10

A New Randomizing Device for the RRT Using Benford’s Law

An Application in an Online Survey

Andreas Diekmann and Marc Hoeglinger

10.1 INTRODUCTION

A crucial feature of any implementation of the Randomized Response Technique (RRT) is the randomizing device. It determines whether or not a particular respondent is required to answer a sensitive question and, consequently, protects respondents’ privacy. Respondents’ cooperation and compliance with the RRT procedure hinge heavily on the device’s ease of use, its trustworthiness, and its availability. Dice, a box with colored balls, a spinner, or cards have frequently been used in face-to-face interviews. But these are difficult to use in paper-and-pencil, online, or telephone surveys because no interviewer is present to provide them to respondents. More commonly, available objects that can be used as randomizing devices, such as coins, are preferable but might still be out of some respondents’ immediate reach. This may lead to RRT break-offs or noncompliance and, as a consequence, invalid measurement. A solution to this problem of availability is to avoid physical randomizing devices and to use questions instead. However, such “randomizing questions” have so far rarely been used, as the range of suitable questions is very restricted.

In this chapter, we present a new randomizing device originally proposed in Diekmann (2012). It uses a “randomizing question” and comprises several desirable properties. Besides its ease of use and its applicability in all survey situations, it allows for increasing the statistical efficiency of the RRT without jeopardizing respondents’ perceived privacy protection. For the latter, the method makes use of Benford’s law and takes advantage of respondents’ misperception of the properties of Benford-distributed numbers such as,
in our example, house numbers. We show how this method can be implemented and we present results of its first large-scale empirical evaluation in an online survey on student cheating. Furthermore, we will discuss the important difference between respondents’ objective privacy protection in RRT designs and their subjectively perceived privacy protection.

### 10.2 THE RANDOMIZED RESPONSE TECHNIQUE (RRT)

The Randomized Response Technique (RRT) is a well-known method to elicit more valid answers to sensitive questions in surveys (originally Warner, 1965; for an overview see Fox & Tracy, 1986 and Krumpal, Jann, Auspurg, & Hermanni, Chapter 11). It provides complete concealment of respondents’ answers by introducing a systematic random error which inhibits any inference of admittance or non-admittance of sensitive behavior from an individual response. This is achieved by a randomizing device such as two dice. In the case of the unrelated-question RRT variant (Horvitz, Shah, & Simmons, 1967; Greenberg, Abul-Ela, Simmons, & Horvitz, 1969), which serves in the following as exemplary case, the randomizing device determines whether a particular respondent has to answer either a sensitive or a non-sensitive question. Respondents could, for instance, be instructed to throw two dice and answer the sensitive question “Have you ever cheated on your taxes?” if their sum is 2 to 8 and to answer the non-sensitive question “Is your mother’s birthday in the months of January through June?” if their sum is 9 to 12. Since only the respondent knows the outcome of the dice throw, no one else is able to infer whether the response given is actually related to the sensitive behavior or not. Accordingly, respondents do not have to fear negative consequences of any kind by admitting a sensitive behavior and should feel free to answer truthfully.

**Estimating the Prevalence of Sensitive Behavior with the RRT**

Even though individual responses are completely concealed, the prevalence of the sensitive behavior can be consistently estimated in the aggregate. The researcher simply takes into account that the observed “yes” responses are not only generated by respondents answering “yes” to the sensitive question but also by respondents answering “yes” to the non-sensitive question. Let $p$ be the probability that respondents are instructed to answer the sensitive
question and $1 - p$ the probability for answering the non-sensitive question whose answer distribution $P(\text{yes}|\text{nonsens.quest.})$ is known. The share of observed “yes” answers is defined as

$$P(\text{yes observed}) = p \cdot P(\text{yes}|\text{sens.quest.}) + (1 - p) \cdot P(\text{yes}|\text{nonsens.quest.}) \quad (10.1)$$

