# An Improved Versatile Optional Randomised Response Technique for Efficient Estimation in Sensitive Surveys

**M. A. Yunusa[1], A. Audu[1, 2], U. Usman[1], K. O. Aremu[3], and A. Adejumobi[4]**

*Collecting reliable data on sensitive issues remains a major challenge because respondents are often reluctant to provide truthful answers to sensitive quantitative questions or may refuse to respond altogether. Consequently, direct questioning on topics such as drug use, child abuse, sexual harassment, sexual behaviour, and cybercrime frequently produces biased data and unreliable estimates. The randomised response technique (RRT) addresses this challenge by protecting respondents' privacy, thereby encouraging truthful responses while preventing the disclosure of individual information to the researcher or any third party. However, many existing RRT models provide only modest improvements in estimation efficiency. This study proposes an improved, versatile, optional randomised response technique for collecting quantitative sensitive data with enhanced estimation efficiency. The statistical properties of the proposed model, including its estimator, variance, privacy level, and a combined measure of efficiency and privacy protection, are derived. A numerical study based on four real datasets demonstrates that the proposed model consistently outperforms the existing versatile optional RRT model in terms of estimation efficiency while maintaining respondent privacy. The findings suggest that the proposed model offers a more effective approach for conducting quantitative sensitive surveys and can improve the quality of statistical inference in studies involving sensitive characteristics.*

**Keywords**: Randomised Response, Sampling, Randomisation device, Privacy level, Estimate

**DOI**: https://doi.org/10.63255/02-1934.25/05

## 1.0 Introduction

Sensitive surveys are surveys that involve investigating sensitive issues such as: cases of rape, cases of examination malpractices, drug use, sexual harassment, cybercrime, cases of illegal earnings of government workers and many more. The problems encountered during sensitive surveys include reporting of false answers and non-response bias on the part of respondents; in such situations, estimation becomes a problem. To get reliable data and estimates, we use the randomised response technique (RRT), a privacy protection technique used to collect reliable data on sensitive topics from respondents.

One of the survey methods of obtaining sensitive information from respondents concerning sensitive issues called the randomized response technique (RRT) was first initiated by Warner (1965) to collect sensitive information on qualitative data and to estimate the prevalence or proportion of sensitive characters like income, tax evasion, abortion cases, drug use, etc, through the use of randomization device like spinner, deck of cards, coins, etc., such that the responses of

[1] Department of Statistics, Usmanu Danfodiyo University, Sokoto, Nigeria
[2] Corresponding Email: ahmed.audu@udusok.edu.ng. The views expressed in this paper are solely those of the author(s) and do not necessarily reflect the views, policies, or official position of the Chartered Institute of Statisticians of Nigeria (CISON), its Council, Management, or any of its affiliated bodies.
[3] Department of Mathematics, Usmanu Danfodiyo University, Sokoto, Nigeria
[4] Department of Mathematics, Kebbi State University of Science and Technology, Aliero, Nigeria

the participants in the survey is protected. Warner (1971) also introduced another randomised response technique (RRT) to obtain sensitive information that is quantitative in nature. In the case of quantitative randomized response technique (RRT), many authors have modified this technique, among them include: Singh and Gorey (2019), Saha (2007,2008), Gupta et al. (2002), Bar Lev et al. (2004), Gjestvang and Singh (2009), Odumade and Singh (2010), Diana and Perri (2010,2011), Singh and Gorey (2019), Narjis and Shabbir (2021), Azeem et al. (2024), Yunusa and Adejumobi (2024), Audu et al. (2025), Yunusa et al. (2025a), Yunusa et al. (2025b). These authors have made remarkable contributions by developing RRT model that can be used in sensitive surveys to obtain reliable data. However, the models developed by these authors in the literature may be efficient, but the privacy level of the models might be jeopardised; on the other hand, the privacy level of the model might be significant, but the model efficiency is jeopardised. To arrive at or decide on a better model between two or more models, Gupta et al. (2018) introduced a metric known as the combined metric of efficiency and privacy level of a randomised response technique (RRT) model to identify a better model. Gupta et al. (2018) suggested that a model with a minimum combined metric of efficiency and privacy level is a better model among several models under study.

Azeem et al. (2024) modified the randomised response model suggested by Gupta et al. (2022), by including the fourth option $TY_i + S$ into the model of Gupta et al. (2022), the modified model achieved more efficiency (lower variance) as compared to Gupta et al. (2022) model, but the privacy level of Gupta et al. (2022) model is higher than that of Azeem et al. (2024) model. The combined metric of both models was the same; this implies that both models are better models, although there is a trade-off between model efficiency and Privacy. The Gupta et al. (2022) model is preferred if the goal of the survey is a high privacy level, and if the priority of the model is model efficiency, Azeem et al. (2024) model is preferred. Despite the efficiency gain in the model of Azeem et al. (2024), this efficiency gain remained limited as the privacy efficiency trade-off persists. The research gap is that the efficiency gain in the work of Azeem et al. (2024) remained limited and needs to be enhanced for better estimates to be obtained.

