

# BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email [info.bmjopen@bmj.com](mailto:info.bmjopen@bmj.com)

# BMJ Open

## Measuring open defecation in India using survey questions: Evidence from a randomized survey experiment

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-030152
Article Type:	Research
Date Submitted by the Author:	01-Mar-2019
Complete List of Authors:	Vyas, Sangita; University of Texas at Austin, Economics Srivastav, Nikhil; University of Texas at Austin Mary, Divya Goel, Neeta Srinivasan, Sujatha Tannirkulam, Ajaykumar Ban, Radu Spears, D; University of Texas at Austin Coffey, Diane; University of Texas at Austin
Keywords:	India, open defecation, sanitation, survey methods, experiment

SCHOLARONE™  
Manuscripts

## Measuring open defecation in India using survey questions: Evidence from a randomized survey experiment

Sangita Vyas, Nikhil Srivastav, Divya Mary, Neeta Goel, Sujatha Srinivasan, Ajaykumar  
Tannirkulam, Radu Ban, Dean Spears, Diane Coffey

**Objectives:** To investigate whether a question about latrine use or open defecation for every member of a household finds different levels of open defecation compared to a household-level question.

**Setting:** Rural India is home to most of the world's open defecation. India's Demographic and Health Survey (DHS) estimates that 54% of households in rural India defecated in the open in 2015-2016. This measure is based on a household-level question that asks about the behavior of everyone in the household in one question. Yet, studies in rural India find substantial open defecation among individuals living in households with latrines, suggesting that household-level questions underestimate true open defecation.

**Participants:** In 2018, we randomly assigned latrine-owning households in the rural parts of four Indian states to receive one of two survey modules measuring sanitation behavior. 1,215 households were asked about latrine use or open defecation individually for every household member. 1,216 households were asked the question used in India's DHS, which asks, at the household level, what type of facility members of the household usually use.

**Results:** Using two methods for comparing open defecation by question, the individual-level question found 20 to 21 (95% CI 16 to 25 for both estimates) percentage points more open defecation than the household-level question, among all households, and 28 to 29 (95% CI 22 to 35 for both estimates) percentage points more open defecation among households that received assistance to construct their latrines.

**Conclusions:** We provide the first evidence that individual-level questions find more open defecation than household-level questions. Because reducing open defecation in India is essential to meeting the Sustainable Development Goals, and exposure to open defecation has consequences for child mortality, health, and development, it is essential to monitor its progress as accurately as possible.

**Registration:** Registry for International Development Impact Evaluations, number 5b55458ca54d1.

**Keywords:** India, open defecation, sanitation, survey methods, experiment

### Affiliations

**Economics, University of Texas at Austin, Austin, TX, USA (S Vyas MPA, D Spears PhD);  
Population Research Center, University of Texas at Austin, Austin, TX, USA (S Vyas MPA,**

1  
2  
3 D Spears PhD, D Coffey PhD); **LBJ School of Public Affairs, University of Texas at Austin,**  
4 **Austin, TX, USA** (N Srivastav BA); **IFMR Lead, Chennai, India** (D Mary MS, S Srinivasan  
5 MBA, A Tannirkulam PhD); **International Initiative for Impact Evaluation (3ie), New Delhi,**  
6 **India** (N Goel PhD); **Bill & Melinda Gates Foundation** (R Ban PhD); **Indian Statistical**  
7 **Institute, Delhi, India** (D Spears PhD, D Coffey PhD).  
8  
9

#### 10 Correspondence to

11 Sangita Vyas, Economics, University of Texas at Austin, Austin, TX 78712, USA;  
12 sangita.vyas@utexas.edu  
13

#### 14 Funding

15 International Initiative for Impact Evaluation, Bill & Melinda Gates Foundation.  
16  
17

#### 18 Role of funders

19 Radu Ban, who is employed on the WSH program of BMGF, contributed to the analysis of the  
20 results. Neeta Goel contributed to coordinating data collection. Both contributed inputs to  
21 drafting the report.  
22  
23

#### 24 Contributors

25 SV NS DS DC contributed to the study design. SV NS DC DS designed the survey instruments.  
26 NS DM SS AT oversaw data collection. SV NS NG coordinated between the 3ie research teams  
27 and the survey team for this study. SV RB DC DS contributed to analysis. All authors  
28 contributed to redrafting the report.  
29  
30

#### 31 Acknowledgements

32 We are very grateful for the collaboration and contributions of the 3ie research teams in  
33 formulating the individual-level questions. The questions are the product of a deliberative  
34 process between the authors and the research teams. We are also grateful for the research teams'  
35 cooperation on this study. In particular, we thank researchers from Oxford Policy Management,  
36 London School of Hygiene and Tropical Medicine, Indian Institute of Public Health  
37 Gandhinagar, Eawag, and Emory University, and the implementing agencies with which they are  
38 working, including World Vision, Coastal Salinity Prevention Cell, Wateraid, and Rural Welfare  
39 Institute, for sharing their census data with us, coordinating field work, and supporting this  
40 endeavor. This study would not have been possible without their collaboration and assistance.  
41  
42  
43

#### 44 Declaration of interests

45 We have no competing interests.  
46

#### 47 Data sharing

48 Deidentified data will be made publicly available upon publication.  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Strengths and limitations

- This is the first study that experimentally tests the difference in open defecation estimated from different survey questions.
- This study provides evidence that the India's Demographic and Health Survey, which asks about the defecation behavior of everyone in the household in one question, substantially underestimates open defecation in India.
- The study shows that measuring open defecation at the individual level is feasible in a large household survey, and finds more open defecation than household-level questions.
- Monitoring open defecation in India is important for understanding progress towards the Sustainable Development Goals.
- Because the households participating in this study are not representative of the rural parts of the states they are in, or rural India, the estimates presented here should not be considered as estimates of open defecation for any of the states, or the country as a whole.

