Journal of Survey Statistics and Methodology (2026) ***00**, 1–28*

THE EXTENDED CROSSWISE MODEL ADJUSTED FOR RANDOM ANSWERING

KHADIGA H. A. SAYED*
MAARTEN J. L. F. CRUYFF
ANDREA PETRÓCZI
PETER G. M. VAN DER HEIJDEN

The Extended Crosswise Model is a popular randomized response design that employs a sensitive and innocuous statement, and asks respondents if one of these statements is true, or if none or both are true. Although the model has a degree of freedom, it is unable to detect random answering. In this article, we propose a new method to detect and correct for random answering. This method makes use of a non-sensitive control statement and a quasi-randomized innocuous statement to which both answers are known, which allows for the detection of and correction for random answering. A simulation study shows that this method yields unbiased estimates of the prevalence of sensitive attribute. For four surveys among elite athletes, we present prevalence estimates of doping use that are corrected for random answering.

KHADIGA H. A. SAYED, MAARTEN J. L. F. CRUYFF, and PETER G. M. VAN DER HEIJDEN is with the Department of Methodology and Statistics, Utrecht University, Utrecht, The Netherlands; KHADIGA H. A. SAYED is with Department of Statistics, Faculty of Economics and Political Science, Cairo University, Egypt; ANDREA PETRÓCZI is with Kingston University London, School of Life Sciences, Pharmacy and Chemistry, Faculty of Health, Science, Social Care and Education, London, UK; ANDREA PETRÓCZI is with ELTE Eötvös Loránd University, Institute of Health Promotion and Sport Sciences, Faculty of Education and Psychology, Budapest, Hungary; PETER G. M. VAN DER HEIJDEN is with Department of Social Statistics and Demography, University of Southampton, Southampton, UK
The authors thank the Editor in Chief, the Associate editor and two anonymous reviewers for their helpful comments and suggestions.
The author(s) received no financial support for the research.
A.P. and M.J.L.F.C. are members of the WADA Doping Prevalence Working Group. A.P. also chaired the Working Group between 2017 and 2023. Working Group members receive no salary for their work for WADA, but are entitled to expenses covered and to receive indemnity for formal meetings and up to five days per year for preparation.
*Address for correspondence to Khadiga Sayed, Department of Methodology and Statistics, Utrecht University, Padualaan 14, 3584 CH Utrecht, The Netherlands; E-mail: k.h.a.sayed@uu.nl; khadigaali@feps.edu.eg.

https://doi.org/10.1093/jssam/smag004

KEY WORDS: Doping use; Goodness-of-fit test; Innocuous statement; Number sequence randomizer; Randomized response; Sensitive attribute.

Statement of Significance

The extended crosswise model is a popular indirect interview design for surveying sensitive topics while protecting respondents' privacy. In this article, we addressed the issue of random answering. Random answering biases the prevalence estimates of the sensitive attribute towards 50 percent, but cannot be detected by the goodness-of-fit test. This article proposes a new method to estimate and correct for random answering by estimating the prevalence of random responders using a non-sensitive control statement for which both the true and randomized answers are known at the individual level. The method is illustrated on data from four surveys on doping use by elite athletes. The results suggest that correcting for random responding can substantially reduce the prevalence estimates for a socially undesirable attribute. However, evaluating the validity of these corrected estimates requires validation studies, in which the true prevalence of the sensitive attribute is known.

1. INTRODUCTION

The randomized response technique is an interview method introduced by Warner (1965) to eliminate evasive response bias when sensitive questions have to be asked. The main idea of this method is that respondents become more inclined to answer sensitive questions truthfully when they believe that their privacy is protected and their individual responses cannot be traced back to them. It involves the use of a randomizer (e.g., a die or a spinner) that perturbs respondents' answers so that no direct link can be established with their true, non-randomized answers. Two meta-analyses (Lensvelt-Mulders et al. 2005; Sagoe et al. 2021) have shown that RR tends to yield more valid responses than direct questioning. However, it has been shown that Warner's design and many of its variants do not completely eliminate evasive response bias due to respondents who give the non-incriminating response irrespective of the outcome of the randomizer (Boeije and Lensvelt-Mulders 2002; Böckenholt et al. 2009). For this reason randomized response designs have been proposed that avoid the use of incriminating responses. The most well-known are the crosswise model (CWM, Yu et al. 2008) and its extension, the extended CWM

(ECWM, Heck et al. 2018). In these designs, respondents are shown two statements: one about a sensitive attribute with unknown prevalence, for example, "I have used illegal drugs," and an innocuous one with known prevalence, for example, "I was born in January or February." Respondents are instructed to indicate whether their answer is "DIFFERENT," that is, one "yes" and one "no" answer, or "SAME," that is, two "yes" or two "no" answers. Unlike designs such as the unrelated question model (Greenberg et al. 1969) and the triangular design (Meisters et al. 2022; Hsieh et al. 2024), these neutral answer options make it easier to give an honest answer, while at the same time they make it more difficult to infer the incriminating answer, and thus to give a evasive answer. The statistical model of the CWM is saturated, that is, it has only one non-redundant randomized response proportion to estimate the prevalence of the sensitive attribute, and therefore does not allow for a goodness-of-fit test. The ECWM extends the CWM by randomly splitting the sample into two non-overlapping sub-samples with complementary probabilities of answering "yes" to the innocuous statement. The model for this design has a degree of freedom that allows for a goodness-of-fit test. Despite these advantages, the additional complexity of the (E)CWM instructions may make it vulnerable to forms of response bias, such as random answering, which can distort prevalence estimates of sensitive attributes.

The (E)CWM has been investigated in several validation studies of socially undesirable attributes such as plagiarism (Coutts et al. 2011; Jann et al. 2012; Hopp and Speil 2019), tax evasion (Korndörfer et al. 2014), xenophobia and Islamophobia (Hoffmann and Musch 2016; Hoffmann et al. 2020; Meisters et al. 2020b), socially desirable attributes such as personal hygiene behaviour during the COVID-19 pandemic (Mieth et al. 2021), and voluntary work in the social sector (Meisters et al. 2023). Compared to the prevalence estimates obtained with the direct questioning method, the prevalence estimates of the (E)CWM were higher for undesirable attributes and lower for desirable ones, and therefore considered more valid according to the "more/less-is-better" criterion (Umesh and Peterson 1991; Sagoe et al. 2021; Schnell et al. 2021). These validation studies provided evidence for the validity of the (E)CWM with a birthday randomizer. In a systematic review on the (E)CWM applications, Sagoe et al. (2021) highlighted some concerns of using the birthday randomizer, for instance, using birth dates as "innocuous question" may undermine the perceived anonymity of the design and reduce its effectiveness in eliciting truthful responses. To address these limitations, Sayed et al. (2022) proposed a refinement of the ECWM that replaces the birthday randomizer with a number-sequence randomizer. This randomizer avoids the use of personal information, enhances the perceived anonymity of the response process, and reduces biases related to informed self-protection or the non-uniform distribution of birthdays. However, the (E)CWM is considered be more prone to random answering than other randomized response designs. The reason for this is that the meaning of its answer categories "DIFFERENT" and "SAME" is not well understood by the respondents, and may therefore incite indifference. Another factor that may contribute to random responding is the symmetrical response format of the (E)CWM (i.e., both answer options are neutral and have no obvious incriminating connotations). This symmetrical property makes the design less susceptible to self-protective responses, and potentially more susceptible to random responding. In studies where respondents were asked directly if they had answered the (E)CWM statement randomly, 2 to 19 percent of the respondents admitted having answered randomly (Enzmann 2017; Schnapp 2019; Meisters et al. 2020a). However, these findings should be interpreted with caution. Respondents who report having answered "randomly" may not necessarily have chosen their answers with equal probability 50 percent, but rather haphazardly, inattentively, or as a result of misunderstanding the design instructions. To our knowledge, there is little direct cognitive-psychological evidence on what respondents actually do when they say they answered "randomly." As noted by Boeije and Lensvelt-Mulders (2002), qualitative cognitive-interview studies are essential for evaluating respondents' comprehension and execution of randomized-response tasks. Future cognitive research would be valuable to clarify the mechanisms underlying random answering in the (E)CWM.

Some studies suggest that the higher presumed validity of the (E)CWM prevalence estimates according to the "more/less is better criterion" may (partly) be explained by random answering (Höglinger et al. 2016; Höglinger and Jann 2018; Enzmann 2017; Höglinger and Diekmann 2017), as random answering biases the prevalence estimates towards 50 percent (Walzenbach and Hinz 2019). Meisters et al. (2023) investigated how the validity of the (E) CWM is affected by random answering by manipulating the direction of social desirability (undesirable vs. desirable) and the prevalence of sensitive attributes (high vs. low). While random responding cannot be ruled out completely, Meisters et al. (2023) found that its influence on prevalence estimates according to the "more/less-is-better" criterion was not substantial. The degree of freedom of the ECWM can detect systematic response biases such as preferring one answer option as being safer (Heck et al. 2018; Cruyff et al. 2024), but it is unable to detect random answering (Heck et al. 2018; Meisters et al. 2023). As a consequence, other solutions have been proposed to address the issue of random answering.