In this present study, an improved versatile optional RRT model with increased efficiency over existing models was proposed, and the properties such as efficiency, privacy level and combined metric of efficiency and privacy level were established theoretically and numerically evaluated.

## 1.1 Literature review

In this section, some pertinent studies on the Randomised Response Technique (RRT) were reviewed,

Warner (1971) introduced an RRT model called the additive model for data collection on sensitive issues. The mathematical expression of the model and its properties were given in section two

Diana and Perri (2011) modified the RRT model of Warner (1971) to obtain a scrambling model with two scramble variables. The modified model had a higher privacy level but possessed poor efficiency. The mathematical expression of the model and properties were given in section two

Gupta et al. (2002) introduced the optional RRT model with two options: either respondents report true responses or scrambled responses with a single scramble variable S. The model had higher efficiency but a poor privacy level as compared to Warner (1971) and Diana and Perri (2011). The mathematical expression of the model and properties were given in section two

The model suggested by Gupta et al. (2022) provides the following three options to the respondents to answer questions in the survey as

1. Report the true response
2. Apply the additive randomisation technique only
3. Apply both additive and multiplicative randomisation techniques together

The model suggested by Azeem et al. (2024) provides the following four options to the respondents to answer questions in the survey as

1. Report the true response
2. Apply the additive randomisation technique only
3. Apply the multiplicative randomisation technique only
4. Apply both additive and multiplicative randomisation techniques together

## 2.0 Methodology

Let a population under study possess a total of N units and let a random sample of n units be chosen from the study population.

Denote Y as the sensitive variable under study, and S as a multiplicative-type random variable or scrambling variable. Also, let $E\left(Y_i\right)=\mu_y$, $Var\left(Y_i\right)=\sigma_y^2$, $E\left(S\right)=0$, $E\left(T\right)=1$, $Var\left(S\right)=\sigma_s^2$ and $Var\left(T\right)=\sigma_T^2$ are the mean and variance of the sensitive variable, mean and variance of the scrambling variables S and T.

$W$ is the probability of respondents considering the sensitive question in the survey to be sensitive, respectively.

### 2.1 Some existing forced RRT models for collection of sensitive information

In the forced RRT models, all the respondents in the survey are expected to scramble their responses, whether the question posed to them in the survey is viewed as sensitive or not sensitive. Among the authors who suggested forced RRT models are Warner (1971), Eichorn and Hayre (1983), Diana and Perri (2011), Yunusa et al. (2025b), etc. Some of these authors' models are presented in this section

Warner (1971) introduced the additive RRT model for the collection of quantitative sensitive information in which all respondents in the survey scramble their responses with a single scrambling variable S as

$$Z_{Wi}=Y_i+S \tag{1}$$

The estimator, variance, privacy level and the combined metric of efficiency and privacy level are as follows:

$$\hat{\mu}_W = \frac{1}{n}\sum_{i=1}^{n} Z_{Wi} \tag{2}$$

$$Var\left(\hat{\mu}_W\right) = \frac{\sigma_y^2 + \sigma_s^2}{n} \tag{3}$$

$$\Delta_W = \sigma_s^2 \tag{4}$$

$$\delta_W = \frac{\sigma_y^2 + \sigma_s^2}{n\sigma_s^2} \tag{5}$$

Diana and Perri (2011), taking motivation from Warner (1971), proposed a randomised response model RRM with the use of multiplicative and additive scrambling together, as in equation (6)

$$Z_{DPi} = TY_i + S \tag{6}$$

The estimator, variance, privacy level and the combined metric of efficiency and privacy level are as follows:

$$\hat{\mu}_{DP} = \frac{1}{n}\sum_{i=1}^{n} Z_{DPi} \tag{7}$$

$$Var\left(\hat{\mu}_{DP}\right) = \frac{\sigma_T^2\left(\mu_y^2 + \sigma_y^2\right) + \sigma_y^2 + \sigma_s^2}{n} \tag{8}$$

$$\Delta_{DP} = \sigma_T^2\left(\mu_y^2 + \sigma_y^2\right) + \sigma_s^2 \tag{9}$$

$$\delta_{DP} = \frac{\sigma_T^2\left(\mu_y^2 + \sigma_y^2\right) + \sigma_y^2 + \sigma_s^2}{n\left[\sigma_T^2\left(\mu_y^2 + \sigma_y^2\right) + \sigma_s^2\right]} \tag{10}$$