## Introduction

Rural India is home to more than half of the world's open defecation.<sup>1</sup> Because the persistence of open defecation threatens gains in child health, the Sustainable Development Goals (SDGs) call for its elimination by 2030. Progress towards eliminating open defecation in rural India will be essential to meeting this goal. India's most recent Demographic and Health Survey (DHS), conducted between January 2015 and December 2016, estimates that 54% of households in rural India defecated in the open, down from 75% in the 2005-2006 DHS.<sup>2,3</sup> This measure is based on a household-level question that asks about the behavior of everyone in the household in the same question.

Recent evidence from studies carried out in India suggests, however, that it is common for individuals living in households with latrines to nevertheless defecate in the open. In rural parts of five north Indian states, Coffey et al. found that 21% of individuals defecated in the open, despite owning a latrine.<sup>4</sup> In rural Tamil Nadu, Yogananth and Bhatnagar report that 54% of respondents defecated in the open despite having a household latrine.<sup>5</sup> In Odisha, Barnard et al. found that less than half of members of households with latrines reported using their latrines at all times.<sup>6</sup>

Experimental studies of sanitation interventions have found similar results. Clasen et al. report on a sanitation intervention in Odisha and note open defecation among individuals living in households with latrines as a reason for not observing impacts on child health outcomes.<sup>7</sup> Patil et al. conducted a sanitation intervention in Madhya Pradesh and experienced a similar problem: modest increases in latrine coverage, and even more modest reductions in open defecation.<sup>8</sup> These findings suggest that open defecation among latrine-owning households is substantial. Since it is probable that latrine use is the socially desirable response to questions on sanitation behavior, measures based on household-level questions, such as those from the DHS, will likely underestimate true open defecation in rural India, particularly among households with latrines.

Because open defecation is an individual behavior, an individual-level survey question may be able to more accurately measure it compared to a household-level question, particularly among households with latrines. We designed this study to experimentally test this hypothesis in rural India. We aimed to investigate whether a balanced question about latrine use or open defecation for every member of a household finds different levels of open defecation compared to a household-level question.

This is the first study to experimentally vary survey methodology to improve upon the measurement of open defecation currently being used. Jenkins et al. study sanitation survey methods, and report on an index they develop for quantifying household excreta disposal.<sup>9</sup> Their study focuses on developing and piloting a new tool rather than comparing different measures.

1  
2  
3 Sinha et al. compare answers to survey questions on latrine use behavior to measures of actual  
4 behavior generated from passive latrine use monitors that were set up in the latrines of  
5 respondents, and find poor to moderate agreement between the two measures.<sup>10</sup> Our study  
6 contributes to this literature by comparing estimates of open defecation obtained from questions  
7 that can be administered in a large household survey, and highlighting potential sources of error  
8 in open defecation measurement.  
9  
10

11  
12 Four years ago, the Government of India launched the Swachh Bharat Mission (SBM), a national  
13 sanitation campaign, which aims to eliminate open defecation in India by 2019. Many latrines  
14 have been constructed in rural India as a result of this campaign. Yet, the effect the SBM has had  
15 on reducing open defecation is still unknown. Because large reductions in open defecation in  
16 India are essential to meeting the SDGs, and because exposure to open defecation has serious  
17 consequences for child mortality, health, and human capital development, it is essential to  
18 monitor its progress as accurately as possible.  
19  
20  
21  
22

## 23 Methods

24  
25  
26 The study received ethical approval for research involving human subjects from the Institute for  
27 Financial Management and Research's Institutional Review Board in India and is registered in  
28 the Registry for International Development Impact Evaluations (RIDIE), number  
29 5b55458ca54d1.  
30  
31  
32

### 33 Sample: mostly latrine-owning households in rural parts of four states

34  
35  
36 This study uses as its sampling frame the study areas of 3ie's Promoting Latrine Use in Rural  
37 India Thematic Window. This Window has funded four independent research teams to conduct  
38 randomized control trials of distinct behavioral campaigns to promote the use of pit latrines in  
39 rural parts of Bihar, Gujarat, Karnata, and Odisha. The study areas are spread across India,  
40 representing different contexts and varying levels of rural open defecation.  
41  
42  
43

44  
45 Because these trials focus on behavioral strategies rather than latrine construction, they are being  
46 carried out in villages that had high levels of coverage of pit latrines at baseline, relative to other  
47 rural parts of the same states. The households that comprise the sampling frame for this study are  
48 those that were identified as having a functional latrine in a census conducted by the research  
49 teams in the villages in which they were working. In all states except for Odisha, only  
50 households that had been excluded from the research teams' samples could be selected for this  
51 study. In Odisha, the sample selected for this study overlaps with the research teams' sample. We  
52 aimed to survey households that own latrines because we expect that an important source of  
53 misreporting of open defecation comes from individuals who do not use the functional latrines  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 that their households own, and that other members of their households use.  
4  
5

6 Figure 1 describes the sample selection. The villages visited in each state were randomly selected  
7 from the full set of villages included in the 3ie research teams' studies. The full set of villages  
8 were selected by the research teams in collaboration with the implementation agencies they were  
9 working with. The research team led by Oxford Policy Management worked with World Vision  
10 in Bihar, the team led by London School of Hygiene and Tropical Medicine worked with Coastal  
11 Salinity Prevention Cell (CSPC) in Gujarat, Eawag worked with Wateraid in Karnataka, and  
12 Emory University worked with the Rural Welfare Institute (RWI) in Odisha. Data for our study  
13 were collected in 22 to 25 villages in each of the four study areas. In most areas, we sampled  
14 more villages than we actually visited in order to facilitate coordination with the research teams.  
15 95 villages were visited in total.  
16  
17  
18  
19

