decreases as π_A increases with values ranging from 0.00104 to 0.00009 while the relative efficiencies range between 1.0026 and 14.6381 (see Table 6).

Table 6. Relative efficiency comparison between the proposed model and Ewemooje *et al.* (2019a) model when P_1= 0.6; P_2= 0.4; α= 25; β= 35; n=200 for varying π_A and π_U.

π_A	π_U	$v(\hat{\pi})$	$v(\hat{\pi}_A)$	RE	π_A	π_U	$v(\hat{\pi})$	$v(\hat{\pi}_A)$	RE
	0.1	0.000616	0.000369	1.666953		0.1	0.001179	0.000574	2.05419
	0.2	0.000616	0.000451	1.364207		0.2	0.001179	0.000586	2.010996
	0.3	0.000616	0.000533	1.155374		0.3	0.001179	0.000598	1.970869
	0.4	0.000616	0.000614	1.002628		0.4	0.001179	0.00061	1.933553
0.1	0.5	0.000616	0.000695	0.886052	0.3	0.5	0.001179	0.000621	1.898819
	0.6	0.000616	0.000776	0.794162		0.6	0.001179	0.000632	1.866467
	0.7	0.000616	0.000856	0.719869		0.7	0.001179	0.000642	1.836318
	0.8	0.000616	0.000935	0.658563		0.8	0.001179	0.000652	1.808212
	0.9	0.000616	0.001014	0.607113		0.9	0.001179	0.000662	1.782007
	1	0.000616	0.001093	0.563322		1	0.001179	0.000671	1.757575
	0.1	0.000794	0.000458	1.733651		0.1	0.001257	0.000563	2.234883
	0.2	0.000794	0.000523	1.519457		0.2	0.001257	0.000557	2.255279
	0.3	0.000794	0.000587	1.353272		0.3	0.001257	0.000552	2.277664
	0.4	0.000794	0.000651	1.220588		0.4	0.001257	0.000546	2.302146
0.15	0.5	0.000794	0.000714	1.112209	0.35	0.5	0.001257	0.00054	2.328848
	0.6	0.000794	0.000777	1.022022		0.6	0.001257	0.000533	2.357906
	0.7	0.000794	0.00084	0.945804		0.7	0.001257	0.000526	2.389473
	0.8	0.000794	0.000902	0.880547		0.8	0.001257	0.000519	2.423725
	0.9	0.000794	0.000964	0.824049		0.9	0.001257	0.000511	2.460857
	1	0.000794	0.001025	0.774659		1	0.001257	0.000503	2.501089
	0.1	0.000947	0.000522	1.816041		0.1	0.001311	0.000526	2.490643
	0.2	0.000947	0.000569	1.66549		0.2	0.001311	0.000504	2.601916
	0.3	0.000947	0.000616	1.538968		0.3	0.001311	0.000481	2.725814
	0.4	0.000947	0.000662	1.431157		0.4	0.001311	0.000458	2.864549
0.2	0.5	0.000947	0.000708	1.338201	0.4	0.5	0.001311	0.000434	3.020886
	0.6	0.000947	0.000754	1.257237		0.6	0.001311	0.00041	3.198326
	0.7	0.000947	0.000799	1.186092		0.7	0.001311	0.000385	3.401361
	0.8	0.000947	0.000844	1.123088		0.8	0.001311	0.00036	3.635866
	0.9	0.000947	0.000888	1.06691		0.9	0.001311	0.000335	3.90966
	1	0.000947	0.000932	1.016511		1	0.001311	0.00031	4.233393
	0.1	0.001076	0.00056	1.919806		0.1	0.001339	0.000465	2.880362
	0.2	0.001076	0.00059	1.823035		0.2	0.001339	0.000425	3.150967
	0.3	0.001076	0.000619	1.736648		0.3	0.001339	0.000385	3.481228
	0.4	0.001076	0.000648	1.659079		0.4	0.001339	0.000344	3.893252
0.25	0.5	0.001076	0.000677	1.589061	0.45	0.5	0.001339	0.000303	4.421607
	0.6	0.001076	0.000705	1.52556		0.6	0.001339	0.000261	5.123548
	0.7	0.001076	0.000733	1.467722		0.7	0.001339	0.000219	6.10128
	0.8	0.001076	0.00076	1.414838		0.8	0.001339	0.000177	7.556836
	0.9	0.001076	0.000787	1.366311		0.9	0.001339	0.000135	9.953349
	1	0.001076	0.000814	1.321637		1	0.001339	0.0000915	14.63817