### 2.1 Some existing optional RRT models for collection of sensitive information

In the optional RRT models, some of the respondents in the survey considered the survey question sensitive and were allowed to scramble their responses. In contrast, some of them considered the question non-sensitive and answered the question posed to them honestly. Among the authors who suggested optional RRT models are: Gupta et al. (2002), Gupta et al. (2022), Singh and Gorey (2019), Azeem et al. (2024), Audu et al. (2025), Yunusa et al. (2025a), etc. Some of these authors' models are presented in this section.

Gupta et al. (2002) were the first to initiate the concept of the optional randomised response technique and proposed the optional randomised response technique (ORRT) model as in (11)

$$Z_{G*i} = \begin{cases} Y_i & \text{with probability} \quad 1-W \\ Y_i T & \text{with probability } W \end{cases} \tag{11}$$

The estimator, variance, privacy level and the combined metric of efficiency and privacy level are as follows:

$$\hat{\mu}_{G*} = \frac{1}{n}\sum_{i=1}^{n} Z_{G*i} \tag{12}$$

$$Var\left(\hat{\mu}_{G*}\right) = \frac{\sigma_y^2 + W\sigma_T^2\left(\sigma_y^2 + \mu_y^2\right)}{n} \tag{13}$$

$$\Delta_{G*} = W\sigma_T^2\left(\mu_y^2 + \sigma_y^2\right) \tag{14}$$

$$\delta_{G*} = \frac{\sigma_y^2 + W\sigma_T^2\left(\sigma_y^2 + \mu_y^2\right)}{nW\sigma_T^2\left(\sigma_y^2 + \mu_y^2\right)} \tag{15}$$

Gupta et al. (2022) proposed an optional randomised response (RRT) model with three options to respond to a question posed to them in a sensitive survey, as in (16)

$$Z_{Gi} = \begin{cases} Y_i & \text{with probability} \quad 1-W \\ Y_i + S & \text{with probability } WA \\ TY_i + S & \text{with probability } W\left(1-A\right) \end{cases} \tag{16}$$

The estimator, variance, privacy level and the combined metric of efficiency and privacy level are as follows:

$$\hat{\mu}_{G} = \frac{1}{n}\sum_{i=1}^{n} Z_{Gi} \tag{17}$$

$$Var\left(\hat{\mu}_{G}\right) = \frac{W\left(1-A\right)\sigma_T^2\left(\sigma_y^2 + \mu_y^2\right) + \sigma_y^2 + W\sigma_s^2}{n} \tag{18}$$

$$\Delta_{G} = W\left(1-A\right)\sigma_T^2\left(\mu_y^2 + \sigma_y^2\right) + W\sigma_s^2 \tag{19}$$

$$\delta_G = \frac{W(1-A)\sigma_T^2(\sigma_y^2+\mu_y^2)+\sigma_y^2+W\sigma_s^2}{n\left[W(1-A)\sigma_T^2(\sigma_y^2+\mu_y^2)+W\sigma_s^2\right]} \tag{20}$$

The scrambling RRT model by Azeem et al. (2024) for the collection of sensitive information when the study variable is sensitive is given as

$$Z_{AZi} = \begin{cases} Y_i \quad \text{with probability} \quad 1-W \\ Y_i+S \quad \text{with probability } WA \\ TY_i \quad \text{with probability } WB \\ TY_i+S \quad \text{with probability } W(1-A-B) \end{cases} \tag{21}$$

The estimator, variance, privacy level and the combined metric of efficiency and privacy level are as follows:

$$\hat{\mu}_{AZ} = \frac{1}{n}\sum_{i=1}^{n} Z_{AZi} \tag{22}$$

$$Var(\hat{\mu}_{AZ}) = \frac{W(1-A)\sigma_T^2(\sigma_y^2+\mu_y^2)+\sigma_y^2+W(1-B)\sigma_s^2}{n} \tag{23}$$

$$\Delta_{AZ} = W(1-A)\sigma_T^2(\mu_y^2+\sigma_y^2)+W(1-B)\sigma_s^2 \tag{24}$$

$$\delta_{AZ} = \frac{W(1-A)\sigma_T^2(\sigma_y^2+\mu_y^2)+\sigma_y^2+W(1-B)\sigma_s^2}{n\left[W(1-A)\sigma_T^2(\sigma_y^2+\mu_y^2)+W(-B)\sigma_s^2\right]} \tag{25}$$

where, $A, B$ are constants to be determined by the investigator during the conduct of the sensitive survey.