One such solution is to provide detailed instructions and use comprehension checks to improve respondents' understanding of the design instructions (Höglinger and Diekmann 2017; Höglinger and Jann 2018; Meisters et al. 2020a). This method presupposes that random answering can be explained by (a lack of) comprehension and is not associated with the sensitive attribute, but it does not allow individual identification of random responders. Respondents who fail comprehension checks may have misunderstood the instructions rather than answered randomly, and conversely, some random responders may still pass comprehension checks by chance. Enzmann (2017) and Schnapp (2019) suggested asking the respondents whether they answered the statements at random, and adjusting the (E)CWM estimate accordingly. This approach also allows for the individual identification of random responders, but it presupposes that random responders do not answer this question randomly. Alternatively, Atsusaka and Stevenson (2023) suggested estimating the prevalence of random responders using a sensitive anchor statement with known prevalence. For instance, for sensitive statement of interest is “In order to avoid paying a traffic ticket, I would be willing to pay a bribe to a police officer,” the anchor statement is “I have paid a bribe to be on the top of a waiting list for an organ transplant,” with a known prevalence of zero. This anchor statement uses the same format as the CWM (i.e., it is paired with an innocuous statement with known prevalence). Due to the CWM format of the anchor statement, however, individual identification of random responders is no longer possible. Consequently, the estimation of the prevalence of random answering is less efficient than in the previously mentioned approaches. Additionally, a challenge of this method is to formulate a relevant sensitive anchor statement that matches the sensitive statement of interest.

The Present Study

In this article, we propose a new method to estimate and correct for random answering in the (E)CWM design. This method involves a non-sensitive control statement with known prevalence of 0 or 1 in combination with an innocuous statement with the probability of a “yes” answer set to either 0 or 1. To avoid confusion, we distinguish clearly between two types of non-sensitive statements used in the ECWM framework; (i) control statement: a non-sensitive statement with a known answer at the individual level. It is used to check respondents’ understanding and adherence to the ECWM instructions, and it allows for estimating the prevalence of random responding, and (ii) innocuous statement: a neutral statement without a sensitive connotation, such as the number sequence randomizer or “I was born in March or April.” The innocuous statement is paired with the sensitive statement of interest to form the crosswise question. To mimic the ordinary ECWM procedure as closely as possible, the suggestion is raised that the answers to the innocuous statement are randomized. This is achieved by using the number sequence randomizer (Sayed et al. 2022), which asks respondents to memorize one number from a sequence of five numbers and to indicate whether this number reappears in a second sequence of five, in which either all or none of the numbers from the first sequence reappear (e.g., see section 5). This method has three major advantages over the previously mentioned methods. First, it tests the comprehension of the ECWM answer procedure by using an ECWM question. Second, in contrast to Enzmann-Schnapp method, the use of a non-sensitive control statement eliminates the risk of social desirability bias. Third, the fact that the true answers to the control question are known allows us to directly check their correctness on the individual level. By attributing the proportion of incorrect "DIFFERENT" or "SAME" answers to random answering, a more efficient estimate of random answering is obtained than the estimates of the previously mentioned methods.

To illustrate the new method, it is applied to data from four ECWM surveys on doping use by elite athletes. These data were analyzed before by Cruyff et al. (2024), who found convincing evidence for *one-saying*. The term one-saying is derived from the answer option "I have One 'yes' and One 'no' answer" used in some of the surveys instead of the equivalent answer option "DIFFERENT". In order to take one-saying into account, we develop our method for both the standard ECWM and the one-sayers model.

The article is structured as follows. Section 2 reviews the ECWM, the one-sayer model and random answering. Section 3 derives the method to correct for random answering, and presents the moment and maximum likelihood estimators of π. Section 4 presents a simulation study to evaluate the proposed method. Section 5 provides a description of the data. It also presents an adjusted procedure to account for random answering because the control statements in these surveys contain measurement errors. Section 6 presents the prevalence estimates of doping corrected for random answering. Section 7 discusses the results of our analyses and ends with some concluding remarks.

2. THE MODELS

This section reviews the ECWM and the one-sayers model using matrix notation, and extends both models to account for random answering. Before presenting the models, we describe our assumptions with respect to one-sayers and random responders in detail.

By one-sayers, we mean respondents who seriously and attentively answer the survey questions but who, when the questions are sensitive, have a tendency to answer evasively. Given that in the ECWM design with a number sequence randomizer it is practically impossible to infer which of the two response options is the incriminating one, one-sayers will select the response that they perceive as the non-incriminating one. Cruyff et al. (2024) provided empirical evidence that this is the "DIFFERENT"/"I have One 'yes' and One 'no' answer" option.

By random responders, we mean respondents who either do not answer the survey questions seriously and/or attentively or do not comprehend the answer instructions of the ECWM sufficiently well to know which answer they should give. In an attempt to find evidence for the former category of random responders, we conducted a series of logistic regressions predicting the probability of doping use from the time it took the respondents to complete the survey. Given that random answering biases these probability estimates towards 50 percent, we expect fast respondents to be more likely to be random responders than those who took their time. Figure 1 presents the estimated prevalence of doping use across completion time quantile bins (blue points), along with the fitted logistic regression curves (red lines). It shows the empirical variability in the data while also showing the smooth trend implied by the logistic regression model: the fast respondents have higher estimated probabilities of doping use than the slower ones in all four surveys. This pattern is consistent with our hypothesis that random responding inflates prevalence estimates among fast respondents, but it does not imply that random answering is

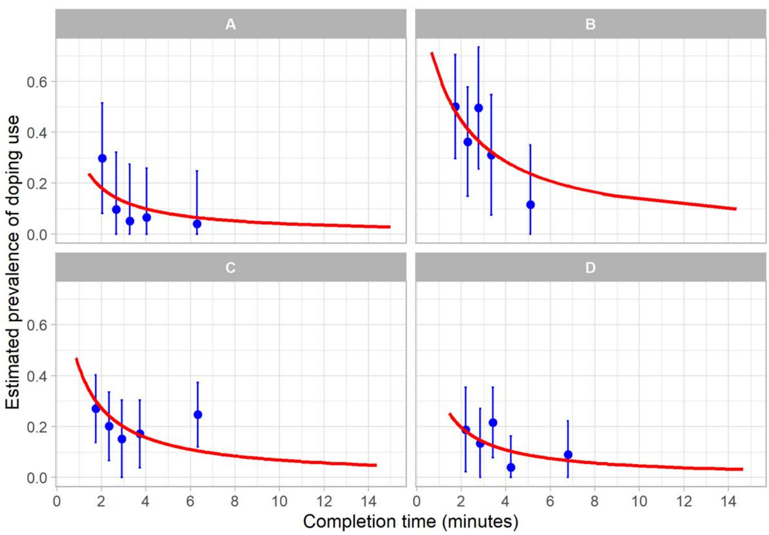

Figure 1 The Estimated Prevalence of Doping Use Across Completion Time Quantile Bins (Blue Points with 95 Percent Confidence Intervals), Together with Fitted Logistic Regression Curves (Red Lines), Shown Separately For Each Survey Event. Bins are based on quantiles of completion time, so that each bin contains approximately the same number of respondents.

the underlying bias mechanism; alternative explanations, such as cognitive or motivational differences, may also play a role. In the article by Cruyff et al. (2024), a similar effect for the doping data was found by analyzing the same dataset with a logistic regression that included a one-sayer parameter and completion time as a covariate for both prevalence and one-saying.

In the models presented below, one-sayers and random responders are treated as two mutually exclusive categories. The motivation for this is that one-sayers are assumed to answer the question seriously, but only edit their response when the questions are perceived as sensitive. Random responders on the other hand are assumed to give random answers to all CWM questions, irrespective of the sensitivity of the question.

2.1 The ECWM

Consider a CWM with a sensitive statement with unknown prevalence π and an innocuous statement with known randomization probability p of answering it with "yes" (e.g. "Is your birthday in the first two months of the year?"). Let π_y^* be the probability of observing randomized response y, for $y \in \{1 \equiv \text{"DIFFERENT"},\ 2 \equiv \text{"SAME"}\}$, π the prevalence of the sensitive attribute, and q the randomization probability of answering the innocuous statement with "yes," for $q = 1 - p \neq .5$. The model consists of the vectors $\boldsymbol{\pi}^*$ with the randomized response probabilities and $\boldsymbol{\pi}$ with the probabilities that the sensitive attribute is present or absent, and a 2×2 transition matrix with the randomization probabilities:

$$\begin{pmatrix} \pi_1^* \\ \pi_2^* \end{pmatrix} = \begin{pmatrix} p & q \\ q & p \end{pmatrix} \begin{pmatrix} \pi \\ 1-\pi \end{pmatrix} = \begin{pmatrix} p\pi + q(1-\pi) \\ q\pi + p(1-\pi) \end{pmatrix}. \tag{1}$$

The ECWM divides the sample into two sub-samples with the respective complementary probabilities p and q of answering "yes" to the innocuous statement. Let $\pi_{y|s}^*$ be the conditional probability of observing randomized response y given membership of sub-sample s, for $s \in \{1, 2\}$. The ECWM is given by

$$\begin{pmatrix} \pi_{1|1}^* \\ \pi_{2|1}^* \\ \pi_{1|2}^* \\ \pi_{2|2}^* \end{pmatrix} = \begin{pmatrix} p & q \\ q & p \\ q & p \\ p & q \end{pmatrix} \begin{pmatrix} \pi \\ 1-\pi \end{pmatrix} = \begin{pmatrix} p\pi + q(1-\pi) \\ q\pi + p(1-\pi) \\ q\pi + p(1-\pi) \\ p\pi + q(1-\pi) \end{pmatrix}. \tag{2}$$

Eq. (2) shows that $\pi^*_{1|1} = \pi^*_{2|2}$ and $\pi^*_{2|1} = \pi^*_{1|2}$. The model estimates the parameter π using two non-redundant observed response frequencies (as the conditional probabilities $\pi^*_{y|s}$ within each sub-sample s need to sum to 1) and therefore has one degree of freedom.