Table 7 shows that for varying π_A and π_U, P_1= 0.7; P_2= 0.3, the variance of the Ewemooje *et al.* (2019a) model increases at all values of π_A from 0.00043 to 0.00135 while the variance of the proposed model also increases as π_U increases when 0.1 ≤

$\pi_A \leq 0.3$ and decreases as π_U increases when $0.35 \leq \pi_A \leq 0.45$. The relative efficiency of the proposed model over Ewemooje *et al.* (2019a) shows that as the sensitive character π_A increases, the relative efficiency increases with the values ranging from 1.0037 to 12.9490.

Table 7. Relative efficiency comparison between the proposed model and Ewemooje *et al.* (2019a) model when P_1= 0.7; P_2= 0.3; α= 25; β= 35; n=200 for varying π_A and π_U.

π_A	π_U	$v(\hat{\pi})$	$v(\hat{\pi}_A)$	RE	π_A	π_U	$v(\hat{\pi})$	$v(\hat{\pi}_A)$	RE
	0.1	0.000634	0.000378	1.67503		0.1	0.001193	0.000587	2.030752
	0.2	0.000634	0.000465	1.361891		0.2	0.001193	0.000602	1.982633
	0.3	0.000634	0.000552	1.148251		0.3	0.001193	0.000615	1.938042
	0.4	0.000634	0.000638	0.993193		0.4	0.001193	0.000629	1.896659
0.1	0.5	0.000634	0.000724	0.875529	0.3	0.5	0.001193	0.000642	1.8582
	0.6	0.000634	0.000809	0.783192		0.6	0.001193	0.000655	1.822421
	0.7	0.000634	0.000894	0.7088		0.7	0.001193	0.000667	1.789101
	0.8	0.000634	0.000979	0.647589		0.8	0.001193	0.000679	1.758048
	0.9	0.000634	0.001063	0.596341		0.9	0.001193	0.00069	1.729089
	1	0.000634	0.001146	0.552808		1	0.001193	0.000701	1.702072
	0.1	0.000811	0.000468	1.73255		0.1	0.00127	0.000577	2.200659
	0.2	0.000811	0.000537	1.510522		0.2	0.00127	0.000573	2.215729
	0.3	0.000811	0.000605	1.33985		0.3	0.00127	0.000569	2.232628
	0.4	0.000811	0.000673	1.20457		0.4	0.00127	0.000564	2.251434
0.15	0.5	0.000811	0.000741	1.094713	0.35	0.5	0.00127	0.000559	2.272238
	0.6	0.000811	0.000808	1.003731		0.6	0.00127	0.000553	2.295143
	0.7	0.000811	0.000875	0.927151		0.7	0.00127	0.000547	2.320265
	0.8	0.000811	0.000941	0.861805		0.8	0.00127	0.000541	2.347734
	0.9	0.000811	0.001007	0.805395		0.9	0.00127	0.000534	2.377699
	1	0.000811	0.001072	0.756207		1	0.00127	0.000527	2.410328
	0.1	0.000963	0.000533	1.807757		0.1	0.001322	0.000542	2.440203
	0.2	0.000963	0.000583	1.650922		0.2	0.001322	0.00052	2.54398
	0.3	0.000963	0.000634	1.520119		0.3	0.001322	0.000497	2.659183
	0.4	0.000963	0.000683	1.409372		0.4	0.001322	0.000474	2.787741
0.2	0.5	0.000963	0.000733	1.314408	0.4	0.5	0.001322	0.000451	2.932044
	0.6	0.000963	0.000782	1.232084		0.6	0.001322	0.000427	3.095093
	0.7	0.000963	0.00083	1.160041		0.7	0.001322	0.000403	3.280704
	0.8	0.000963	0.000879	1.096473		0.8	0.001322	0.000379	3.493806
	0.9	0.000963	0.000926	1.039974		0.9	0.001322	0.000354	3.740884
	1	0.000963	0.000974	0.989432		1	0.001322	0.000328	4.030638
	0.1	0.001091	0.000573	1.904498		0.1	0.00135	0.000482	2.801958
	0.2	0.001091	0.000605	1.802396		0.2	0.00135	0.000441	3.057695
	0.3	0.001091	0.000637	1.711794		0.3	0.00135	0.000401	3.368273
	0.4	0.001091	0.000669	1.630872		0.4	0.00135	0.00036	3.753384
0.25	0.5	0.001091	0.0007	1.558175	0.45	0.5	0.00135	0.000318	4.243433
	0.6	0.001091	0.000731	1.492525		0.6	0.00135	0.000276	4.887965
	0.7	0.001091	0.000761	1.432962		0.7	0.00135	0.000234	5.773544
	0.8	0.001091	0.000791	1.37869		0.8	0.00135	0.000191	7.066269
	0.9	0.001091	0.000821	1.329049		0.9	0.00135	0.000148	9.13035
	1	0.001091	0.00085	1.283481		1	0.00135	0.000104	12.94899