20 Up to 40 households in each village were randomly assigned to receive the household or  
21 individual questions. In some villages, fewer than 40 households were assigned because fewer  
22 than 40 households met the eligibility criteria. The survey team visited as many assigned  
23 households as it could in these villages, given time constraints, and availability of household  
24 members. On average, the survey team interviewed 25 households per village. Data collection  
25 took place between March and July 2018.  
26  
27  
28

29 In each state, data collection took place after the 3ie research teams had conducted their censuses  
30 and baselines, but before they had started their interventions. Since in all states, the households  
31 visited in this study were also visited at the time of the census, response bias may be a concern.  
32 This would not, however, impact the internal validity of this study since randomization generates  
33 equal response bias, in expectation, across treatment arms.  
34  
35  
36  
37

### 38 Randomization and masking: random variation in latrine use questions at the 39 household level 40 41 42

43 We randomly assigned the type of latrine use question administered in the survey at the  
44 household level. Roughly half of the households were assigned individual-level questions on  
45 latrine use. The other half were assigned a household-level question. One of the authors who was  
46 not involved in data collection carried out the randomization using a random number generator in  
47 Stata. Because of the nature of the study, it was not possible to blind the respondents or  
48 surveyors to the type of survey question administered in the survey. However, in the interest of  
49 data quality, respondents were not explicitly told that the primary purpose of the survey, which  
50 took approximately 25 minutes to complete, was to measure open defecation. Additionally,  
51 surveyors did not know which survey question had been assigned to a household until starting  
52 the survey with the household.  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

The individual-level questions asked for every household member age five or older whether the individual defecated in the open or used the latrine. The preface to this series of questions was: “I have seen that some people defecate in the open, and some people use the latrine. Now I want to ask about where you and your family members defecate.” Then, the surveyor asked the following question for each individual in the household, and coded the answer in a household roster: “The last time [*name of household member*] defecated, did [*name of household member*] defecate in the open or use the latrine?” The answer options included latrine, open, and somewhere else. Surveyors used the last option, which meant that the household member defecated in a bedpan, cloth, or other place, in less than 0.5% of cases. Because the priming statement and the behavior question include both open defecation and latrine use, they are balanced between the two different behaviors and could reduce social desirability bias. The surveyor asked household members who were participating in the interview directly about their behavior, and asked the main respondent, in most cases an adult female member of the household, to report on the behavior of their family members who were not participating in the interview.

The rest of the households were assigned the household-level question used in India’s DHS: “What kind of toilet facility do members of your household usually use?”<sup>11</sup> The answer codes were also the same as those used in India’s DHS: flush to piped sewer system, flush to septic tank, flush to pit latrine, flush to somewhere else, flush to don’t know where, ventilated improved pit or biogas latrine, pit latrine with slab, pit latrine without slab or open pit, twin pit or composting toilet, dry toilet, and no facility or uses open space or field. We also included an individual-level question on mobile ownership or preferring vegetarian food versus non-vegetarian food in the surveys that asked the household-level question so that both types of surveys would take approximately the same amount of time to complete.

There are three main factors that differ between the two types of latrine use survey modules: the level of aggregation, the reference period, and the presence of a priming statement. Therefore, the differences in reported open defecation that we observe reflect the fact that the two sets of questions vary on all of these factors combined.

## Statistical analyses

The outcome of interest is reported open defecation. For the household-level questions, we created a dummy variable that is equal to one if the response was “no facility or uses open space or field,” and zero otherwise. The unit of observation for households assigned the individual-level question is the individual, while the unit of observation for those assigned the household-level question is the household. Therefore, in order to directly compare and test the significance of differences in reported open defecation between the two question types, we construct estimates that use the same unit of observation. We impute individual-level open defecation from

1  
2  
3 responses to the household-level questions, and household-level open defecation rates from  
4 responses to the individual-level questions. To construct individual-level open defecation using  
5 the household-level questions, we assign the answer from the household question to each  
6 individual in the household. Similarly, to construct household-level open defecation using the  
7 individual-level questions, we average the responses among individuals in the household, and  
8 assign this average as the household value. Our main analysis tests differences in measured open  
9 defecation by question type. We show pooled results, as well as results by study area.  
10  
11  
12

13  
14 We also conduct sub-group analyses. First, we investigate whether the difference in reported  
15 open defecation by question type depends on whether the latrine was constructed privately, or  
16 with assistance from the government or an NGO. In practice, assistance to construct latrines  
17 often comes from the government, but sometimes NGOs get involved in facilitating the  
18 implementation of the government program. As part of the SBM, the Government of India assists  
19 rural households to construct latrines either by providing financial assistance directly to  
20 households so they can construct their own latrines, or by local government officials constructing  
21 latrines for households. In the discussion that follows, we will describe a household as having  
22 “received help” if it received financial assistance or a partially or completely constructed latrine  
23 from the government or an NGO.  
24  
25  
26  
27

28 The Indian government promotes and constructs latrines with pits that are approximately 60  
29 cubic feet.<sup>12</sup> However, many rural Indians aspire to construct latrines with pits that are much  
30 larger, so that they can avoid emptying the pit, a task that is associated with ritual pollution.<sup>13–  
31 15,19</sup> Compared to latrines constructed privately, those constructed with government help are less  
32 likely to be used due to concerns over purity and pit emptying. Since a large fraction of rural  
33 households are likely to receive latrines with help from the government as a result of the SBM, it  
34 is important to explore how much open defecation different types of latrine use questions  
35 measure, based on having received help to construct the latrine.  
36  
37  
38  
39

40 Second, we investigate whether the difference in measured open defecation between the two  
41 question types is statistically different for males compared to females. Sex differences are an  
42 important aspect to explore because observational studies have found consistently higher open  
43 defecation among latrine owners for males compared to females.<sup>4</sup> This observation could reflect  
44 greater demand for latrine use among females due to, for instance, greater psychosocial stress  
45 experienced when defecating in the open,<sup>16</sup> or it could be because of cultural norms that keep  
46 females in their reproductive years inside the home.  
47  
48  
49  
50

51 We cluster standard errors in all analyses at the village level. Statistical analyses were conducted  
52 using Stata (version 11).  
53  
54  
55  
56  
57  
58  
59  
60

## Public involvement

The individual-level questions used in this study are the product of a deliberative process between the authors and research teams from Oxford Policy Management, London School of Hygiene and Tropical Medicine, Indian Institute of Public Health Gandhinagar, Eawag, and Emory University. The findings of this study contribute to our understanding of the scale of an important public health problem.