2.2 The ECWM Correcting for Random Answering

To account for random responders in the ECWM, let γ denote the prevalence of random responders. Assuming equal probabilities of random answering for all randomized responses, the model is given by

$$\begin{pmatrix} \pi^*_{1|1} \\ \pi^*_{2|1} \\ \pi^*_{1|2} \\ \pi^*_{2|2} \end{pmatrix} = \begin{pmatrix} (1-\gamma)p + .5\gamma & (1-\gamma)q + .5\gamma \\ (1-\gamma)q + .5\gamma & (1-\gamma)p + .5\gamma \\ (1-\gamma)q + .5\gamma & (1-\gamma)p + .5\gamma \\ (1-\gamma)p + .5\gamma & (1-\gamma)q + .5\gamma \end{pmatrix} \begin{pmatrix} \pi \\ 1-\pi \end{pmatrix}. \qquad (3)$$

We refer to this model as "ECWM+RA," where RA stands for random answer. In (3) the equality relations $\pi^*_{1|1} = \pi^*_{2|2}$ and $\pi^*_{2|1} = \pi^*_{1|2}$ are not affected by random answering. This means that multiple combinations of γ and π give the same parameter vector $\boldsymbol{\pi}^*$, and therefore the model is not identified, as was noted before by Heck et al. (2018).

2.3 The One-Sayers Model Correcting for Random Answering

The one-sayers model (Cruyff et al. 2024) accounts for evasive respondents who answer "DIFFERENT" (or equivalently "I have One 'yes' and One 'no' answer"), irrespective of the outcome of the randomizer. With θ denoting the prevalence of one-sayers, the model is given by

$$\begin{pmatrix} \pi^*_{1|1} \\ \pi^*_{2|1} \\ \pi^*_{1|2} \\ \pi^*_{2|2} \end{pmatrix} = \begin{pmatrix} (1-\theta)p + \theta & (1-\theta)q + \theta \\ (1-\theta)q & (1-\theta)p \\ (1-\theta)q + \theta & (1-\theta)p + \theta \\ (1-\theta)p & (1-\theta)q \end{pmatrix} \begin{pmatrix} \pi \\ 1-\pi \end{pmatrix}, \qquad (4)$$

This model shows that $\pi^*_{1|1} \neq \pi^*_{2|2}$ and $\pi^*_{2|1} \neq \pi^*_{1|2}$ for $\theta > 0$. Now each combination of θ and π yields a unique parameter vector $\boldsymbol{\pi}^*$, and therefore the model is identified.

The one-sayers model that additionally accounts for random answering (One-sayers + RA) is given by

$$\begin{pmatrix} \pi^*_{1|1} \\ \pi^*_{2|1} \\ \pi^*_{1|2} \\ \pi^*_{2|2} \end{pmatrix} = \begin{pmatrix} (1-\gamma-\theta)p+\theta+.5\gamma & (1-\gamma-\theta)q+\theta+.5\gamma \\ (1-\gamma-\theta)q+.5\gamma & (1-\gamma-\theta)p+.5\gamma \\ (1-\gamma-\theta)q+\theta+.5\gamma & (1-\gamma-\theta)p+\theta+.5\gamma \\ (1-\gamma-\theta)p+.5\gamma & (1-\gamma-\theta)q+.5\gamma \end{pmatrix} \begin{pmatrix} \pi \\ 1-\pi \end{pmatrix}, \tag{5}$$

with the restriction that $\theta+\gamma\leq 1$ because one-saying and random answering are mutually exclusive response categories. A diagram of the ECWM accounting for the one-saying and random answer is depicted in figure 2. This model is over-parameterized, but it can be identified by fixing the parameter γ. To do so, a reasonable estimate of the prevalence of random responders has to obtained by other means (e.g., by the use of a control question).

2.4 Expected Bias Resulting from Non-Zero γ and θ Parameters

In the presence of one-saying and random responding, the standard ECWM estimator $\widehat{\pi}$ of model (2) yields a biased estimate of π (i.e., $\mathbb{E}(\widehat{\pi}) = \pi + Bias$). The size of this bias equals $(\theta+\gamma)(.5-\pi)$ (see appendix A on OSF (https://osf.io/ekcjb/?view_only=e5ad20e51f2c4b4ea3816d5281350782), please see the supplementary data online for a derivation). Figure 3 shows the expected estimates of π when response bias due to random answering and one-saying is not taken into account.

The top-left plot in figure 3 depicts the expected estimates of π when random answering is present but there is no one-saying. It shows that as the value of γ increases, the expected estimate of π approaches 50 percent. In the case that all respondents answer randomly, $\mathbb{E}(\widehat{\pi}) = 0.5$, irrespective of the value of π. The remaining three plots show the expected estimates of π for the prevalence of one-saying $\theta \in \{.1, .2, .3\}$ with the restriction that $\gamma+\theta\leq 1$, because one-saying and random answering are two mutually exclusive phenomena. This explains why fewer values of γ are shown with the increasing values of θ. The plots show similar patterns for γ as in the top-left panel, but with higher expected estimates of π as the value of γ increases. For instance when $\pi = .25$, the standard ECWM of (2) yields biased expected estimates of $\widehat{\pi} \in \{.31, .38, .44, .5\}$ for $\gamma \in \{.25, .5, .75, 1\}$, and $\theta = 0$. For $\gamma \in \{.25, .5, .75\}$, and $\theta = .1$, the ECWM yields biased expected estimates of $\widehat{\pi} \in \{.34, .4, .46\}$, respectively. In summary, the figure shows that both random answering and one-saying bias the prevalence estimates towards 0.5.

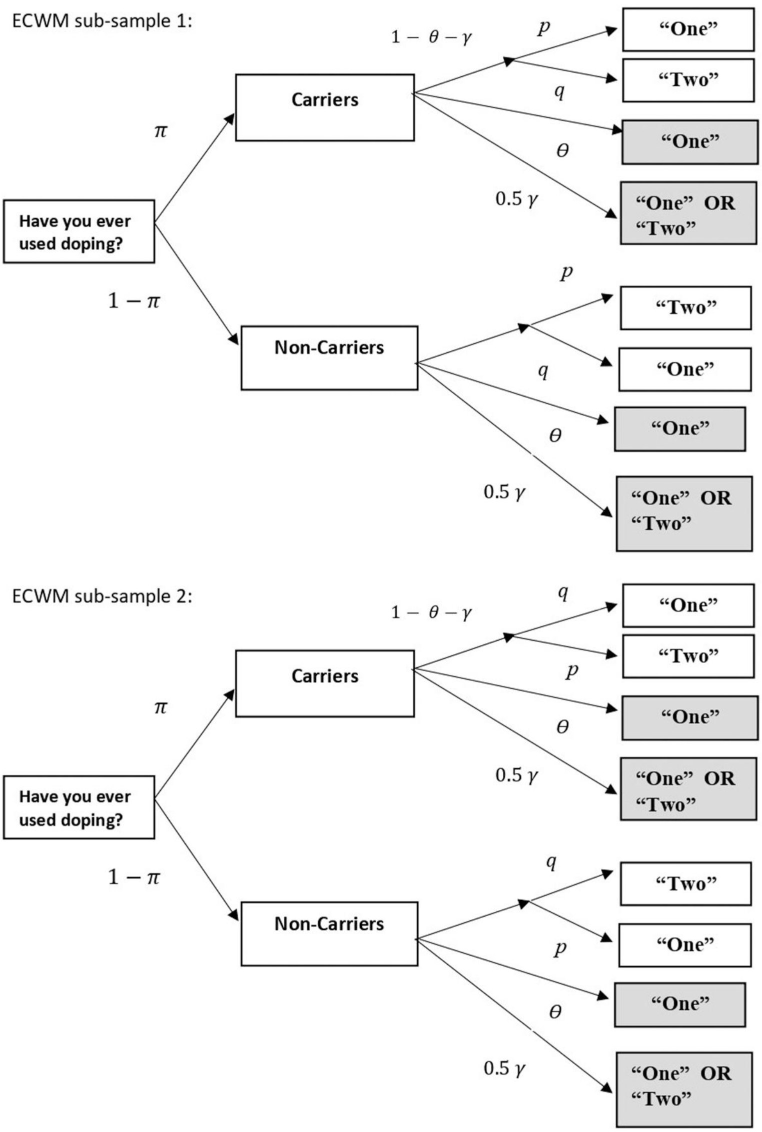

Figure 2 The one-Sayers Model Accounting for Random Answering, with Self-Protective "One" Answers And Random Answer In Gray. The ECWM with random answering is obtained as a special case of this model by setting the one-saying probability $\boldsymbol{\theta} = 0$; hence, no separate diagram is provided for that case. The parameter $\boldsymbol{\pi}$ represents the unknown prevalence of the sensitive attribute, q and p denote the known probabilities of answering "yes" to the innocuous question in sub-samples 1 and 2, $\boldsymbol{\theta}$ denotes the prevalence of one-sayers and $\boldsymbol{\gamma}$ represents the prevalence of random responders.