## 3 Proposed Improved Optional Randomised Response Technique (RRT) Model

Having studied the Randomised Response Technique (RRT) Model suggested by Azeem et al. (2024), and identified the gap in the previous research, that is, the efficiency gain in the previous study remained limited, hence, an improved model was proposed as follows by replacing the option $TY_i + S$ in the model of Azeem et al. (2024) by $Y_i + ST$ for increased efficiency gain as:

$$Z_{ymi} = \begin{cases} Y_i & with\ probability\ \ 1-W \\ Y_i + S & with\ probability\ WA \\ TY_i & with\ probability\ WB \\ Y_i + ST & with\ probability\ W(1-A-B) \end{cases} \tag{26}$$

where S and T are the scrambling variables distributed normally with $\left(E(S)=0, V(S)=\sigma_s^2\right)$, $\left(E(T)=1, V(T)=\sigma_T^2\right)$ and they are independent of the sensitive variable Y, and W is the proportion of respondents who considered the survey question to be sensitive; A and B are constants to be determined by the investigator whose range of values is 0 to 1 inclusive.

The estimator of the proposed improved RRT model is given as

$$E\left(Z_{ymi}\right) = \mu_{ym} = \frac{1}{n}\sum_{i=1}^{n} Z_{ymi} \tag{27}$$

**3.1 Variance, Privacy level and combined metric of efficiency and privacy of the model**

The properties of the proposed model, such as variance, privacy level of the model, and combined metric of efficiency and privacy level are defined in this section.

The measure of model efficiency is defined as

$$Var\left(\mu_{ymi}\right) = \frac{1}{n}\left[E\left(Z_{ymi}^2\right)^2 - \left(E\left(Z_{ymi}\right)\right)^2\right] \tag{28}$$

where, $E(T)=1,\ E\left(T^2\right)=\sigma_T^2+1,\ E\left(Y_i\right)=\mu_y,\ E\left(Y_i^2\right)=\mu_y^2+\sigma_y^2,\ E\left(S^2\right)=\sigma_s^2,\ E(S)=\mu_s=0.$

According to Yan et al. (2008), the measure for judging a model with higher privacy is defined in (29) as

$$\Delta_{ymi} = E\left(Z_{ymi}-Y\right)^2 \tag{29}$$

According to Gupta et al. (2018), the combined metric of efficiency and privacy level for judging a better model is defined in (30) as

$$\delta_{ymi} = \frac{Var\left(\hat{\mu}_{ymi}\right)}{\Delta_{ymi}} \tag{30}$$

To get the variance of the estimator of the proposed model, we use the relation in (28), and get (31)

$$Var\left(\hat{\mu}_{ymi}\right)=\frac{1}{n}\left[\begin{array}{l}(1-W)E\left(Y_i^2\right)+WAE\left(\left(Y_i+S\right)^2\right)+WBE\left(\left(TY_i\right)^2\right)\\ W\left(1-A-B\right)BE\left(\left(Y_i+ST\right)^2\right)-\left(\begin{array}{l}(1-W)E\left(Y_i\right)+WAE\left(Y_i+S\right)\\ +WBE\left(Y_i+S\right)+E\left(Y_i+ST\right)\end{array}\right)^2\end{array}\right] \tag{31}$$

$$Var\left(\hat{\mu}_{ymi}\right)=\frac{1}{n}\left[\begin{array}{l}(1-W)E\left(Y_i^2\right)+WAE\left(Y_i^2+S^2+2Y_iS\right)+WBE\left(T^2Y_i^2\right)\\ W\left(1-A-B\right)BE\left(Y_i^2+S^2T^2+2Y_iST\right)-\left(\begin{array}{l}(1-W)E\left(Y_i\right)+WAE\left(Y_i+S\right)\\ +WBE\left(Y_i+S\right)+E\left(Y_i+ST\right)\end{array}\right)^2\end{array}\right] \tag{32}$$

By simplifying (32) and applying the relation in (28), we obtain the variance of the estimator of the proposed model as in (33)

$$Var\left(\hat{\mu}_{ymi}\right)=\frac{1}{n}\left[\sigma_y^2+W\sigma_s^2\left[\sigma_T^2\left(1-A\right)+1\right]+WB\left[\sigma_T^2\left(\mu_y^2+\sigma_y^2\right)-\sigma_s^2\left(\sigma_T^2+1\right)\right]\right] \tag{33}$$