5. Application of the Proposed Model

The proposed method and direct method (DM) were used in collecting information on examination malpractices prevalence among students at a Nigerian university. Two hundred (200) instruments were administered using two decks of cards consisting of the sensitive question *"have you ever been involved in examination malpractices?"* and unrelated question "*do you love soccer*?" as the randomized devices. The respondents were given proper education on how to use the randomized devices with appropriate demonstration. They were also assured of confidentiality by ensuring that responses given cannot be traced to respondents; hence, they willingly participated in the survey. The respondents were directly (DM) asked the sensitive question *"have you ever been involved in examination malpractices?"*. If "yes" answer is obtained, he/she does not use the randomized device but if he/she answers "no", then he/she is instructed to choose one of the two decks of cards at random and then respond accordingly without revealing question answered to the interviewer. The two randomized devices R_1 and R_2 consist of two unrelated questions (the sensitive question with probability $P_1 = 0.7$, and unrelated question with probability $1\text{-}P_1 = 0.3$ for R_1 while $P_2 = 0.3$ and $P_2 = 0.7$ for R_2).

The age distribution of the sampled respondents ranges between 16 and 29 years with the age group 20–24 years having the higher percentage of 58.0% and about three-quarters of them are male (74.0%). The true proportion of respondents who answered "yes" to the unrelated question (π_U) "*do you love soccer*?" is 0.45. The estimate of examination malpractices prevalence and their associated coefficient of variation (CV) are presented in Table 8. The DM estimate prevalence of examination malpractices at 19.0% compared to 23.0% for the proposed method. The standard error associated with DM is 0.028 (CV = 14.6%) while the proposed model is 0.026 (CV = 11.5%).

However, contrary to what was reported by Jann et al., (2012) where Crosswise Model (CM) produced higher estimate with higher standard error, the proposed method produced higher estimate with lower standard error as against the DM. Hence, the proposed model performs better than the DM in line with earlier works of Jann *et al.* (2012), Ewemooje *et al.,* (2017), Cobo *et al.,* (2016) and Ewemooje *et al.,* (2019b).

Table 8. Comparative analysis of the proposed model versus the direct method

Method	π_A	$V(\hat{\pi}_A)$	$S.E(\hat{\pi}_A)$	$C.V(\hat{\pi}_A)$
Direct Method	0.19	0.00077	0.028	14.6%
Proposed Model	0.23	0.00070	0.026	11.5%

6. Conclusion

The unrelated design has been shown to improve efficiency of a randomized response method and to reduce distrust of the respondents; hence, we proposed a new Randomized Response Model (RRM) which consists of the unrelated questions in dichotomous randomized response model. To ensure better efficiency, the proportion of the sensitive attribute must be at least 0.2 and not greater than 0.5. The variance of the proposed model decreases as the proportion of the sensitive attribute π_A and unrelated attribute π_U increases as against the Ewemooje *et al.* (2019a) model, which increases as the proportion of the sensitive attribute increases. The relative efficiency of the proposed model over Ewemooje *et al.* (2019a) reduces as π_U increases when $0.05 \leq \pi_A \leq 0.3$ and increases as π_U increases when $0.35 \leq \pi_A \leq 0.45$. Also, as the sample size increases from 50 to 500, the relative efficiency of the proposed model stood at 9.93 while as P_1 increases and P_2 decreases, the relative efficiency reduces from 7.09 to 6.23. Application of the proposed model also revealed it efficiency over the direct method in estimating the prevalence of examination malpractices among university students. The direct method estimated the prevalence of examination malpractices among university students at 19.0% while the proposed method estimated it at 23.0%. Hence, the proposed model is shown to be more efficient than the direct method and Ewemooje *et al.* (2019a) model as the proportion of people belonging to the sensitive attribute increases.