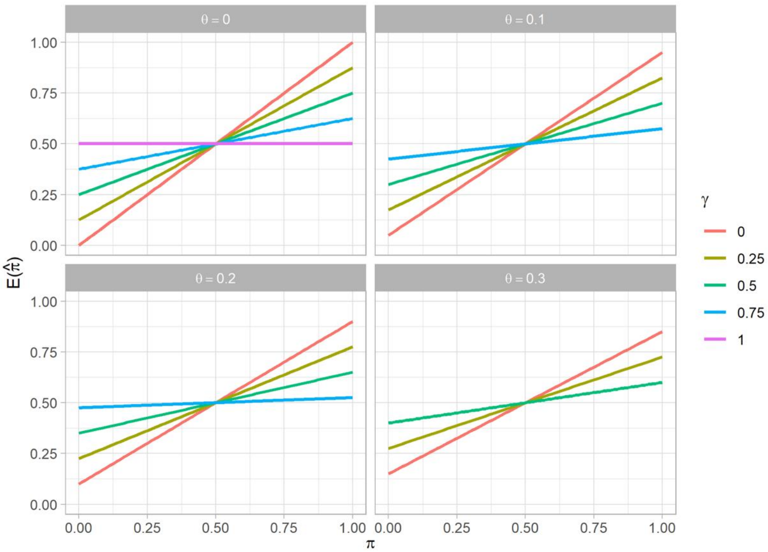

Figure 3 The Expectation $\mathbb{E}(\hat{\pi})$ Under the ECWM as a Function γ, θ and π.

3. ESTIMATION

This section presents moment and maximum likelihood estimates of the model parameters presented in the previous section.

3.1 Estimation of γ on the Basis of the Control Question

In this section, we derive our method to correct for random answering. It estimates its prevalence γ on the basis of the incorrect answers to the non-sensitive control question under the assumptions that (i) all respondents know the correct answer to the control statement, (ii) random responders answer all (E)CWM statements randomly, and (iii) random answering is independent of the sensitivity of the ECWM statement.

If all respondents know the correct answer to the control statement, the prevalence of random answering is estimated as $\hat{\gamma} = 2 \cdot e_c$, where e_c denotes the observed proportion of errors on the control question. The multiplication by two is necessary because on average half of the random responders will answer the control question correctly by chance.

Our estimate of γ is derived from a control statement about a non-sensitive characteristic with known prevalence of 0 or 1. Since this statement is not sensitive, the answers are assumed not to be affected by one-saying. The control statement is combined with a number sequence statement with probability of answering "yes" set to either 0 or 1, which makes it statistically equivalent to a direct question. Since the answers to both the control and the number sequence statement are known, the incorrect "DIFFERENT/SAME" answers can be individually identified and attributed random answering. Since random responders have 50 percent chance to have a correct "DIFFERENT/SAME" answer, our estimate of γ is twice the proportion of incorrect "DIFFERENT/SAME" answers.

In appendix C, please see the supplementary data we compare the power and the relative efficiency of our estimator of γ with respect to the Atsusaka and Stevenson (2023) and Schnapp (2019) estimators. The power curves show that our estimate of γ as twice the error rate of the control question: (i) it is about 3 to 20 times more efficient than the Atsusaka and Stevenson (2023) estimator when γ approaches .5, and .05, respectively, (ii) it is about 2 to 3 times more efficient than the Schnapp (2019) estimator, and (iii) it requires smaller sample sizes to test for γ with the same power.

3.2 Moment Estimation

The unidentified ECWM correcting for random answering (3) is identified by using a value for γ derived from the non-sensitive control statement. The estimator $\widehat{\pi}_{ra}$ of the sensitive attribute is given by

$$\widehat{\pi}_{ra} = \frac{\pi^*_{11} + \pi^*_{22} - (1-\gamma)q - .5\gamma}{(p-q)(1-\gamma)}. \tag{6}$$

where $\pi^*_{ys} = (n_s/n)\pi^*_{y|s}$ is the unconditional probability of observing response y in sub-sample s, n_s is the sub-sample size and $n = \sum_s n_s$ the total sample size. The moment estimator in (6) is computed by plugging in the observed sample proportions n_{ys}/n as estimates for π^*_{ys}, and $\widehat{\gamma}$ for γ. This estimator is identical to the ones presented by Schnapp (2019) and Atsusaka and Stevenson (2023), but the difference is in the method of estimating γ (see appendix C, please see the supplementary data for relative comparison).

For the one-sayers model correcting for random answering (5), the moment estimator of θ is

$$\widehat{\theta} = \pi^*_{1|1} + \pi^*_{1|2} - 1, \tag{7}$$

which is identical to the estimator of one-sayers in model (4), that is, the presence of random responders does not affect the estimate of one-sayers. The estimator $\widehat{\pi}_{one+ra}$ of the sensitive attribute is

$$\widehat{\pi}_{one+ra} = \frac{p\pi^*_{2|2} - q\pi^*_{2|1} - \gamma(p-.5)}{(p-q)(\pi^*_{2|1} + \pi^*_{2|2} - \gamma)}, \quad p \neq 0.5 \tag{8}$$

The analytical variances of the estimators $\widehat{\pi}_{ra}$ and $\widehat{\pi}_{ra+one}$ are presented in appendix A, please see the supplementary data on OSF (https://osf.io/ekcjb/?view_only=e5ad20e51f2c4b4ea3816d5281350782). Substituting $\gamma = 0$ in (6) and (8) yields the moment estimators of π of the standard ECWM of (2), and the one-sayers model of (4), respectively.

3.3 Maximum Likelihood Estimation

The parameters π and/or θ of models (3) and (5) can alternatively be estimated by maximization of the log-likelihood

$$\ln \ell(\pi, \theta \mid \boldsymbol{n}, \widehat{\gamma}) = \boldsymbol{n}' \ln \boldsymbol{\pi}^*, \tag{9}$$

where $\boldsymbol{n}$ is the vector with the observed randomized response frequencies n_{ys} corresponding to the elements $\pi^*_{y|s}$. If the model includes the parameter γ, then $\widehat{\gamma}$ is treated as a fixed value.

4. SIMULATION STUDY

This section presents a simulation study to evaluate the performance of the proposed method. In $10,000$ samples of $n = 5,000$ we drew "doping users" with prevalence $\pi \in \{.05, .1, .2, \}$ and, independently thereof, random responders with probability $\gamma \in \{.05, .1, .2\}$. For the control statement, we generated the responses given a true prevalence of 100 percent in combination with an innocuous statement with probabilities of a TRUE answer of zero in one sub-sample and of one in the other sub-sample to check the answers at the individual level. We then let the random responders answer the sensitive doping statement and the non-sensitive control statement randomly and let the remaining respondents answer according to the ECWM design. We then fitted the ECWM and ECWM-RA, for which we used twice the percentage of errors on the control statement as an estimate of γ.

Figure 4 depicts the estimates of the ECWM and ECWM-RA models, and shows that the latter estimates π unbiasedly, indicating that the estimates of γ as twice error rate of the control statement are unbiased and irrespective the values of π. appendix B, please see the supplementary data on OSF (https://osf.io/ekcjb/?view_only=e5ad20e51f2c4b4ea3816d5281350782) also includes some examples for smaller sample sizes. These show that the maximum likelihood estimator of (9) is biased in case that random answering results in estimates of π on the boundary. In that case the moment estimator of (6) is unbiased because it allows for negative estimates of π.

5. DATA

The data are from four surveys on doping use which were conducted as a part of the World Anti-Doping Agency (WADA) anti-doping program.

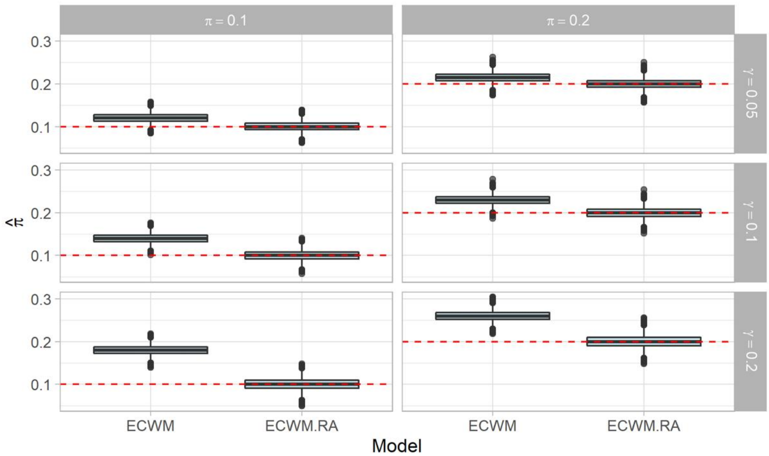

Figure 4 Box-Plots of the Estimates Under the ECWM, and the Model Corrects for Random Answering "ECWM-RA" At Different Values of $\gamma \in \{.05, .1, .2\}$.