## Results

Table 1 shows summary statistics for households assigned the two types of latrine use questions. The total sample consisted of 2,431 households, which were approximately equally divided across question type in each of the study areas. The means presented in Table 1 show that there were no significant differences on measures relevant for latrine use between households assigned different types of latrine use questions. Households in both groups had approximately the same number of household members, fraction female, fraction Hindu, educational attainment of the household head, and asset ownership of 13 assets, including mobile phone, electricity, radio, television, fan, mosquito net, bicycle, motorcycle, car, chair, gas stove, pressure cooker, and shoes for everyone in the family. It is important for the validity of the results that the sample is balanced on religious composition, since studies have documented an association between household religion and latrine use.<sup>17,18</sup>

Table 1 also shows that, as the study design intended, most households in both groups had a latrine. Conditional on having a latrine, 64% of households had gotten help from the government or an NGO to build the latrine, and the average pit size was 180 cubic feet. It is important that the sample is balanced on these two characteristics because, compared to latrines constructed privately, those constructed with government help are less likely to be used because of concerns over purity and pit emptying.<sup>13–15,19</sup> Finally, among households with latrines, approximately 80% of them appeared to the surveyor to be in use upon observation.

Figure 2 presents the main results of the study; it shows means and 95% confidence intervals from the individual-level (shown as dashed red bars) and the household-level (shown as solid blue bars) questions, for the full dataset and for different sub-samples. Observations are individuals for the individual-level estimates, and households for the household-level estimates. The first set of estimates shown in the figure uses the full sample. The second, third, fourth, and fifth sets of estimates break the sample up by project area. In the full sample, and in each project area, the individual questions find more open defecation. The sixth set of estimates uses only households with latrines and finds similar results as the full sample. The size of the difference in reported open defecation between question types is large and consistent. No matter how the data

1  
2  
3 are broken up, the individual-level, balanced latrine use questions find significantly higher rates  
4 of open defecation than the household-level question.  
5  
6

7 Table 2 shows actual and imputed open defecation rates, measured at the individual and  
8 household levels. Columns 1 and 3 are the estimates shown in Figure 2. Columns 2 and 4 show  
9 imputed open defecation rates, at the individual and household level, respectively. Imputed  
10 values are calculated based on the method described in the Statistical Analyses section. Column  
11 5 shows the difference in the measured rate of open defecation between the individual and  
12 household questions when observations are individuals, and column 6 shows the same difference  
13 when observations are households. In the full sample, the individual-level, balanced questions  
14 find 21 (95% CI 16 to 25) percentage points more open defecation than the household-level  
15 question when observations are individuals, and 20 (95% CI 16 to 25) percentage points more  
16 open defecation when observations are households. Notably, the individual-level questions  
17 measure consistently higher levels of open defecation in the full sample and in all subsamples,  
18 irrespective of how the difference is calculated. All differences are significant at the one percent  
19 level.  
20  
21  
22  
23  
24  
25

26 The seventh set of estimates shown in Figure 2 show reported open defecation from individual-  
27 and household-level questions among households that received help to construct their latrines.  
28 Comparing these estimates to the sixth set of estimates in the figure, which include all  
29 households with latrines, suggests that the household-level question underestimates open  
30 defecation by more among those that received help to construct their latrines, compared to those  
31 that did not. The seventh row in Table 2 shows that, using different methods for computing  
32 differences, individual-level questions measure 28 to 29 (95% CI 22 to 35 for both estimates)  
33 percentage points more open defecation than the household-level question, among those that  
34 received help to construct their latrines, compared to those that did not.  
35  
36  
37  
38

39 The first two columns in Table 3 test whether the difference in measured open defecation  
40 between the two question types is statistically different among households that received help to  
41 construct their latrines compared to households that did not receive help to construct their  
42 latrines. The first column in Table 3 uses individuals as observations, and the second column  
43 uses households. The coefficients in the third row represent the difference-in-differences  
44 estimate. The difference in measured open defecation between the two question types is 18 (95%  
45 CI 11 to 25) to 19 (95% CI 11 to 26) percentage points larger among households that received  
46 help to construct their latrines, compared to households that did not, depending on the method  
47 used to compute the difference. Among households that did not receive help to construct their  
48 latrines, the individual-level questions find 10 (95% CI 5 to 15) percentage points more open  
49 defecation than the household-level question.  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 The last column in Table 3 investigates whether the difference in measured open defecation  
4 between the two question types is statistically different among males compared to females. Using  
5 the full sample, individual questions find 4 (95% CI 2 to 7) percentage points more open  
6 defecation than the household question among males compared to females. Breaking the sample  
7 up by sex, the individual questions find 23 (95% CI 18 to 28) percentage points more open  
8 defecation than the household question among males, and 19 (95% CI 14 to 24) percentage  
9 points more among females. The household question underestimates open defecation by more  
10 among males compared to females.  
11  
12  
13  
14