By rearranging the equation, we get the share of respondents answering “yes” to the sensitive question, and, hence, the prevalence of the sensitive behavior under the condition that respondents complied to the RRT instructions:

$$P(\text{yes}|\text{sens.quest.}) = \frac{P(\text{yes observed}) - (1 - p) \cdot P(\text{yes}|\text{nonsens.quest.})}{p} \quad (10.2)$$

The variance of the RRT estimator is then given by (e.g., Fox & Tracy, 1986, p. 19):

$$\text{var}(P(\text{yes}|\text{sens.quest.})) = \frac{P(\text{yes observed}) \cdot (1 - P(\text{yes observed}))}{n \cdot p^2} \quad (10.3)$$

The variance is inversely related to $p^2$, hence the lower the probability that respondents have to answer the sensitive question, the higher the variance of the estimator of the sensitive behavior. Respondents’ privacy protection comes at the cost of a lower statistical efficiency of the RRT estimator.

The RRT Randomizing Device and Its Requirements

The RRT randomizing device serves to introduce randomness into the answering process of survey respondents and, therefore, is the central part of any RRT implementation. The principal requirements a randomizing device has to meet are ease of use, trustworthiness, and availability. Ease of use means that respondents are able to carry out the randomization quickly and without too much effort. Throwing two dice, for instance, does not have to be explained and takes only seconds if dice are readily available.

Trustworthiness regarding the randomizing device means that respondents understand that the outcome of the randomization procedure is truly random and that they believe that the outcome is not detectable by somebody else. The first aspect of trustworthiness, understanding, is assured for well-known randomizing procedures such as throwing dice, flipping a coin, or drawing a card from a deck. Nevertheless, true randomness may be put into question if uncommon or novel random devices are used, such as picking numbers on a screen or using digits of a phone number.
Randomness may also be put into question when the outcome distribution is susceptible to manipulation. This is the case with most “virtual” randomizing devices implemented in online surveys, such as digital coins, dice, or spinners (see Peeters, Lensvelt-Mulders, & Lasthuizen, 2010; Coutts & Jann, 2011 for implementations).

The second aspect of trustworthiness, confidence in the undetectability of the outcome, is often an issue when the RRT is used in interviewer-administered surveys. Respondents might suspect that the interviewer is somehow able to observe the outcome of the randomization procedure. Twenty percent of respondents instructed to draw colored chips from a box in an RRT survey indicated they believed the interviewer knew which chip they would draw—making RRT pointless for these respondents (Wiseman, Moriarty, & Schafer, 1975). A similar issue arises with virtual randomizing devices in online surveys whose outcome might be suspected of being traceable. Undetectability, furthermore, might be questioned when respondents’ answers to “randomizing questions” are used in place of a physical randomizing device. A randomizing question, that is, a question which serves as a randomizing device, may be asked if the distribution of a particular attribute in the surveyed population is known. For instance, the number of persons whose birthday falls in a particular month of the year (“If your birthday is between January and March, please answer the following question: . . . If your birthday is between April and December, please answer the following question: . . .”). However, responses to randomizing questions of this type are still detectable in principle if they refer to respondents themselves or to their relatives and, thus, raise suspicion.

Availability, finally, means that the randomizing device should be within respondents’ reach during the survey. Availability is guaranteed if an interviewer is present to hand over the randomizing device or if the randomizing device is sent out together with a paper-and-pencil questionnaire. In online and telephone surveys, however, the use of a physical randomizing device is almost always problematic. Dice or cards, for instance, are rarely within respondents’ reach. Sending these devices to respondents in advance works in some situations (see de Jong, Pieters, & Fox, 2010 for an example), yet it is costly and still does not guarantee that respondents have the device actually at hand when they answer the survey. The same holds for more common devices such as coins or banknotes. Even though they are available to all respondents in principle, having to get up from the computer to get one’s wallet leads some respondents to skip the randomization procedure. The only safe strategy for self-administered and telephone surveys regarding availability is—in our view—to avoid any physical randomizing device and to use what we call a “randomizing question”.
Questions on birthdays or other known demographics have been used frequently as non-sensitive questions in the unrelated-question RRT design (Horvitz et al., 1967). But they have rarely been used as a randomizing device for the first step in the RRT procedure to determine whether the sensitive or the non-sensitive question has to be answered. In one of the few early RRT studies that applied such a randomizing question, Brown (1975, as cited in Fox & Tracy, 1986, p. 61f.) used a demographic question on respondents’ mothers’ dates of birth in order to determine whether a sensitive or a non-sensitive question had to be answered subsequently. Besides the apparent advantage of availability in all survey situations the use of a randomizing question also entails some caveats. Detectability has already been mentioned. In addition, it is usually difficult to find one or more suitable randomizing questions as the set of possible questions with known response distribution in the surveyed population is usually very restricted.