A previous analysis of these data showed the presence of one-saying (Cruyff et al. 2024). Data analysis was approved by the Ethics Review Board of the Faculty of Social and Behavioural Sciences of Utrecht University in application 22–0185.

Procedure and Measures

Surveys A and D were accessed via a web-link and administered on an online platform, while Surveys B and C also offered the possibility of administration on mobile phones/tablets. The respondents in these surveys were elite athletes over the age of 16 years. The numbers of athletes that completed the surveys A to D were respectively 354, 325, 915, and 813. In these surveys, the sensitive statement "I have intentionally used a prohibited substance or method without a Therapeutic Use Exemption (TUE) in the last 12 months." was paired with the number sequence randomizer as the innocuous statement (Sayed et al. 2022). Figure 5 shows how this works. Respondents are first presented with a sequence of five randomly generated two-digit numbers, and are asked to memorize one. Then they are shown a second sequence of five numbers in which either one or four numbers of the first sequence reappear, with the innocuous statement B asking whether the memorized number is in this sequence. In figure 2 only one number reappears. The respondents are then asked to indicate whether their answers to these two statements are the "SAME" (both statements are true or false) or "DIFFERENT" (only one statement is true). In this example, the probability of answering "yes" to the

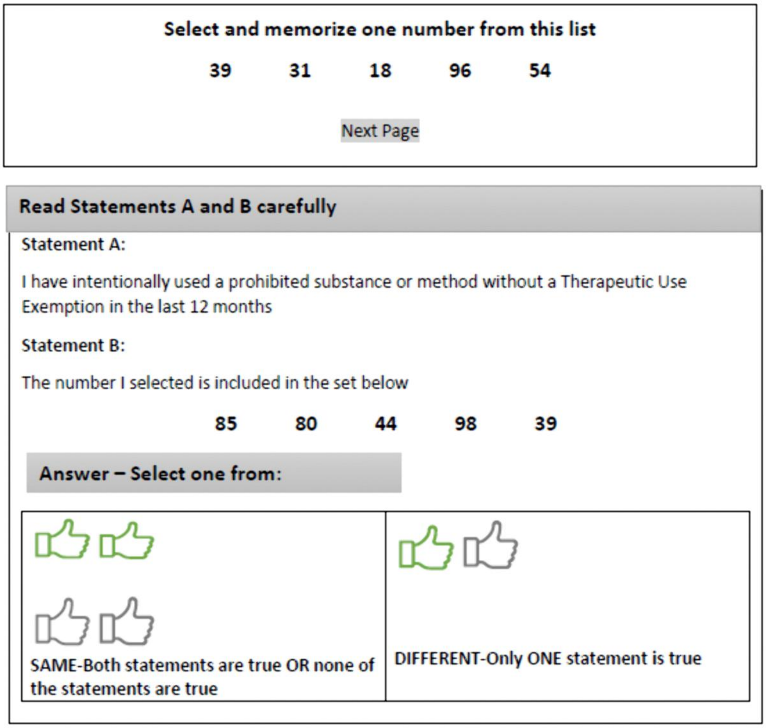

Figure 5 Example of the Number Sequence Randomizer for the Statement on Doping Use.

innocuous statement B is $p = 1/5$ because only one of the five numbers from the first sequence reappears in the second. In the surveys, respondents were randomly assigned to a condition with $p = 1/5$ (when one number reappears) or with $p = 4/5$ (when four numbers reappear).

To check the understanding and adherence to the ECWM instructions, the athletes were also presented with the control statement A "I am a licensed/accredited athlete" in combination with a number sequence statement B. The answers to statement B were quasi-randomized, because the probability of a "yes" answer set either to 1 by letting all the numbers of the first sequence reappear in the second sequence, or to 0 by letting none of the numbers reappear. The probability of a "yes" answer to the control question was assumed to be 1, because for all athletes in the surveys having a license/accreditation is mandatory. However, it appeared that the athletes were not always aware of being licensed/accredited. The reason for this is that the license/accreditation it is often arranged by their sports associations. Consequently, some of the athletes may have believed that they were not licensed/accredited, and may therefore have given the wrong answer to the control question.

Adjusting for Measurement Error in the Control Question

This section describes a procedure to correct for random answering when the correct answer of the control question is not known by some respondents. For the data at hand, we expect that some respondents answered the control question incorrectly because they did not know the correct answer (see section 5). If we denote the prevalence of these respondents by ϕ, then $e_c = .5\gamma + \phi$ (see appendix A, please see the supplementary data for a derivation) is a mixture of random respondents with prevalence γ and respondents who think they are not licensed/accredited with prevalence ϕ. Under the assumption that ϕ is independent of doping use, the conclusion is justified that ϕ has no effect on the validity of the answers to the doping statement. The problem then reduces to the estimation of γ.

The procedure to estimate γ is as follows. Let $\widehat{\pi}_{in}$ and $\widehat{\pi}_{out}$ respectively denote the doping prevalence estimate of the ECWM (2) with the respondents who answered the control question incorrectly in- and excluded from the data. Their exclusion implies that approximately half of the random respondents are eliminated from the data, since the other half answered the control question correctly by chance. If random answering is completely absent we expect that $\widehat{\pi}_{in} \approx \widehat{\pi}_{out}$. If not, the expectation is that $\widehat{\pi}_{in} > \widehat{\pi}_{out}$, because random answering biases the prevalence estimate toward 0.5. Since the exclusion of the incorrect answers to the control question excludes only half of the random responders, the difference between the two estimates $\Delta\widehat{\pi} = \widehat{\pi}_{in} - \widehat{\pi}_{out}$ is due to $.5\gamma$. Given the linearity of the relationship between $\mathbb{E}(\widehat{\pi})$ and γ as depicted in figure 3, the estimate $\widehat{\pi}_{ra}$ that is corrected for random answering is thus given by

$$\widehat{\pi}_{ra} = \widehat{\pi}_{out} - \Delta\widehat{\pi}. \tag{10}$$

The estimate of γ can be obtained by rearranging (6) to yield a closed-form expression for γ in terms of $\widehat{\pi}_{ra}$, the randomization probabilities p, q, and the observed unconditional probabilities π^*_{11} π^*_{22} as

$$\widehat{\gamma}^{ra} = \frac{\pi^*_{11} + \pi^*_{22} - (p-q)\widehat{\pi}_{ra} - q}{(p-q)\widehat{\pi}_{ra} - (p-0.5)}. \tag{11}$$

The same procedure is followed for the one-sayers model (4) by estimating $\widehat{\pi}_{one,\ in}$ and $\widehat{\pi}_{one,\ out}$ with the one-sayers model, and defining $\Delta\widehat{\pi}_{one} = \widehat{\pi}_{one,\ in} - \widehat{\pi}_{one,\ out}$ and $\widehat{\pi}_{ra+one} = \widehat{\pi}_{one,\ out} - \Delta\widehat{\pi}_{one}$. By rearranging (8), the estimate of γ is given by:

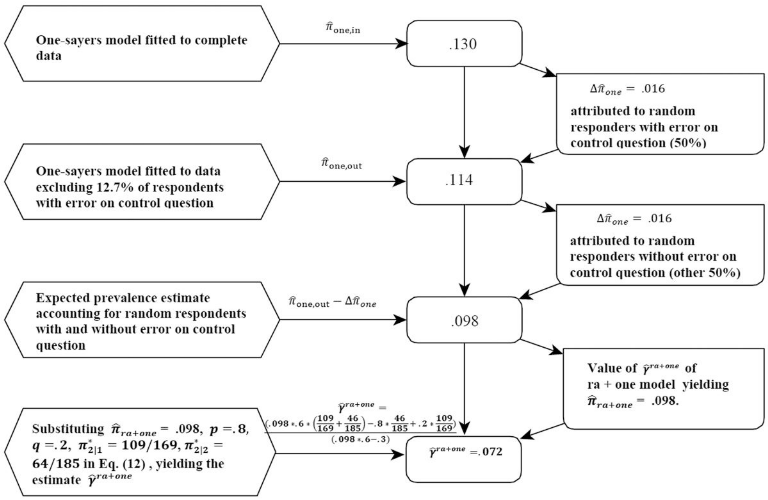

Figure 6 Schematic Representation of the Estimation of γ in Survey A.

$$\widehat{\gamma}^{ra+one} = \frac{\widehat{\pi}_{ra+one}(p-q)(\pi^*_{2|1} + \pi^*_{2|2}) - p\pi^*_{2|2} + q\pi^*_{2|1}}{\widehat{\pi}_{ra+one}(p-q) - (p-.5)}. \tag{12}$$

Given that the adapted procedure to estimate γ is rather complex; it is illustrated for Survey A in the next section with figure 6.

6. RESULTS

In this section, we present the prevalence estimates of doping use for the Surveys A to D of the ECWM (2), the one-sayers model (4), and the one-sayers model with correction for the control statement (5).