## 15 Discussion

16  
17  
18 Our findings show that in our sample, individual-level, balanced questions find 20 to 21 (95%  
19 confidence intervals 16 to 25 for both estimates) percentage points more open defecation than the  
20 household-level question. This is both a statistically significant and practically important  
21 difference. This study presents compelling evidence that India's DHS, which provides the most  
22 recent nationally representative estimates of open defecation for rural India, and other surveys  
23 that ask household-level questions, greatly underestimate open defecation among households  
24 with latrines.  
25  
26  
27  
28

29 We also found that the difference in reported open defecation between the two question types is  
30 significantly greater for households that received help to construct their latrines compared to  
31 households that did not. Among households that received help to construct their latrines, the  
32 individual-level questions find 28 to 29 (95% confidence intervals 22 to 35 for both estimates)  
33 percentage points more open defecation than the household-level question. This suggests that as  
34 more and more households receive government assistance for a latrine through the Swachh  
35 Bharat Mission, household-level questions will become even less accurate at estimating open  
36 defecation.  
37  
38  
39  
40

41 The larger difference in measured open defecation between the two question types among  
42 households that received help compared to those that did not is likely arising from higher rates of  
43 open defecation among households that received help to construct their latrines. There are  
44 several reasons that could explain why households receiving assistance may be less likely to use  
45 their latrines. First, these households are likely to have lower demand for latrine use, compared  
46 to households that built latrines on their own. Second, households that received help have latrines  
47 with pits that are on average 150 cubic feet smaller than the pits of latrines in households that did  
48 not receive help. Because of concerns over ritual purity, rural Indians are less likely to use  
49 latrines with pits that need to be emptied manually every few years, like the latrines that are  
50 promoted and constructed by the government.<sup>13-15,19</sup> Whether only one, both, or other factors are  
51 leading to more open defecation among households that received help, the individual-level  
52 questions are better able to capture this open defecation than the household-level question.  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4 We also find a statistically significant difference in reported open defecation between the two  
5 question types for males compared to females. The difference between the individual-level,  
6 balanced questions and the household-level question is 4 (95% CI 2 to 7) percentage points more  
7 for males compared to females. This supports evidence that, conditional on latrine ownership,  
8 males are more likely to defecate in the open compared to females. Individual-level questions  
9 understate the difference in open defecation between the two sexes by less than the household-  
10 level questions. Compared to the difference in reported open defecation by receiving help to  
11 construct the latrine, differences by sex are not as large.  
12  
13  
14  
15

16 Measuring open defecation at the individual level is feasible. Our survey team's experience  
17 suggests that adding the balanced, individual-level questions on use to a survey that already  
18 contains a household roster increases survey time by about 2 minutes, on average. Of course, the  
19 amount of time required to ask the individual-level questions depends on the number of  
20 individuals in the household.  
21  
22  
23

24 The household-level question asked in the DHS also collects information on the types of latrines  
25 that households own, data that is still of great interest to researchers and practitioners. Therefore,  
26 individual questions on use, combined with a separate question on the types of latrines that  
27 households own, would satisfy both goals: evaluating latrine infrastructure, and measuring open  
28 defecation as accurately as possible.  
29  
30  
31

32 A limitation of our study is that the samples from these project areas are not representative of the  
33 rural parts of the states they are in, nor are they collectively representative of rural India. The  
34 households in this study are much more likely to have a latrine than the average rural Indian  
35 household, and therefore, the individuals in this study are more likely to use a latrine. For this  
36 reason, the estimates presented here should not be considered as estimates of open defecation for  
37 any of the states, or the country as a whole. Rather, they show a large and significant difference  
38 in reported open defecation based on the type of question asked.  
39  
40  
41  
42

43 Measuring open defecation at the individual-level in a large household survey is doable and will  
44 provide a more accurate estimate of open defecation in rural India. Since reducing open  
45 defecation in India is important for meeting the SDGs, and since open defecation is an important  
46 factor contributing to poor health among children in India, it is important to measure its progress  
47 as accurately as possible.  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## References

- 1 WHO and Unicef. *Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines*. Geneva: World Health Organization (WHO) and the United Nations Children's Fund (UNICEF), 2017.
- 2 International Institute for Population Sciences (IIPS) and ICF. National Family Health Survey (NFHS-4), 2015-16 [Dataset]. IAHR74DT.DTA. India. Mumbai: IIPS [Producers]. ICF [Distributor], 2017.
- 3 International Institute for Population Sciences (IIPS) and Macro International. National Family Health Survey (NFHS-3), 2005–06 [Dataset]. IAHR52DT.DTA. India. Mumbai: IIPS [Producers]. ICF [Distributor], 2007.
- 4 Coffey, D, A Gupta, P Hathi, et al. Revealed preference for open defecation. *Economic & Political Weekly* 2014, 49(38), 43-55.
- 5 Yogananth, N, & T Bhatnagar. Prevalence of open defecation among households with toilets and associated factors in rural south India: an analytical cross-sectional study. *Transactions of The Royal Society of Tropical Medicine and Hygiene* 2018, 112(7), 349-360.
- 6 Barnard, S, P Routray, F Majorin, et al. Impact of Indian Total Sanitation Campaign on latrine coverage and use: a cross-sectional study in Orissa three years following programme implementation. *PloS One* 2013, 8(8), e71438.
- 7 Clasen, T, S Boisson, P Routray, et al. Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial. *The Lancet Global Health* 2014, 2(11), e645-e653.
- 8 Patil, SR, BF Arnold, AL Salvatore, et al. The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: a cluster randomized controlled trial. *PLoS Medicine* 2014, 11(8), e1001709.
- 9 Jenkins, MW, MC Freeman, & P Routray. Measuring the safety of excreta disposal behavior in India with the new Safe San Index: Reliability, validity and utility. *International Journal of Environmental Research and Public Health* 2014, 11(8), 8319-8346.
- 10 Sinha, A, CL Nagel, E Thomas, WP Schmidt, B Torondel, S Boisson, & TF Clasen. Assessing latrine use in rural India: a cross-sectional study comparing reported use and passive latrine use monitors. *The American Journal of Tropical Medicine and Hygiene* 2016, 95(3), 720-727.