**Respondents’ Objective and Subjectively Perceived Privacy Protection**

The core rationale underlying the RRT is that respondents understand that their answers remain totally concealed and that thus admitting sensitive behavior bears no risk at all. Respondents’ privacy protection is supposed to lead to more truthful answers and hence to an increase in data validity. Because the deterministic link between individual survey response and admittance of a sensitive behavior is broken by introducing randomness into the answering process, respondents’ protection is guaranteed in all RRT designs. Nonetheless, a probabilistic link between individual response and sensitive behavior remains. The strength of the probabilistic link depends on the particular RRT design and on the true prevalence of the sensitive behavior under question. The researcher directly influences it by defining the RRT design parameter $p$, the probability with which respondents have to answer the sensitive question. A higher $p$ increases the correlation between the individual response and the admittance of sensitive behavior. As a consequence, “respondents’ jeopardy” (Fox & Tracy, 1986, p. 32), defined as $P(\text{sens.behavior}|\text{yes answer})$, the probability that a respondent giving a “yes” response actually admitted the sensitive behavior under question, increases.

However, the choice of $p$ not only influences respondents’ jeopardy or—conversely—respondents’ privacy but also the variance of the RRT estimator as shown in the preceding section. From this fact originates the researcher’s dilemma in choosing an appropriate $p$ for an RRT design: on the one hand, $p$ should be low in order to provide a high level of privacy protection to respondents; on the other hand, $p$ should be as high as possible in order to
obtain an efficient estimator (see Lensvelt-Mulders, Hox, & Heijden, 2005 for statistical implications of the choice of RRT design parameters).

Yet, as Moriarty and Wiseman (1976) already pointed out, it is essential to distinguish between the objective $p$ of an RRT design, and $p$ and the privacy protection as perceived by respondents. Only the latter affects respondents’ trust as well as compliance and, as a consequence, the validity of measurements obtained through the RRT. Even though a correlation between the objective value of $p$ and respondents’ perceived privacy protection may be expected, there is virtually no knowledge about this empirical relation. Studies on the effect of different values of $p$ on respondents’ trust in the RRT, on perceived privacy protection, and on data validity are almost nonexistent and the RRT literature gives no empirically grounded advice on which $p$ to choose. A study by Soeken and Macready (1982) is the only exception known to us. They found a slight decrease in respondents’ perceived privacy protection with increasing $p$ and a statistically significantly lower perceived protection for $p = .91$ compared to values of $p \leq .84$.

10.3 BENFORD RRT: A NEW RANDOMIZING DEVICE USING BENFORD’S LAW

In this section we present Benford RRT, a new randomizing device (originally suggested in Diekmann, 2012), which fulfills the stated requirements of a good RRT randomizing device. At the core of Benford RRT lies a randomizing question on the first digit of an acquaintance’s address house number. First digits of house numbers follow, as we will show in the next section, a known distribution, namely the Benford distribution. This fact can be used to obtain a suitable randomizing device that is applicable in all circumstances. Furthermore, we show how Benford RRT increases the efficiency of the RRT estimator by exploiting the divergence between respondents’ objective privacy protection and their subjectively perceived privacy protection. This divergence is particularly high in the case of Benford RRT due to the “Benford illusion”, the substantial misperception of the frequency of Benford distributed numbers.