Before presenting the parameter estimates of these models, table 1 shows how the estimates of γ for the one-sayers model were obtained. The column $2e_c$ of table 1 shows twice the observed error rates on the control statement, which would have been our estimate of γ if all athletes would have known that they were accredited/licensed. The columns $\widehat{\pi}_{one,\ in}$ and $\widehat{\pi}_{one,\ out}$ show the prevalence estimates of doping use obtained with the one-sayers model, and $\Delta\widehat{\pi}_{one}$ is the difference between the two. As expected in the presence of random answering, the $\Delta\widehat{\pi}_{one}$ are all positive. The $\widehat{\pi}_{one+ra} = \widehat{\pi}_{one,\ out} - \Delta\widehat{\pi}_{one}$ are prevalence estimates corrected for random answering. The values of $\widehat{\gamma}$ are obtained by fitting the one-sayers model to the data with all respondents included, with $\widehat{\gamma}$ chosen such that these models yield $\widehat{\pi}_{ra+one}$. The $\widehat{\gamma}$ values are all smaller than $2e_c$, suggesting that part of incorrect answers to the control question are due to ignorance with respect to the accreditation/licensing.

Table 1. Estimation of $\widehat{\gamma}$ for the One-Sayers Model

	$2e_c$	$\widehat{\pi}_{one,\ in}$	$\widehat{\pi}_{one,\ out}$	$\Delta\widehat{\pi}_{one}$	$\widehat{\pi}_{ra+one}$	$\widehat{\theta}$	$\widehat{\gamma}$
A	0.254	0.130	0.114	0.016	0.098	0.106	0.072
B	0.270	0.357	0.331	0.026	0.306	0.115	0.233
C	0.394	0.213	0.194	0.019	0.176	0.131	0.099
D	0.268	0.129	0.126	0.003	0.124	0.088	0.013

To illustrate the procedure in more detail, figure 6 presents it schematically for Survey A. The first layer shows a prevalence estimate of the one-sayers model for the complete data of .130. The second layer shows an estimate of .114 when the respondents with an incorrect answer to the control question are deleted from the data. These include both random responders and respondents who did not know the correct answer. The difference of .016 is attributed to random answering, because the deletion of respondents who did not know the correct answer is not expected to have an effect on the prevalence estimate under the assumption that not knowing the correct answer to the control question is independent of doping use. The difference of .016 is assumed to be due to 50 percent of the random responders, because the other 50 percent answered the control statement correctly by chance. So the .114 estimate is adjusted for 50 percent of the random responders, and therefore the third layer adjusts for the 50 percent random responders who answer the control statement correctly. The doping prevalence estimate corrected for random answering is therefore $\widehat{\pi}_{ra+one} = .098$. In the fourth level the estimated prevalence of random responding $\widehat{\gamma}^{ra+one}$ is obtained directly by substituting $\widehat{\pi}_{ra+one} = .098$, $p = .8$, $q = .2$, and the observed conditional proportions $\pi^*_{2|1} = 109/169$, $\pi^*_{2|2} = 46/185$ into (12), yielding $\widehat{\gamma}^{ra+one} = .072$.

Table 2 shows the prevalence estimates of of doping of the three models. The goodness-of-fit statistics show that for none of the surveys the ECWM fits the data, and therefore we started by fitting the one-sayers model. This model is saturated, and therefore yields a zero G^2 statistic. The column with $\%\widehat{\pi}_{ECWM}$ shows the percentage reduction in the corrected estimates relative to the uncorrected estimates of the ECWM. For instance, in survey A the corrected estimate for one-saying and random answer is $(.098/.166) * 100 = 59\%$ of the uncorrected estimate of the ECWM. The correction for one-saying results in a reduction of the uncorrected estimates of approximately 5 percent to 20 percent. The model with corrections for one-saying and random answering yields prevalence estimates that are between 18.2 percent and 25.9 percent lower than the uncorrected estimates, but for Survey A the difference of 41 percent is much larger. Here 21.7 percent is due to one-saying, so that the further reduction of 19.3 percent is due to random answering.

Table 2. Uncorrected and Corrected Doping Prevalence Estimates $\widehat{\pi}$.

Survey	Model	$\widehat{\pi}$ (95% CI)	$\%\widehat{\pi}_{ECWM}$	$\widehat{\theta}$ (95% CI)	Goodness-of-fit
A	ECWM	0.166 (0.086, 0.246)	100		$G_1^2 = 4.8,\ p = .028$
	+ one-saying	0.130 (0.039, 0.222)	78.3	0.106 (0.010, 0.202)	$G_0^2 = 0$
	+ ra[a]	0.098 (0.000, 0.220)	59.0	0.106 (0.010, 0.202G	$G_0^2 = 0$
B	ECWM	0.374 (.285, 0.462)	100		$G_1^2 = 4.4,\ p = .036$
	+ one-saying	0.357 (0.225, 0.485)	95.4	0.115 (0.009, 0.222)	$G_0^2 = 0$
	+ ra[a]	0.306 (0.181, 0.431)	81.8	0.115 (0.009, 0.222)	$G_0^2 = 0$
C	ECWM	0.251 (0.200, 0.302)	100		$G_1^2 = 17.3,\ p < .001$
	+ one-saying	0.213 (0.152, 0.273)	84.9	0.131 (0.070, 0.192)	$G_0^2 = 0$
	+ ra[a]	0.176 (0.090, 0.262)	70.1	0.131 (0.070, 0.192)	$G_0^2 = 0$
D	ECWM	0.161 (0.108, 0.214)	100		$G_1^2 = 7.5,\ p = .006$
	+ one-saying	0.129 (0.069, 0.189)	80.1	0.088 (0.025, 0.151)	$G_0^2 = 0$
	+ ra[a]	0.124 (0.051, 0.197)	77.0	0.088 (0.025, 0.151)	$G_0^2 = 0$

[a]95 percent confidence intervals of $\widehat{\pi}$ obtained with the non-parametric bootstrap.

To account for the uncertainty in the estimate $\widehat{\gamma}$, the 95 percent confidence intervals of $\widehat{\pi}_{one+ra}$ are obtained with the non-parametric bootstrap. For $10,000$ bootstrap samples of the doping statement and the control question $\widehat{\pi}_{one+ra}$ are estimated in the same way as for the original data, and from these estimates the 95 percent confidence intervals are obtained with the percentile method.

7. DISCUSSION

In this study, we proposed a new method to estimate and correct for random answering of the (E)CWM. It estimates the prevalence of random responders by employing a non-sensitive control statement for which both the true and randomized answers are known at the individual level. This method identifies (half of the) random responders by their incorrect answers to the control statement. This may include inattentive respondents who answer fast, but also respondents who try their best to comprehend the instructions but fail to do so and thus take longer to complete the survey. The results of the simulation study show that correction for random answering yields unbiased estimates of the true prevalence of sensitive attribute if random answering is independent of the status on the sensitive attribute.

The analyses of the data have shown that our corrections for random answer invariably result in substantially lower prevalence estimates for a socially undesirable attribute (doping use) with a relatively low prevalence. These findings were obtained by fitting only the ECWM model with both random answering and one-saying (i.e., ECWM + RA + one-saying). We did not apply ECWM +RA to the survey data because empirical evidence for one-saying was found in all four surveys. Our results are in line with the previous studies (Höglinger et al. 2016; Enzmann 2017; Höglinger and Diekmann 2017; Höglinger and Jann 2018; Schnapp 2019), suggesting that higher prevalence estimates of sensitive attributes cannot be interpreted as evidence for a successful control of evasive response bias and random answering provides an alternative explanation for these high estimates. Unfortunately, the significance of these effects cannot be tested with goodness-of-fit tests, because the one-sayers model is a saturated model. That is, it perfectly reproduces the observed data, resulting in a test statistic of $G_0^2 = 0$. Introducing the random answering parameter γ alongside the one-saying parameter θ does not alter this statistic. However, this does not imply that the model fits the data adequately, it simply reflects that the model is saturated. This limitation should be taken into account when interpreting the corrected prevalence estimates. For the data at hand, the quality of the estimates may have been be negatively affected by measurement errors in the answers to the control statement due to the ignorance of being accredited/licensed. These measurement errors unnecessarily complicated the analyses but, as we will argue, can be easily avoided by taking the necessary precautions.

If all athletes would have been aware of being licensed/accredited, γ could have simply been estimated by twice the error rate e_c on the control statement. Since this was not the case, the rather complicated method presented in section 5 and table 1 had to be applied to distinguish between random responders and respondents who were unaware of being accredited/licensed. While this method seems to be valid under the assumption that the difference between the estimates with and without the incorrect answers to the control statement are due to random answering, it decreases the reliability of the γ estimates. These errors can be avoided by formulating an unambiguous, non-sensitive control statement. An example is to show respondents a picture of a street either with or without a bicycle in it, and ask them if there is a bicycle in the picture. Combined with a quasi-randomized number sequence statement this would leave no room for error.

A remarkable finding was the high error rate of 19.7 percent on the control question of Survey C, while in the other three surveys it ranges from 12.7 percent to 13.5 percent. A substantial part of the surveys was administered in the registration hall where the athletes collected their accreditation pass, so that it seems likely that these athletes should have been aware of being accredited. While it is important to understand the causes of this high error rate, we do not have a conclusive explanation for it as yet. In a future study we will explore potential explanations for it.