References

ADEBOLA, F. B., ADEPETUN, A. O., (2011). A new Tripartite Randomized Response Technique. Journal of the Nigerian Association of Mathematical Physics, 19, pp. 119–122.

ADEBOLA, F. B., ADEDIRAN, A. A., EWEMOOJE, O. S., (2017). Hybrid tripartite randomized response technique. Communications in Statistics-Theory and Methods, 46(23), pp. 11756–11763; https://doi.org/10.1080/03610926.2016.1277760.

COBO, B., RUEDA, M. M., LÓPEZ –TORRECILLAS, F., (2016). Application of randomized response techniques for investigating cannabis use by Spanish university students, International Journal of Methods in Psychiatric Research; https://doi.org/10.1002/mpr.1517.

EWEMOOJE, O. S., (2017). Estimating two sensitive characters with equal probabilities of protection, Cogent Mathematics, 4, pp. 1–14.

EWEMOOJE, O. S., ADEBOLA, F. B., AMAHIA, G. N., (2019a). Alternative Unbiased Estimator in Dichotomous Randomized Response Technique, Communication in

Mathematics and Statistics, 7(4), pp. 383–400; https://doi.org/10.1007/s40304-018-0145-x.

EWEMOOJE, O. S., ADEBOLA, F. B., ADEDIRAN, A. A., (2018). A Stratified Hybrid Tripartite Randomized Response Technique, Gazi University Journal of Science, 31(4), pp. 1246–1266.

EWEMOOJE O. S., ADEBOLA, F. B., AWOGBEMILA, A. T., (2019b). Substance Use disorder prevalence using tripartite randomized technique, 14(1), pp. 55–61. https://dx.doi.org/10.4314/njtr.v14i1.8.

EWEMOOJE, O. S., AMAHIA, G. N., ADEBOLA, F. B., (2017). Estimating prevalence of induced abortion and multiple sexual partners using improved randomized response technique for two sensitive attributes, Communications in Statistics: Case Studies, Data Analysis and Applications, 3(1-2), pp. 21–28; https://doi.org/10.1080/23737484.2017.1398057.

FOX, J-P., VEEN, D., KLOTZKE K., (2019). Generalized Linear Mixed Models for Randomized Responses, Methodology (2019), 15(1), pp. 1–18; https://doi.org/10.1027/1614-2241/a000153.

GREENBERG, B., ABUL-ELA, A., SIMMONS, W., HORVITZ, D., (1969). The Unrelated question Randomized Response Technique: Theoretical framework, Journal of American Statistical Association, 64, pp. 529–539.

HOGLINGER, M., JANN, B., (2018). More is not always better: An experimental individual-level validation of the randomized response technique and the crosswise model, PLoS ONE 13(8): e0201770; https://doi.org/10.1371/journal.pone.0201770.

HORVITZ, D. G., SHAH, B. V., SIMMONS, W. R., (1967). The unrelated question Randomized Response Technique, Proceedings of social statistical section American Statistical Association, pp. 65–72.

HUSSAIN, Z., SHABBIR, J., (2007). Randomized use of Warner's randomized response Technique, Interstat: April #7.http://interstat.statjournals.net/INDEX/Apro7.html

JANN, B., JERKE, J., KRUMPAL, I., (2012). Asking sensitive question using the crosswise model, an experimental survey measuring plagiarism. Public Opinion Quarterly, 76, pp. 32–49; https://doi.org/10.1093/poq/nfr036.

MANGAT, N. S., SINGH, R., (1990). An alternative randomized response procedure, Biometrika, 77, pp. 439–442.

WARNER, S. L., (1965). Randomized response: a survey technique for eliminating evasive answers bias, Journal of the American Statistical Association, 60, pp. 63–69.

YU, J.-W.., TIAN, G.-L., TANG, M.-L., (2008). Two new models for survey sampling with sensitive characteristic; design and analysis. Metrika, 67, pp. 251–263; https://doi.org/10.1007/s00184-007-0131-x.