Benford’s Law of First Digits

First digits of many real-life data follow a particular distribution with low digits (i.e., “1”) occurring more often than larger digits (i.e., “9”). This fact
has been discovered and the distribution formalized by Newcomb (1881) and later Benford (1938). It is nowadays widely known as Benford’s law. Benford’s law states that the probabilities of first digits $d = 1, 2, \ldots, 9$ are

$$P(d) = \log_{10}(1 + 1/d)$$  \hspace{1cm} (10.4)$$

First digits of the population of countries, the size of lakes, numbers in tax declarations or in newspaper articles, and many other data have all been shown to follow this distribution (e.g., Diekmann, 2012).

In principle, all of these data sources could be used as a randomizing device for Benford RRT. The empirical fit to the Newcomb-Benford distribution should in any case be carefully tested, as the preconditions which produce Benford-distributed first digits might not be fulfilled. For instance, the first digits of numbers in the Bible do follow a Benford distribution, with the exception of the digit 7, which is overrepresented (Huengerbuehler, 2007). Benford (1938) already hypothesized that first digits of house numbers follow a Benford distribution and found supporting evidence using the American Men of Science directory. Diekmann (2012) examined the same, using the Swiss telephone directory. Figure 10.1 shows

**FIGURE 10.1**
Comparison of the empirical distribution of first digits of house numbers from the Swiss phone directory (TwixTel34, $N \approx 3$ million) with the Benford distribution. Numbers compiled by Stefan Wehrli.
that the empirical distribution of first digits of house numbers of Swiss addresses almost perfectly fits the Benford distribution, hence, obeys Benford's law. In a subsequent test, respondents of a general population survey were asked to indicate the house number of an acquaintance. The distribution of the first digits of house numbers generated through this process was in line with the theoretical Benford distribution (Diekmann, 2012). This gives empirical support to the assumption that first digits of house numbers of acquaintances generated by survey respondents approximately follow the Benford distribution.

**Implementing Benford RRT**

Benford RRT uses a question on the first digit of the house number of an acquaintance’s address as a randomizing question. It can be implemented as follows (see also Figure 10.2):

*Please think of an acquaintance of yours whose address you know. Now take the first digit of this person’s house number.*

*If this digit is 1 to 5, please answer the following question: Have you ever cheated on your taxes?*

*If this digit is 6 to 9, please answer this question: Is your mother’s birthday in the months of January through June?*

In this example of an unrelated-question RRT design, the first digit of the house number determines whether a respondent subsequently has to answer a sensitive or a non-sensitive question. $p$ is defined as .78 by choosing the range of digits $\{1, 2, 3, 4, 5\}$ leading to the sensitive question and $\{6, 7, 8, 9\}$ to the non-sensitive question; but, of course, other values are possible.

**FIGURE 10.2**

Benford RRT in the unrelated-question RRT design.
As first digits of house numbers follow the Benford distribution, a question on the first digit of a randomly chosen address’s house number becomes a naturally occurring randomizing device with a known outcome distribution without the need for any physical artifact such as dice or coins. This makes it suitable for any survey situation.

The “Benford Illusion”

The use of a Benford question as randomizing device for the RRT bears the additional advantage that respondents usually underestimate the probability of the occurrence of Benford-distributed low digits because they typically assume a rather uniform distribution. Survey respondents, when explicitly asked, substantially underestimated the probability of the occurrence of the first digits 1 to 4 in house numbers by .09 with a subjective estimate of 0.61 (N = 295; Diekmann, 2012). By making use of that misperception—the Benford illusion—the tradeoff between statistical efficiency and respondents’ perceived privacy protection in RRT designs is relaxed. A higher probability $p$ that respondents are instructed to answer the sensitive question may be chosen without provoking respondents’ privacy concerns because respondents’ subjectively perceived $p$ is substantially lower than the objective $p$.