The current paper has some limitations concerning the assumptions underlying the present models. First, it is assumed that respondents who answer the control statement randomly do so consistently for the sensitive question. While this assumption is necessary for identification in our setting with only a single sensitive question per respondent in the four surveys, it is not empirically testable within the current model. As pointed out by both the referees and the associate editor, respondents might treat the non-sensitive control question with less attention, as the purpose of the ECWM format may be less apparent, which could result in more random responding on the control question. Conversely, respondents might take the sensitive question more seriously. If true, this asymmetry would lead to an overestimation of γ and, consequently to a misleading overcorrection of the prevalence estimate of the sensitive question. To investigate the potential effect of such an asymmetry, we carried out a small simulation study (see appendix D, please see the supplementary data). We generated 10,000 pairs of randomized responses to the sensitive doping question and the control question from a sample of size $n = 1,000$, with different combinations of random responding probabilities $\gamma_c \in \{.2, .3\}$ for the control item and $\gamma_s \in \{.05, .1, 15, .2\}$ for the sensitive item, with the constraint that $\gamma_c \geq \gamma_s$. The results show that when $\gamma_c > \gamma_s$, using the control question to estimate γ leads to downward-biased corrected prevalence estimates. For example, with $\gamma_c = .20$ and a true prevalence $\pi = .10$, the corrected estimates deviated from the true value of π on average by .026 to .075 percentage points for $\gamma_s \in \{.05, .1, 15\}$. Summary statistics of the simulation results can be found in table 1 of the appendix file D, please see the supplementary data on OSF (https://osf.io/ekcjb/?view_only=e5ad20e51f2c4b4ea3816d5281350782). An alternative to assuming equal random responding across control and sensitive question would be to treat the control-based estimate of γ as prior information about the sensitive-question parameter. This approach relaxes the assumption of equality while still making use of the control question. Such an extension would be best implemented using a Bayesian framework, and further research is recommended to investigate this.

Second, our approach to estimating the prevalence of random responding (γ) ideally requires a non-sensitive control question with a known prevalence of 0 or 1. In such cases, γ can be estimated directly as twice the error rate on the control question, and no assumption about the independence between ignorance of the true answer of the control question and the sensitive question is needed. For the current data, however, this condition was not met, which necessitated the more complex procedure described in section 5. This assumption cannot be empirically verified within the present model and may be implausible in some contexts, for example if doping users are more likely to have obtained legal advice and thus be informed about their accreditation status. Therefore, the corrected prevalence estimates reported in this paper should be interpreted as conditional on this assumption.

Third, it is assumed that the prevalence of random responding (γ) and the prevalence of one-saying (θ) are constant across respondents and across sensitive questions. In principle, these parameters could vary between respondents or across items, and random-effects models could be used to capture such variation (see, e.g., Böckenholt and van der Heijden (2007); Wlömert et al. (2019); Fox (2005)). However, such extensions require multi-item survey designs with several sensitive and control questions, while our application (elite-athlete surveys) involves only a single sensitive item. For this reason, we restricted our focus to the single-item case, which is both practically relevant and identifiable. Extending the model to multi-item designs with random effects is an important avenue for future methodological research.

Fourth, our correction method assumes that participants in the survey had an equal chance of being selected. In other words, it is based on a simple random sampling approach. This assumption is reasonable given the sampling procedures followed in collecting the data from the elite athletes in the four surveys. However, considering complex sample designs (Arnab 2025; Chaudhuri 2001) like stratification, clustering, and unequal probabilities of selection can be the appropriate selection method in many practical surveys especially for multistage recruitment. For our method, it might be interesting to investigate how the demographic characteristics of elite athletes like gender, competition level, and survey completion time affect the likelihood to answer inattentively or evasively by one-saying.

Fifth, the question order may influence the likelihood of random responding, but in the four surveys, athletes were first presented with the sensitive questions, followed by the control question. In our future research, we will investigate the impact of randomized order on the prevalence of random responding, as question order may influence respondents' comprehension or motivation.

A further limitation of the present paper concerns the behavioral assumptions underlying the distinction between random answering and one-saying. While these two behaviors are mutually exclusive, the psychological reality is more complex. In practice, respondents' behavior may reflect a more complex interaction of distrust in the anonymization process, misunderstanding of the randomization procedure, inattentiveness, or deliberate self-protection. Individual differences in cognitive processing, perceived anonymity, and motivation further complicate this picture, making it impossible to determine whether "random answering" truly corresponds to random choice (i.e., selecting each response option with probability 0.5). This limitation underscores the need for future research that allows these behaviors to vary across individuals and questions, for example by analyzing one-saying and random answering in a design with multiple sensitive and control questions.

The main aim of randomized response techniques like the (E)CWM is to increase respondents' willingness to answer sensitive questions truthfully by protecting their privacy and ensuring that their individual answers cannot be linked back to them. However, the effectiveness of these methods depends not only on their statistical properties but also on respondents' psychological trust and understanding of the anonymization mechanism (Landsheer et al. 1999). In a qualitative study of the forced response design by Boeije and Lensvelt-Mulders (2002), some respondents admitted to have modified their responses because they feared being falsely incriminated when forced by the randomizer to give the incriminating "yes" response. In this article, respondents' answers to the number sequence randomizer are not known to investigators, ensuring that individual responses can not be traced. However, respondents may suspect that their answers can still be traced to them, which can undermine the perceived anonymity of the method. Indeed, Cruyff et al. (2024) found empirical evidence for one-saying bias in the four surveys analyzed here and proposed a correction for it, while other studies have demonstrated the presence of random responding (Enzmann 2017; Schnapp 2019; Walzenbach and Hinz 2019). Other forms of evasive behavioral responses, however, may also occur. In this sense, the ECWM including the correction proposed in this paper cannot by themselves overcome psychological distrust or misunderstanding.

A final remark concerns the reasons for random responding. In survey methodology, such behavior is well documented and often attributed to factors such as cognitive load, time pressure, survey fatigue, and general inattentiveness or low motivation. In the context of ECWM, there is little cognitive research directly addressing why respondents might answer randomly. The additional complexity of the ECWM instructions may increase cognitive burden and misunderstanding, which could increase the likelihood of inattentive or haphazard responding. Our findings should therefore be interpreted with caution; the statistical model identifies random answering patterns, but it does not reveal the psychological processes that generate them. Future cognitive psychology research would therefore be valuable to better understand the mechanisms behind random answering in randomized-response formats, for example by the use of qualitative thinking-aloud research, as done by Boeije and Lensvelt-Mulders (2002) in the cognitive survey labs.

Overall, the findings of this article suggest that correcting for random responding result in substantially lower prevalence estimates for a socially undesirable attribute. However, the validity of these corrected estimates cannot be assessed within the present study. Evaluating the validity of the corrections requires carrying out validation studies in which the true prevalence of the sensitive attribute is known and that allow for a comparison of the prevalence estimates obtained with and without correction. Additionally, the corrections we presented in this article are suboptimal due to measurement errors in the answers to the control question because of the athletes who were unaware of the correct answer to the control question about accreditation or licensing. The suggested improvements of the control question would avoid such measurement errors. In future surveys we use such control questions in order to obtain better estimates of the prevalence of random answering and, consequently, obtain more valid estimates of the prevalence of the sensitive attribute.

SUPPLEMENTARY MATERIAL

Supplementary materials are available online at academic.oup.com/jssam. All supplementary files are available online on the OSF repository via: https://osf.io/ekcjb/?view_only=e5ad20e51f2c4b4ea3816d5281350782.

DATA AVAILABILITY

The data sets analyzed during the current study are available on the GitHub page https://github.com/MaartenCruyff/RRsp1, R code to reproduce the results and an appendix of all derivations are available on the OSF repository via: https://osf.io/ekcjb/?view_only=e5ad20e51f2c4b4ea3816d5281350782.

REFERENCES

Arnab, R. (2025), "Randomized Response Techniques for Complex Survey Designs," in *Indirect Methods of Data Collection and Analysis from Surveys*, eds. A. Bandyopadhyay, P. Bandyopadhyay, B. Mukherjee, and B. Sury, Singapore: Springer Nature Singapore, pp. 109–156. https://doi.org/10.1007/978-981-97-6005-3_5.

Atsusaka, Y., and Stevenson, R. T. (2023), "A Bias-Corrected Estimator for the Crosswise Model with Inattentive Respondents," *Political Analysis*, 31, 134–148. https://doi.org/10.1017/pan.2021.43.

Böckenholt, U., Barlas, S., and van der Heijden, P. G. M. (2009), "Do Randomized-Response Designs Eliminate Response Biases? An Empirical Study of Non-Compliance Behavior," *Journal of Applied Econometrics*, 24, 377–392. https://doi.org/10.1002/jae.1052.

Böckenholt, U., and van der Heijden, P. G. M. (2007), "Item Randomized-Response Models for Measuring Noncompliance: Risk-Return Perceptions, Social Influences, and Self-Protective Responses," *Psychometrika*, 72, 245–262. https://doi.org/10.1007/s11336-005-1495-y.

Boeije, H., and Lensvelt-Mulders, G. (2002), "Honest by Chance: A Qualitative Interview Study to Clarify Respondents'(Non-) Compliance with Computer-Assisted Randomized Response," *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 75, 24–39. https://doi.org/10.1177/075910630207500104.