To obtain the privacy level of the proposed models, we use the relation (29), and we get (34)

$$\Delta_{ymi}=\left[\begin{array}{l}(1-W)E\left(Y_i-Y_i\right)^2+WAE\left(Y_i+S-Y_i\right)^2\\ +WBE\left(TY_i-Y_i\right)^2+\left(1-WA-WB\right)E\left(Y_i+ST-Y_i\right)^2\end{array}\right] \tag{34}$$

$$\Delta_{ymi}=\left[WAE\left(S^2\right)+WBE\left(T^2Y_i^2+Y_i^2-2TY_i^2\right)+\left(1-WA-WB\right)E\left(S^2T^2\right)\right] \tag{35}$$

By simplifying (35) and applying the relation in (28), we obtain the privacy level of the proposed model as in (36)

$$\Delta_{ymi}=\sigma_y^2+W\sigma_s^2A+WB\sigma_T^2\left(\mu_y^2+\sigma_y^2\right)+W\left(1-A-B\right)\sigma_s^2\left(\sigma_T^2+1\right) \tag{36}$$

To obtain the combined metric of efficiency and privacy level, we use the relation in (33) and (36) as

$$\delta_{ymi}=\frac{\sigma_y^2+W\sigma_s^2\left[\sigma_T^2\left(1-A\right)+1\right]+WB\left[\sigma_T^2\left(\mu_y^2+\sigma_y^2\right)-\sigma_s^2\left(\sigma_T^2+1\right)\right]}{n\left(\sigma_y^2+W\sigma_s^2A+WB\sigma_T^2\left(\mu_y^2+\sigma_y^2\right)+W\left(1-A-B\right)\sigma_s^2\left(\sigma_T^2+1\right)\right)} \tag{37}$$

## 4. Efficiency Comparisons

The improved model is more efficient than the model suggested by Gupta et al. (2002), if the following condition is satisfied

If $Var\left(\hat{\mu}_{ymi}\right)-Var\left(\hat{\mu}_{G*i}\right)<0$ , then

$$\frac{\left[\sigma_s^2\left[\sigma_T^2\left(1-A\right)+1\right]+B\left[\sigma_T^2\left(\mu_y^2+\sigma_y^2\right)-\sigma_s^2\left(\sigma_T^2+1\right)\right]\right]}{\sigma_T^2\left(\sigma_y^2+\mu_y^2\right)}<1 \tag{38}$$

The improved model is more efficient than the model suggested by Gupta et al. (2022), if the following condition is satisfied

If $Var\left(\hat{\mu}_{ymi}\right)-Var\left(\hat{\mu}_{Gi}\right)<0$ , then

$$\frac{\sigma_s^2\left[\sigma_T^2\left(1-A\right)+1\right]+B\left[\sigma_T^2\left(\mu_y^2+\sigma_y^2\right)-\sigma_s^2\left(\sigma_T^2+1\right)\right]}{\left(1-A\right)\sigma_T^2\left(\sigma_y^2+\mu_y^2\right)+\sigma_s^2}<1 \tag{39}$$

The improved model is more efficient than the model suggested by Azeem et al. (2024), if the following condition is satisfied

If $Var\left(\hat{\mu}_{ymi}\right)-Var\left(\hat{\mu}_{AZi}\right)<0$ , then

$$\frac{\left[\sigma_s^2\left[\sigma_T^2\left(1-A\right)+1\right]+B\left[\sigma_T^2\left(\mu_y^2+\sigma_y^2\right)-\sigma_s^2\left(\sigma_T^2+1\right)\right]\right]}{\left(1-A\right)\sigma_T^2\left(\sigma_y^2+\mu_y^2\right)+\sigma_y^2+\left(1-B\right)\sigma_s^2}<1 \tag{40}$$

## 5. Numerical illustration

To demonstrate the performance of the improved model and existing models, the following datasets were used. The parameters used for the numerical illustration are the population size (N), sample size (n), population mean for Y, S, T, are $\left(\mu_y\right)$, $\left(\mu_s\right)$, $\left(\mu_T\right)$ and population variance for Y, S, T are $\left(\sigma_y^2\right)$, $\left(\sigma_y^2\right)$, $\left(\sigma_y^2\right)$ and $W$, the probability of answering the sensitive questions and a sensitivity analysis was also carried out with the data by changing A and B, A and B are constants to be determined by the interviewer, assuming values between 0 and 1.