The idea that a good randomizing device for the RRT should bear the property that respondents perceive $p$ as smaller than the objective $p$, was originally brought up by Moriarty and Wiseman (1976). They investigated respondents’ perception of $p$ for different randomizing devices and found that using two dice had the desired property. Respondents heavily underestimated the outcome probability of a throw of two dice being 4 to 10 by .13 with a median perceived probability of .70; a misperception bias that is similar in magnitude to the one of Benford RRT. In this sense, Benford RRT can be seen as a substitute for the throw of two dice in interview situations where no interviewer is present to provide respondents with dice.

10.4 AN APPLICATION IN A SURVEY ON STUDENT CHEATING

Data and Design

We implemented Benford RRT in an online student survey on exam cheating and plagiarism at two major Swiss universities (Hoeglinger, Diekmann, & Jann, 2013). All students enrolled at the two institutions were contacted...
via their official university email address in spring 2011. Out of a total of 19,410 students, 6,494 completed the survey, resulting in a response rate of 33% (RR1; AAPOR, 2011). Two hundred and one respondents who partially completed, i.e., reached the part of the questionnaire with the sensitive questions, are also included in the following analyses. Respondents who had neither sat an exam nor submitted a paper (386), or who had poor German language skills (230), as well as 67 respondents with incomplete data have been excluded, leaving us with a sample of 6,012 observations. The subsequent analyses are, furthermore, restricted to 1,001 respondents who were surveyed in direct questioning mode and 994 surveyed using Benford RRT.

Survey respondents were asked five sensitive questions about their own cheating behavior, using either direct questioning, Benford RRT, or one of four other RRT variants, which will not be discussed here. Assignment to one of these sensitive question techniques was randomized. The wording of the sensitive questions was identical in all conditions. Benford RRT was implemented in an unrelated-question RRT design as presented in the preceding section. Half of the respondents were directed to the sensitive question with probability $p = .70$, the other half with $p = .78$. This allowed the investigation of whether a different $p$ has any effect on respondents’ admittance of sensitive behavior or on their perceived privacy protection. The unrelated non-sensitive questions consisted of five questions on respondents’ mothers’ dates of birth, with answer distributions of $P(\text{yes}|\text{nonsens. quest.}) \cong .5$ (see endnote 2 for the question wording). Their order was randomized to offset any effects of a particular unrelated question.

**Results**

In order to evaluate Benford RRT, in the following section we compare prevalence estimates of respondents’ admittance of sensitive behavior resulting from Benford RRT and from direct questioning (DQ). Assuming that respondents only falsely deny but never falsely admit a sensitive behavior, higher prevalence estimates are interpreted as a result of more respondents answering truthfully. According to this "more-is-better assumption", which is the basis of all comparative RRT studies (e.g., Lensvelt-Mulders, Hox, van der Heijden, & Maas, 2005), higher prevalence estimates of one method indicate its superior validity. Due to the experimental design, i.e., the fact that respondents were randomly assigned to either Benford RRT or direct questioning, differences in prevalence estimates can be interpreted as causal effects of the particular sensitive question technique. RRT point estimates and standard errors are calculated using a generalization
of the formulae from the first section to the case where different values of \( p \) and \( P(\text{yes}|\text{nonsens.quest.}) \) for subgroups of respondents are used. The procedure is implemented in the Stata module rrreg (Jann, 2008).