- 1  
2  
3 11 International Institute for Population Sciences (IIPS) and ICF. National Family Health  
4 Survey (NFHS-4), 2015-16. Household Questionnaire. India. Mumbai: IIPS, 2017.  
5  
6 12 Ministry of Drinking Water and Sanitation. Guidelines for Swachh Bharat Mission  
7 (Gramin). Government of India, New Delhi, India, 2017.  
8  
9  
10 13 Coffey, D, A Gupta, P Hathi, et al. Understanding open defecation in rural India:  
11 Untouchability, pollution, and latrine pits. *Economic & Political Weekly* 2017, 52(1), 59-  
12 66.  
13  
14 14 O'Reilly, K, R Dhanju, & E Louis. Subjected to sanitation: Caste relations and sanitation  
15 adoption in rural Tamil Nadu. *The Journal of Development Studies* 2017, 53(11), 1915-  
16 1928.  
17  
18 15 Routray, P., WP Schmidt, S Boisson, T Clasen, & MW Jenkins. Socio-cultural and  
19 behavioural factors constraining latrine adoption in rural coastal Odisha: an exploratory  
20 qualitative study. *BMC Public Health* 2015, 15(1), 880.  
21  
22  
23 16 Sahoo, KC, KR Hulland, BA Caruso, et al. Sanitation-related psychosocial stress: a  
24 grounded theory study of women across the life-course in Odisha, India. *Social Science  
25 & Medicine* 2015, 139, 80-89.  
26  
27  
28 17 Vyas, S, & D Spears. Sanitation and Religion in South Asia: What Accounts for  
29 Differences across Countries? *The Journal of Development Studies* 2018, 54(11), 1-17.  
30  
31  
32 18 Geruso, M, & D Spears. Neighborhood sanitation and infant mortality. *American  
33 Economic Journal: Applied Economics* 2018, 10(2), 125-62.  
34  
35  
36 19 Coffey, D, & D Spears. *Where India goes: Abandoned toilets, stunted development and  
37 the costs of caste*. Harper Collins, India, 2017.  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



**Table 1: Randomization balance: No significant differences in observed means between households assigned different latrine use questions**

	individual (1)	household (2)	difference (3)
Number of households	1,215	1,216	-1
Number of households by project			
World Vision (in Bihar)	316	313	3
Coastal Salinity Prevention Cell (in Gujarat)	309	319	-10
Wateraid (in Karnataka)	297	296	1
Rural Welfare Institute (in Odisha)	293	288	5
Household members	5·685 (0·0950)	5·604 (0·101)	0·0803 (-0·117 to 0·278)
Female	0·491 (0·00485)	0·489 (0·00490)	0·00190 (-0·0115 to 0·0153)
Hindu	0·967 (0·00882)	0·960 (0·0102)	0·00737 (-0·00383 to 0·0186)
Household head completed at least eight years of schooling	0·288 (0·0192)	0·319 (0·0199)	-0·0310 (-0·0675 to 0·00552)
Count of assets (max 13)	8·202 (0·111)	8·234 (0·117)	-0·0327 (-0·196 to 0·130)
Has latrine	0·943 (0·00795)	0·946 (0·00781)	-0·00251 (-0·0215 to 0·0165)
Got help from government or NGO to build toilet (given has toilet)	0·625 (0·0304)	0·656 (0·0274)	-0·0309 (-0·0692 to 0·00747)
Pit size (cubic feet, given has toilet)	179·0 (14·33)	180·7 (14·27)	1·666 (-25·94 to 22·61)
Toilet looks used (given has toilet)	0·805 (0·0234)	0·794 (0·0233)	0·0103 (-0·0243 to 0·0450)

Standard errors clustered by village under means in columns 1 and 2. 95% CI under differences in column 3.

\*\* $p < 0.01$ , \*  $p < 0.05$ .

**Table 2: Individual-level, balanced latrine use questions find significantly higher rates of open defecation than household level questions**

unit of observation: question type:	individuals		households		difference (1)-(2) (5)	difference (4)-(3) (6)
	individual	household (imputed)	household	individual (imputed)		
	(1)	(2)	(3)	(4)		
1. Full sample						
Estimates	0.324 (0.0288)	0.115 (0.0150)	0.122 (0.0158)	0.326 (0.0289)	0.209** (0.163 to 0.254)	0.204** (0.160 to 0.248)
n (individuals or households)	13,070	13,070	2,431	2,431		
2. World Vision sample (in Bihar)						
Estimates	0.153 (0.0361)	0.0184 (0.00796)	0.0256 (0.0112)	0.154 (0.0321)	0.134** (0.0659 to 0.202)	0.128** (0.0719 to 0.184)
n (individuals or households)	3,675	3,675	629	629		
3. Coastal Salinity Prevention Cell sample (in Gujarat)						
Estimates	0.192 (0.0358)	0.0848 (0.0212)	0.0721 (0.0174)	0.169 (0.0299)	0.108** (0.0355 to 0.180)	0.0973** (0.0417 to 0.153)
n (individuals or households)	3,340	3,340	628	628		
4. Wateraid sample (in Karnataka)						
Estimates	0.558 (0.0580)	0.184 (0.0365)	0.193 (0.0406)	0.550 (0.0575)	0.374** (0.262 to 0.486)	0.358** (0.247 to 0.468)
n (individuals or households)	3,112	3,112	593	593		
5. Rural Welfare Institute sample (in Odisha)						
Estimates	0.434 (0.0484)	0.204 (0.0284)	0.208 (0.0294)	0.450 (0.0463)	0.230** (0.150 to 0.309)	0.242** (0.170 to 0.313)
n (individuals or households)	2,943	2,943	581	581		
6. Households with toilets only						
Estimates	0.289 (0.0284)	0.0734 (0.0137)	0.0791 (0.0150)	0.291 (0.0287)	0.215** (0.166 to 0.265)	0.211** (0.162 to 0.261)
n (individuals or households)	12,366	12,366	2,296	2,296		
7. Households that received help						
Estimates	0.375 (0.0370)	0.0895 (0.0181)	0.0968 (0.0193)	0.378 (0.0368)	0.285** (0.220 to 0.351)	0.281** (0.217 to 0.346)
n (individuals or households)	7,958	7,958	1,470	1,470		