**Population 1 : (Source: Azeem et al. (2024)**

$$N=1000, n=500, \mu_y=5, \sigma_y^2=2, \mu_s=0, \sigma_s^2=20, \mu_T=1, \sigma_T^2=10, A=0.3, B=0.5$$
$$W=\left\{0.80, 0.847, 0.867, 0.890, 0.917, 0.92, 0.93, 0.942, 0.957, 0.962\right\}$$

**Population 2 : (Source: Azeem et al. (2024)**

$N = 1000, n = 500, \mu_y = 5, \sigma_y^2 = 2, \mu_s = 0, \sigma_s^2 = 20, \mu_T = 1, \sigma_T^2 = 10, A = 0.5, B = 0.3$

$W = \{0.80, 0.847, 0.867, 0.890, 0.917, 0.92, 0.93, 0.942, 0.957, 0.962\}$

Sensitivity analysis was considered using A=0.3 and B=0.5, as in the case of population 1 and then A=0.5 and B=0.3, as in the case of population 2. These constants A and Bare chosen based on the choice of the interviewer.

**Table 1: Results of variance of Gupta et al. (2022), Azeem et al. (2024) and proposed model using population 1**

| $W$ | $Var(Z_{Gi})$ | $Var(Z_{AZi})$ | $Var(Z_{ymi})$ |
|---|---|---|---|
| 0.800 | 0.384 | 0.3224 | 0.3000 |
| 0.847 | 0.3580 | 0.3411 | 0.3174 |
| 0.867 | 0.3664 | 0.3491 | 0.3248 |
| 0.890 | 0.3760 | 0.3582 | 0.3333 |
| 0.917 | 0.3873 | 0.3690 | 0.3433 |
| 0.920 | 0.3886 | 0.3702 | 0.3444 |
| 0.930 | 0.3927 | 0.3741 | 0.3481 |
| 0.942 | 0.3978 | 0.3789 | 0.3525 |
| 0.957 | 0.4040 | 0.3849 | 0.3581 |
| 0.962 | 0.4061 | 0.3869 | 0.3599 |

**Table 2: Results of variance of Gupta et al. (2022), Azeem et al. (2024) and proposed model using population 2**

| $W$ | $Var(Z_{Gi})$ | $Var(Z_{AZi})$ | $Var(Z_{ymi})$ |
|---|---|---|---|
| 0.800 | 0.2520 | 0.2424 | 0.2200 |
| 0.847 | 0.3666 | 0.2564 | 0.2327 |
| 0.867 | 0.2727 | 0.2624 | 0.2381 |
| 0.890 | 0.2799 | 0.2692 | 0.2443 |
| 0.917 | 0.2883 | 0.2773 | 0.2516 |
| 0.920 | 0.2892 | 0.2782 | 0.2524 |

| | | | |
|---|---|---|---|
| 0.930 | 0.2923 | 0.2811 | 0.2551 |
| 0.942 | 0.2960 | 0.2847 | 0.2583 |
| 0.957 | 0.3007 | 0.2892 | 0.2629 |
| 0.962 | 0.3022 | 0.2907 | 0.2637 |

**Table 3: Results of privacy level of Gupta et al. (2022), Azeem et al. (2024) and proposed model using population 1**

| $W$ | $\Delta_{Gi}$ | $\Delta_{AZi}$ | $\Delta_{ymi}$ |
|---|---|---|---|
| 0.800 | 169.2 | 161.2 | 150 |
| 0.847 | 179 | 170.55 | 158.7 |
| 0.867 | 183.2 | 174.55 | 162.4 |
| 0.890 | 188 | 179.1 | 166.65 |
| 0.917 | 193.65 | 184.5 | 171.65 |
| 0.920 | 194.3 | 185.1 | 172.2 |
| 0.930 | 196.35 | 187.05 | 174.05 |
| 0.942 | 198.9 | 189.45 | 176.25 |
| 0.957 | 202 | 192.45 | 179.05 |
| 0.962 | 203.05 | 193.45 | 179.95 |

**Table 4: Results of privacy level of Gupta et al. (2022), Azeem et al. (2024) and proposed model using population 2**

| $W$ | $\Delta_{Gi}$ | $\Delta_{AZi}$ | $\Delta_{ymi}$ |
|---|---|---|---|
| 0.800 | 126 | 121.2 | 110 |
| 0.847 | 133.3 | 128.2 | 116.35 |
| 0.867 | 136.65 | 131.2 | 119.05 |
| 0.890 | 139.95 | 134.6 | 122.15 |
| 0.917 | 144.15 | 138.65 | 125.8 |

| | | | |
|---|---|---|---|
| 0.920 | 144.6 | 139.1 | 126.2 |
| 0.930 | 146.15 | 140.55 | 127.55 |
| 0.942 | 148 | 142.35 | 129.15 |
| 0.957 | 150.35 | 144.6 | 131.45 |
| 0.962 | 151.1 | 145.35 | 131.85 |