Chaudhuri, A. (2001), "Using Randomized Response from a Complex Survey to Estimate a Sensitive Proportion in a Dichotomous Finite Population," *Journal of Statistical Planning and Inference*, 94, 37–42. https://doi.org/10.1016/S0378-3758(00)00210-X.

Coutts, E., Jann, B., Krumpal, I., and Näher, A.-F. (2011), "Plagiarism in Student Papers: Prevalence Estimates Using Special Techniques for Sensitive Questions," *Jahrbücher für Nationalökonomie und Statistik*, 231, 749–760. https://doi.org/10.1515/jbnst-2011-5-612.

Cruyff, M. J. L. F., Sayed, K. H. A., Petróczi, A., and van der Heijden, P. G. M. (2024), "The One-Sayers Model for the Extended Crosswise Design," *Royal Statistical Society Series A: Statistics in Society*, 1–18. https://doi.org/10.1093/jrsssa/qnae009.

Enzmann, D. (2017), "Die anwendbarkeit des crosswise-modells zur prüfung kultureller unter schiede sozial erwünschten antwortverhaltens," in *Methodische probleme von mixed-mode-ansätzen in der umfrageforschung*, eds. S. Eifler, F. Faulbaum, Wiesbaden, Hessen, Germany: Springer Fachmedien Wiesbaden, pp. 239–277. https://doi.org/10.1007/978-3-658-15834-7_10.

Fox, J.-P. (2005), "Randomized Item Response Theory Models," *Journal of Educational and Behavioral Statistics*, 30, 189–212. https://doi.org/10.3102/10769986030002189.

Greenberg, B. G., Abul-Ela, A.-L. A., Simmons, W. R., and Horvitz, D. G. (1969), "The Unrelated Question Randomized Response Model: Theoretical Framework," *Journal of the American Statistical Association*, 64, 520–539. https://doi.org/10.1080/01621459.1969.10500991.

Heck, D. W., Hoffmann, A., and Moshagen, M. (2018), "Detecting Nonadherence Without Loss in Efficiency: A Simple Extension of the Crosswise Model," *Behavior Research Methods*, 50, 1895–1905. https://doi.org/10.3758/s13428-017-0957-8.

Hoffmann, A., Meisters, J., and Musch, J. (2020), "On the Validity of Non-Randomized Response Techniques: An Experimental Comparison of the Crosswise Model and the Triangular Model," *Behavior Research Methods*, 52, 1768–1782. https://doi.org/10.3758/s13428-020-01349-9.

Hoffmann, A., and Musch, J. (2016), "Assessing the Validity of Two Indirect Questioning Techniques: A Stochastic Lie Detector versus the Crosswise Model," *Behavior Research Methods*, 48, 1032–1046. https://doi.org/10.3758/s13428-015-0628-6.

Höglinger, M., and Diekmann, A. (2017), "Uncovering a Blind Spot in Sensitive Question Research: False Positives Undermine the Crosswise-Model Rrt," *Political Analysis*, 25, 131–137. https://doi.org/10.1017/pan.2016.5.

Höglinger, M., and 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, e0201770. https://doi.org/10.1371/journal.pone.0201770.

Höglinger, M., Jann, B., and Diekmann, A. (2016), "Sensitive Questions in Online Surveys: An Experimental Evaluation of Different Implementations of the Randomized Response Technique and the Crosswise Model," *Survey Research Methods*, 10, 171–187. https://doi.org/10.18148/srm/2016.v10i3.6703.

Hopp, C., and Speil, A. (2019), "Estimating the Extent of Deceitful Behaviour Using Crosswise Elicitation Models," *Applied Economics Letters*, 26, 396–400. https://doi.org/10.1080/13504851.2018.1486007.

Hsieh, S.-H., Perri, P. F., and Hoffmann, A. (2024), "Prevalence Estimates for Covid-19-Related Health Behaviors Based on the Cheating Detection Triangular Model," *BMC Public Health*, 24, 2523. https://doi.org/10.1186/s12889-024-19819-6.

Jann, B., Jerke, J., and Krumpal, I. (2012), "Asking Sensitive Questions Using the Crosswise Model: An Experimental Survey Measuring Plagiarism," *Public Opinion Quarterly*, 76, 32–49. https://doi.org/10.1093/poq/nfr036.

Korndörfer, M., Krumpal, I., and Schmukle, S. C. (2014), "Measuring and Explaining Tax Evasion: Improving Self-Reports Using the Crosswise Model," *Journal of Economic Psychology*, 45, 18–32. https://doi.org/10.1016/j.joep.2014.08.001.

Landsheer, J. A., van der Heijden, P., and van Gils, G. (1999), "Trust and Understanding, Two Psychological Aspects of Randomized Response," *Quality and Quantity*, 33, 1–12. https://doi.org/10.1023/A:1004361819974.

Lensvelt-Mulders, G. J., Hox, J. J., van der Heijden, P. G. M., and Maas, C. J. (2005), "Meta-Analysis of Randomized Response Research: Thirty-Five Years of Validation," *Sociological Methods & Research*, 33, 319–348. https://doi.org/10.1177/0049124104268664.

Meisters, J., Hoffmann, A., and Musch, J. (2020a), "Can Detailed Instructions and Comprehension Checks Increase the Validity of Crosswise Model Estimates?," *PloS One*, 15, e0235403. https://doi.org/10.1371/journal.pone.0235403.

———. (2020b), "Controlling Social Desirability Bias: An Experimental Investigation of the Extended Crosswise Model," *PloS One*, 15, e0243384. https://doi.org/10.1371/journal.pone.0243384.

Meisters, J., Hoffmann, A., Musch, J., Meisters, J., Hoffmann, A., and Musch, J. (2022), "A New Approach to Detecting Cheating in Sensitive Surveys: The Cheating Detection Triangular Model," *Sociological Methods & Research*, 1–41. https://doi.org/10.1177/00491241211055764.

Meisters, J., Hoffmann, A., and Musch, J. (2023), "More than Random Responding: Empirical Evidence for the Validity of the (Extended) Crosswise Model," *Behavior Research Methods*, 55, 716–729. https://doi.org/10.3758/s13428-022-01819-2.

Mieth, L., Mayer, M. M., Hoffmann, A., Buchner, A., and Bell, R. (2021), "Do They Really Wash Their Hands? prevalence Estimates for Personal Hygiene Behaviour during the Covid-19 Pandemic Based on Indirect Questions," *BMC Public Health*, 21, 12–18. https://doi.org/10.1186/s12889-020-10109-5.

Sagoe, D., Cruyff, M., Spendiff, O., Chegeni, R., De Hon, O., Saugy, M., van der Heijden, P. G. M., and Petróczi, A. (2021), "Functionality of the Crosswise Model for Assessing Sensitive or Transgressive Behavior: A Systematic Review and Meta-Analysis," *Frontiers in Psychology*, 12, 655592–655519. https://doi.org/10.3389/fpsyg.2021.655592.

Sayed, K. H. A., Cruyff, M. J. L. F., van der Heijden, P. G. M., and Petróczi, A. (2022), "Refinement of the Extended Crosswise Model with a Number Sequence Randomizer: Evidence from Three Different Studies in the UK," *Plos One*, 17, e0279741. https://doi.org/10.1371/journal.pone.0279741.

Schnapp, P. (2019), "Sensitive Question Techniques and Careless Responding: Adjusting the Crosswise Model for Random Answers," *Methods, Data, Analyses: A Journal for*

Quantitative Methods and Survey Methodology (mda), 13, 307–320. https://doi.org/10.12758/mda.2019.03.

Schnell, R., Thomas, K., Schnell, R., and Thomas, K. (2021), "A Meta-Analysis of Studies on the Performance of the Crosswise Model," *Sociological Methods & Research*, 1–26. https://doi.org/10.1177/0049124121995520.

Umesh, U. N., and Peterson, R. A. (1991), "A Critical Evaluation of the Randomized Response Method: Applications, Validation, and Research Agenda," *Sociological Methods & Research*, 20, 104–138. https://doi.org/10.1177/0049124191020001004.

Walzenbach, S., and Hinz, T. (2019), Pouring water into wine: Revisiting the advantages of the crosswise model for asking sensitive questions,. *Survey Methods: Insights from the Field*. https://doi.org/10.1007/s11199-018-0969-6.

Warner, S. L. (1965), "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias," *Journal of the American Statistical Association*, 60, 63–66.

Wlömert, N., Pellenwessel, D., Fox, J.-P., and Clement, M. (2019), "Multidimensional Assessment of Social Desirability Bias: An Application of Multiscale Item Randomized Response Theory to Measure Academic Misconduct," *Journal of Survey Statistics and Methodology*, 7, 365–397. https://doi.org/10.1093/jssam/smy013.

Yu, J.-W., Tian, G.-L., and Tang, M.-L. (2008), "Two New Models for Survey Sampling with Sensitive Characteristic: Design and Analysis," *Metrika*, 67, 251–263. https://doi.org/10.1007/s00184-007-0131-x.

Journal of Survey Statistics and Methodology, 2026, 00, 1–28
https://doi.org/10.1093/jssam/smag004
Article