Figure 10.3 presents comparisons of prevalence estimates for the five surveyed sensitive cheating behaviors between direct questioning (DQ) and Benford RRT. In the left panel, prevalence point estimates with 95% confidence intervals specified by the lines on both sides of the point estimates are depicted. Estimates range from 17.8% of students admitting having copied in an exam to 1.5% of students admitting partial paper plagiarism. Clearly discernible is the pattern of estimates resulting from Benford RRT

<table>
<thead>
<tr>
<th>Cheating Behavior</th>
<th>DQ Estimate</th>
<th>Benford RRT Estimate</th>
<th>Difference in Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copy in exam</td>
<td>17.8%</td>
<td>17.3%</td>
<td>0.5%</td>
</tr>
<tr>
<td>Notes in exam</td>
<td>9.1%</td>
<td>12.9%</td>
<td>3.8%</td>
</tr>
<tr>
<td>Drugs for exam</td>
<td>3.4%</td>
<td>4.5%</td>
<td>1.1%</td>
</tr>
<tr>
<td>Partial plagiarism</td>
<td>2.9%</td>
<td>7.8%</td>
<td>4.9%</td>
</tr>
<tr>
<td>Severe plagiarism</td>
<td>1.5%</td>
<td>2.4%</td>
<td>0.9%</td>
</tr>
</tbody>
</table>

FIGURE 10.3
Comparison of prevalence estimates of cheating between direct questioning (DQ) and Benford RRT. Lines indicate 95% CIs.* N varies between the different items because questions on cheating in exams have been asked only of respondents who sat in at least one exam; questions on plagiarism only of respondents who have handed in a paper.

*Wording of the sensitive questions (translated from German):
Copy in exam: “In your studies, have you ever copied from other students during an exam?”
Notes in exam: “In your studies, have you ever used illicit crib notes in an exam (including notes on mobile phones, calculators, or similar)?”
Drugs for exam: “In your studies, have you ever used prescription drugs to enhance your performance in an exam?”
Partial plagiarism: “In your studies, have you ever handed in a paper containing a passage deliberately taken from someone else’s work without citing the original?”
Severe plagiarism: “In your studies, have you ever had someone else write a large part of a submitted paper for you or have you handed in someone else’s paper as your own?”
being higher than the corresponding DQ estimates except for the first item, “copy in exam”, where the Benford RRT estimate is marginally lower by 0.6 \([-5.0; 3.9]\) percentage points. Note that confidence intervals for Benford RRT estimates are considerably larger than for DQ, which is due to the RRT’s inherently lower statistical efficiency.

Differences between Benford RRT and DQ estimates are portrayed in the right panel of Figure 10.3. If confidence intervals do not include the zero line, prevalence estimates between Benford RRT and DQ differ significantly at the 95% level. Results show a significant difference only for one out of the five sensitive items, namely “partial plagiarism”, where the Benford RRT estimate is 4.9 \([.8; 9.0]\) percentage points higher than the DQ estimate. For the item “notes in exam” the Benford RRT estimate is 3.8 \([-2.2; 7.8]\) percentage points higher than the DQ estimate; but with a \(p\)-value of .06 the difference just misses the conventional significance level.

Further analysis showed that the survey break-off rate for Benford RRT was almost twice as high as for direct questioning but remained with 2.2% of respondents within an acceptable level. Considering that answering the sensitive questions took respondents 175 seconds with Benford RRT and only 53 seconds with DQ, this is no surprise. Respondents’ self-stated trust in the survey’s anonymity and privacy protection measures was lower for Benford RRT (73% do trust) than for DQ (81%).\(^1\) The RRT procedure seems, initially, to intensify privacy concerns among respondents. However, the risk of disclosure, i.e., the risk that any respondents’ cheating behavior will be exposed because of the survey, is considered lower in the case of Benford RRT (79% see no risk) compared to DQ (71%).\(^2\)

Finally, we compared prevalence estimates and respondents’ perceived privacy protection for Benford RRT designs with different levels of privacy protection, i.e., with different values of \(p\), the probability with which respondents are instructed to answer the sensitive question. Using \(p = .70\) and \(p = .78\), results showed no significant differences in prevalence estimates and no discernible pattern of one RRT design performing systematically differently from the other (see detailed results in the online appendix). Furthermore, respondents’ assessment of anonymity and privacy protection, as well as risk of disclosure, did not differ between the two conditions. Choosing \(p = .78\) instead of \(p = .70\) had clearly no effect on prevalence estimates or respondents’ perception of privacy. Yet, the choice of \(p\) affects statistical efficiency. Therefore, \(p = .78\) is the preferred choice for an implementation of the Benford RRT. Possibly, even a higher \(p\) than \(p = .78\) could be chosen without affecting respondents’ privacy and data validity.
10.5 CONCLUSIONS