**Table 5: Results of combined metric of efficiency and privacy level of Gupta et al. (2022), Azeem et al. (2024) and proposed model using population 1**

| $W$ | $\delta_{Gi}$ | $\delta_{AZi}$ | $\delta_{ymi}$ |
|---|---|---|---|
| 0.800 | 0.002 | 0.002 | 0.002 |
| 0.847 | 0.002 | 0.002 | 0.002 |
| 0.867 | 0.002 | 0.002 | 0.002 |
| 0.890 | 0.002 | 0.002 | 0.002 |
| 0.917 | 0.002 | 0.002 | 0.002 |
| 0.920 | 0.002 | 0.002 | 0.002 |
| 0.930 | 0.002 | 0.002 | 0.002 |
| 0.942 | 0.002 | 0.002 | 0.002 |
| 0.957 | 0.002 | 0.002 | 0.002 |
| 0.962 | 0.002 | 0.002 | 0.002 |

**Table 6: Results of combined metric of efficiency and privacy level of Gupta et al. (2022), Azeem et al. (2024) and proposed model using population 2**

| $W$ | $\delta_{Gi}$ | $\delta_{AZi}$ | $\delta_{ymi}$ |
|---|---|---|---|
| 0.800 | 0.002 | 0.002 | 0.002 |
| 0.847 | 0.002 | 0.002 | 0.002 |
| 0.867 | 0.002 | 0.002 | 0.002 |
| 0.890 | 0.002 | 0.002 | 0.002 |

| | | | |
|---|---|---|---|
| 0.917 | 0.002 | 0.002 | 0.002 |
| 0.920 | 0.002 | 0.002 | 0.002 |
| 0.930 | 0.002 | 0.002 | 0.002 |
| 0.942 | 0.002 | 0.002 | 0.002 |
| 0.957 | 0.002 | 0.002 | 0.002 |
| 0.962 | 0.002 | 0.002 | 0.002 |

## 5.1 Discussion of Results

Tables 1 and 2 show the variances of Gupta et al. (2022), Azeem et al. (2024) and the proposed model using populations 1 and 2. It can be seen from Tables 1 and 2 that the variance of the proposed model is smaller than the variances of the Gupta et al. (2022) and Azeem et al. (2024) models at various values of W, A and B. This implies that the proposed model is more efficient; that is, it is the model with the smaller variance. Theoretically, the proposed model performs better than Gupta et al. (2022) and Azeem et al. (2024) with the introduction of the fourth option Y+ST, instead of TY+S in the model of Azeem et al. (2024). This modification reduced the variance of the proposed model.

Tables 3 and 4 show the privacy levels of Gupta et al. (2022), Azeem et al. (2024) and the proposed model using populations 1 and 2. It can be seen from Tables 3 and 4 that the privacy level of the proposed model is smaller than the privacy levels of Gupta et al. (2022) and Azeem et al. (2024) models at various values of W, A and B. It can also be seen that the privacy level of Azeem et al. (2024) is smaller than that of Gupta et al. (2022). The practical implication of the reduced privacy level of the proposed model is that the model cannot conceal the sensitive information collected from sensitive surveys as compared to the Gupta et al. (2022) and Azeem et al. (2024) models. Theoretically, the proposed model's privacy level is reduced as a result of the introduction of the fourth option Y+ST, instead of TY+S in the model of Azeem et al. (2024).

Tables 5 and 6 show combined metrics of Gupta et al. (2022), Azeem et al. (2024), and the proposed model using populations 1 and 2. It can be seen from Tables 5 and 6 that the combined metric of the proposed model is approximately the same as the combined metrics of Gupta et al. (2022) and Azeem et al. (2024) models at various values of W, A and B. Selecting the most appropriate model becomes challenging because all models exhibit identical combined metrics.
In the previous study, a trade-off exists: trying to improve efficiency leads to a reduced privacy level. The Gupta et al. (2022) model achieved a better privacy level (high privacy level), but poor efficiency (larger variance) when compared with the Azeem et al. (2024) model with a poor privacy level (lower privacy level) and better efficiency (lower variance). In the present study, the

Azeem et al. (2024) model achieved a better privacy level (high privacy level), but poor efficiency (large variance) when compared with the proposed model with a poor privacy level (lowest privacy level) and highly efficient (lowest variance). Hence, the implication of this reduced privacy and improved efficiency is that the model can only be applied in sensitive surveys that require a precise estimate of the sensitive study variable without considering much privacy of the model; this is a condition for using the proposed model.