Standard errors clustered by village under means in columns 1 through 4. 95% CI under differences in columns 5 and 6. \*\* $p < 0.01$ , \* $p < 0.05$ . Household imputed refers to estimates of individual open defecation imputed from answers to the household-level question. Each individual in the household is given the same answer as the household-level answer. Individual imputed refers to estimates of household open defecation imputed from answers to the individual level questions. The household estimates are constructed by averaging open defecation among individuals in the household.

**Table 3: Sub-group analyses**

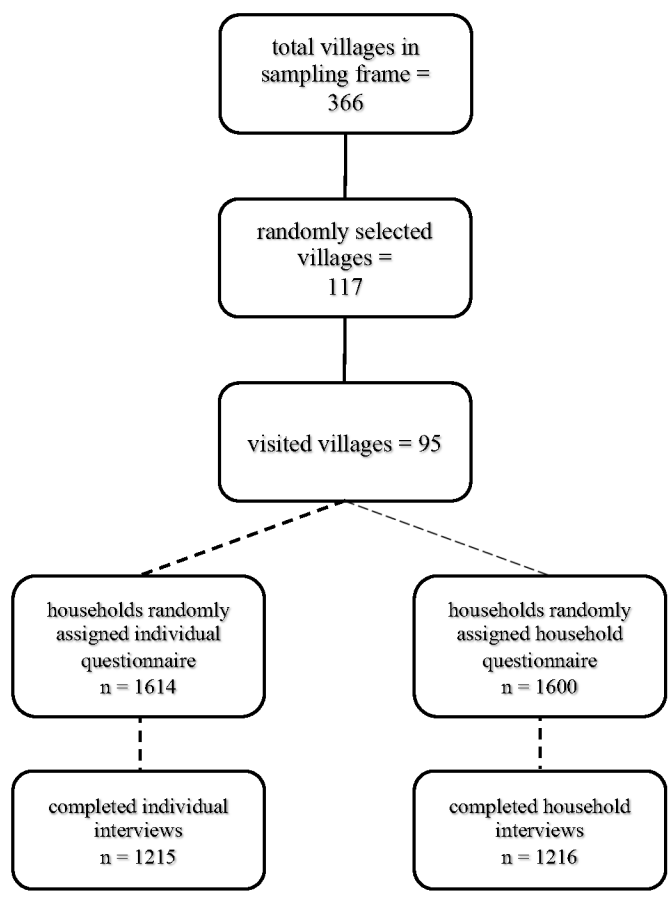
unit of observation:	individuals (1)	households (2)	individuals (3)
Individual-level question	0.100** (0.0548 to 0.146)	0.0992** (0.0566 to 0.142)	0.229*** (0.182 to 0.275)
Received help for construction	0.0470* (0.00962 to 0.0844)	0.0514** (0.0134 to 0.0893)	
Individual-level question X Received help	0.185** (0.113 to 0.257)	0.182** (0.114 to 0.251)	
Female			0.00134 (-0.0118 to 0.0144)
Individual-level X Female			-0.0409*** (-0.0646 to -0.0172)
Constant	0.0425** (0.0180 to 0.0670)	0.0455** (0.0201 to 0.0708)	0.115*** (0.0846 to 0.145)
n (individuals or households)	12,366	2,296	13,070

95% CI under coefficients, calculated using standard errors clustered by village. \*\* $p < 0.01$ , \*  $p < 0.05$ .

1  
2  
3 *Figure 2 caption: Figure shows means and confidence intervals. Red, dashed lines indicate responses to individual-*  
4 *level, balanced latrine use questions, and blue solid lines indicate responses to the household-level question. Unit of*  
5 *observation is individuals for individual-level questions and households for household-level questions. Confidence*  
6 *intervals are computed using standard errors clustered by village.*  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