In this chapter we have introduced Benford RRT, a new randomizing device for the RRT based on Benford’s law, and have presented results of an empirical evaluation of the method. The new randomizing device uses a randomizing question and does not need any physical artifact. Therefore, it is particularly suitable for self-administered surveys and telephone surveys. In addition, it allows for increasing the statistical efficiency of the RRT without jeopardizing respondents’ perceived privacy protection by taking advantage of the Benford illusion, namely, respondents’ misperception of Benford-distributed first digits.

Benford RRT performed well in our online survey on student cheating behavior. In one out of five items it generated a higher, and, under the more-is-better assumption, a more valid, estimate of sensitive behavior than direct questioning. A second item estimate was substantially higher, but with \( p = .06 \) missed conventional significance levels. No Benford RRT estimate was substantially lower than the DQ estimates, and all Benford RRT estimates were positive and meaningful. In contrast to other RRT online implementations (see, for instance, Coutts, Jann, Krumpal, & Naeher, 2011; Coutts & Jann, 2011; Peeters, 2006), the problem of severely negatively biased or even negative estimates did not arise in our implementation of Benford RRT. It should be noted, however, that a new RRT variant, the Crosswise Method (Tian, Yu, Tang, & Geng, 2007), which was also implemented in our study, performed even better than Benford RRT and seems to be another well-performing, promising method to survey sensitive questions (see Hoeglinger, Jann, & Diekmann, 2013).

Results also showed that an increase of the probability \( p \) with which respondents are instructed to answer the sensitive question by .08 to \( p = .78 \) had no effect on estimates nor on respondents’ perceived privacy protection. Thus, it is safe to choose \( p \) as high as \( p = .78 \) when implementing Benford RRT. Future studies should address in more detail how far \( p \) can be increased without endangering data validity. It remains unclear, though, whether a decrease or increase of \( p \) within a reasonable range has no effect on respondents’ perceived privacy protection in other RRT designs or whether this is somehow related to the Benford illusion, a special property of Benford RRT. Results suggest, in any case, that respondents’ perception of privacy protection is mainly driven by other design considerations than the mere choice of \( p \).

Whether the increase in more truthful answers achieved through Benford RRT justifies the additional burden put on respondents and the need for
bigger sample sizes in order to compensate for the RRT’s lower statistical efficiency depends on two things: the sensitivity of the topic surveyed and whether a sizeable respondent sample is actually available. If an implementation of the RRT is considered, however, Benford RRT seems to be a well-performing RRT variant that is easily implemented not only, but particularly, in survey situations where no interviewer is present.

ACKNOWLEDGEMENTS

We thank Ben Jann for his support in the design and conduct of this study as well as for valuable suggestions on improving the manuscript.

FUNDING

This research was supported by a grant from the German Research Foundation, DI 292/5 to Andreas Diekmann.

NOTES

1. Wording of the question: “Please be honest: How much do you trust our measures for guaranteeing survey participants’ anonymity and privacy protection?” Response categories “very much” and “quite a bit” have been coded as respondent does trust; response categories “partly”, “rather not”, and “not at all” as respondent does not trust.

2. Wording of the question: “In your opinion, how likely is it that it can be traced back to a particular respondent whether they carried out one of the surveyed sensitive behaviors (copying in exam, crib notes, plagiarism, etc.)?” Response categories “impossible” and “very unlikely” have been coded as respondent sees no risk; response categories “rather unlikely”, “rather likely”, and “very likely” as respondent sees a risk.

BIBLIOGRAPHY


Andreas Diekmann and Marc Hoeglinger