## 6. Conclusion

The study developed an improved randomised response technique for seeking information about sensitive issues in a sensitive survey. The main aim of the study is to develop an efficient model, although the privacy level may be jeopardised, which is the limitation of the model. Also, the combined metrics of Gupta et al. (2022), Azeem et al. (2024) and the proposed model are the same. This means that the proposed model, Gupta et al. (2022) and Azeem et al. (2024) models are better models for collecting sensitive information. For investigating sensitive issues that require a better estimate, the proposed model is used, but if the priority of the survey is better privacy, Gupta et al. (2022) is recommended. Hence, the proposed model and its estimator are recommended for use in real-life surveys involving sensitive issues and when the study aims to obtain a precise estimate of the population mean of the sensitive study variable. This study can be replicated under the stratified sampling design, since the current study used a simple random sampling design.

**References**

Audu, A., Yunusa, M. A., Aphane, M., & Lekganyane, M. (2025). Calibrated optional randomized response techniques for efficient and robust estimation of quantitative sensitive variables. *Scientific Reports, 15*, 1–22.

Azeem, M., & Salam, A. (2024). A versatile optional randomized response technique for use with sensitive surveys. *VFAST Transactions on Mathematics, 12*(1), 176–188.

Bar-Lev, S. K., Bobovitch, E., & Boukai, B. (2004). A note on randomised response models for quantitative data. *Metrika, 60*, 255–260.

Cochran, W. G. (1940). The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce. *The Journal of Agricultural Science*, *30*, 262–275.

Diana, G., & Perri, P. F. (2010). New scrambled response models for estimating the mean of a sensitive quantitative character. *Journal of Applied Statistics*, *37*(11), 1875–1890.

Diana, G., & Perri, P. F. (2011). A class of estimators of quantitative sensitive data. *Statistical Papers, 52*(3), 633–650.

Eichhorn, B. H., & Hayre, L. S. (1983). Scrambled randomised response methods for obtaining sensitive quantitative data. *Journal of Statistical Planning and Inference, 7*(4), 307–316.

Gjestvang, C. R., & Singh, S. (2009). An improved randomised response model: Estimation of mean. *Journal of Applied Statistics, 36*(12), 1361–1367.

Gupta, S., Gupta, B., & Singh, S. (2002). Estimation of sensitivity level of personal interview survey questions. *Journal of Statistical Planning and Inference, 100*(2), 239–247.

Gupta, S., Shabbir, J., Mehta, S., & Khalil, S. (2018). A unified measure of respondent privacy and model efficiency in quantitative RRT models. *Journal of Statistical Theory and Practice, 12*(3), 506–511.

Gupta, S., Zhang, J., Khalil, S., & Sapra, P. (2022). Mitigating lack of trust in quantitative randomised response technique models. *Communications in Statistics: Simulation and Computation,* 1–9.

Narjis, G., & Shabbir, J. (2021). An efficient new scrambled response model for estimating sensitive population mean in successive sampling. *Communications in Statistics: Simulation and Computation*, 1–18.

Odumade, O., & Singh, S. (2010). An alternative to the Bar-Lev, Bobovitch and Boukai randomised response model. *Sociological Methods & Research, 39*, 206–221.

Saha, A. (2007). A simple randomized response technique in complex surveys. *Metron, 65*, 59–66.

Saha, A. (2008). A randomised response technique for quantitative data under unequal probability sampling. *Journal of Statistical Theory and Practice, 2*(4), 589–596.

Singh, H. P., & Gorey, S. M. (2019). A remark on Gupta, Gupta and Singh optional randomised response model. *Journal of Advances in Mathematics and Statistics, 15*(1), 43–73.

Warner, S. L. (1965). Randomised response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association, 60*(309), 63–69.

Warner, S. L. (1971). The linear randomised response model. *Journal of the American Statistical Association, 66*(336), 884–888.

Yan, Z., Wang, J., & Lai, J. (2008). An efficiency and protection degree-based comparison among quantitative randomised response strategies. *Communications in Statistics: Theory and Methods, 38*(3), 400–408.

Yunusa, M. A., & Adejumobi, A. (2024). Some log-type estimators of population mean of a sensitive variable using an additive scrambling model. *Oriental Journal of Physical Sciences, 9*(2), 1–10.

Yunusa, M. A., Audu, A., & Adejumobi, A. (2025a). Robust optional randomised response technique models (RORRT) for seeking information about sensitive issues. *International Journal of Scientific Research in Mathematical and Statistical Sciences, 12*(3), 1–7.

Yunusa, M. A., Audu, A., & Alafara, A. A. (2025b). Development of some randomised response technique (RRT) models with better efficiency and privacy level. *International Journal of Scientific Research in Mathematical and Statistical Sciences, 12*(3), 97–101.