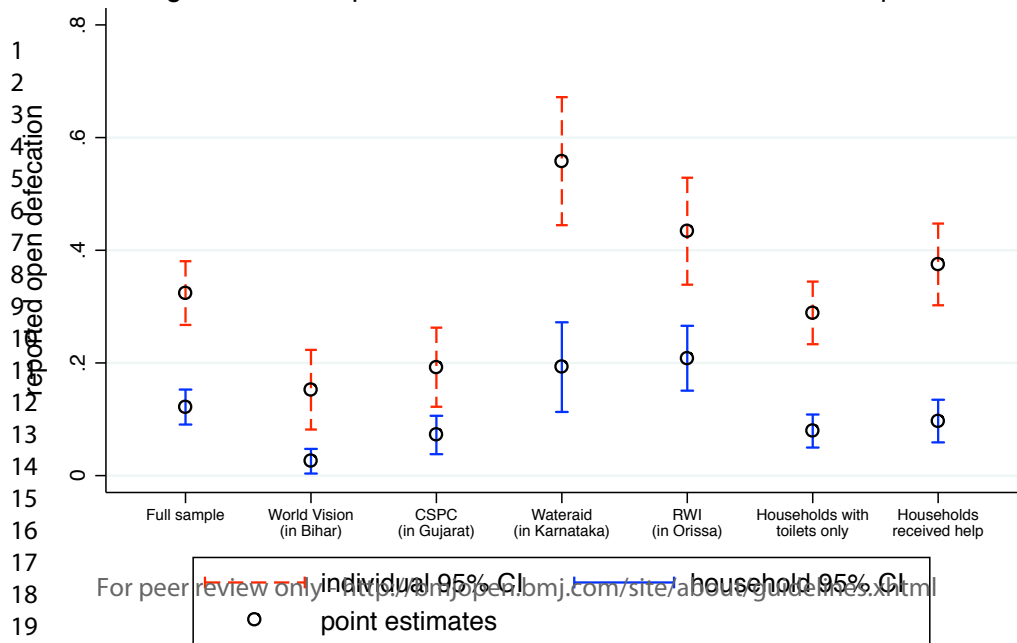
For peer review only

Figure 1: Sample selection



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Figure 2: Individual-level, balanced questionnaire use questions find significantly higher rates of open defecation than the household-level question





## CONSORT 2010 checklist of information to include when reporting a randomised trial\*

Section/Topic	Item No	Checklist item	Reported on page No
<b>Title and abstract</b>			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	1
<b>Introduction</b>			
Background and objectives	2a	Scientific background and explanation of rationale	4
	2b	Specific objectives or hypotheses	4
<b>Methods</b>			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	6
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	n/a
Participants	4a	Eligibility criteria for participants	5
	4b	Settings and locations where the data were collected	5
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	6-7
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	7
	6b	Any changes to trial outcomes after the trial commenced, with reasons	n/a
Sample size	7a	How sample size was determined	6
	7b	When applicable, explanation of any interim analyses and stopping guidelines	n/a
<b>Randomisation:</b>			
Sequence generation	8a	Method used to generate the random allocation sequence	6
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	6
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	6
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	6
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those	n/a

		assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	7
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	7-8
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	8
<b>Results</b>			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	figure 1
	13b	For each group, losses and exclusions after randomisation, together with reasons	figure 1, page 6
Recruitment	14a	Dates defining the periods of recruitment and follow-up	6
	14b	Why the trial ended or was stopped	n/a
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	16
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	16
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	17
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	17
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	18
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	n/a
<b>Discussion</b>			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	12
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	12
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	10-12
<b>Other information</b>			
Registration	23	Registration number and name of trial registry	5
Protocol	24	Where the full trial protocol can be accessed, if available	n/a
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	2

\*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see [www.consort-statement.org](http://www.consort-statement.org).



# BMJ Open

## Measuring open defecation in India using survey questions: Evidence from a randomized survey experiment

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-030152.R1
Article Type:	Original research
Date Submitted by the Author:	27-Jul-2019
Complete List of Authors:	Vyas, Sangita; University of Texas at Austin, Economics and Population Research Center Srivastav, Nikhil; University of Texas at Austin, LBJ School of Public Affairs Mary, Divya; Institute for Financial Management and Research, IFMR LEAD Goel, Neeta; International Initiative for Impact Evaluation Dehli, N/A Srinivasan, Sujatha; Institute for Financial Management and Research, IFMR LEAD Tannirkulam, Ajaykumar; Institute for Financial Management and Research, IFMR LEAD Ban, Radu; Bill and Melinda Gates Foundation, Evidence and Measurement, WSH Program Spears, D; University of Texas at Austin, Economics and Population Research Center Coffey, Diane; University of Texas at Austin, Sociology and Population Research Center
<b>Primary Subject Heading</b>:	Public health
Secondary Subject Heading:	Health policy
Keywords:	India, open defecation, sanitation, survey methods, experiment

SCHOLARONE™  
Manuscripts

# Measuring open defecation in India using survey questions: Evidence from a randomized survey experiment

Sangita Vyas, Nikhil Srivastav, Divya Mary, Neeta Goel, Sujatha Srinivasan, Ajaykumar Tannirkulam, Radu Ban, Dean Spears, Diane Coffey

**Objectives:** To investigate differences in reported open defecation between a question about latrine use or open defecation for every household member and a household-level question.

**Setting:** Rural India is home to most of the world's open defecation. India's Demographic and Health Survey (DHS) 2015-2016 estimates that 54% of households in rural India defecate in the open. This measure is based on a question asking about the behavior of all household members in one question. Yet, studies in rural India find substantial open defecation among individuals living in households with latrines, suggesting that household-level questions underestimate true open defecation.

**Participants:** In 2018, we randomly assigned latrine-owning households in rural parts of four Indian states to receive one of two survey modules measuring sanitation behavior. 1,215 households were asked about latrine use or open defecation individually for every household member. 1,216 households were asked the household-level question used in India's DHS: what type of facility do members of the household usually use?

**Results:** We compare reported open defecation between households asked the individual-level questions and those asked the household-level question. Using two methods for comparing open defecation by question type, the individual-level question found 20 to 21 (95% CI 16 to 25 for both estimates) percentage points more open defecation than the household-level question, among all households, and 28 to 29 (95% CI 22 to 35 for both estimates) percentage points more open defecation among households that received assistance to construct their latrines.

**Conclusions:** We provide the first evidence that individual-level questions find more open defecation than household-level questions. Because reducing open defecation in India is essential to meeting the Sustainable Development Goals, and exposure to open defecation has consequences for child mortality and development, it is essential to accurately monitor its progress.

**Registration:** Registry for International Development Impact Evaluations, number 5b55458ca54d1.

**Keywords:** India, open defecation, sanitation, survey methods, experiment

## Affiliations

**Economics, University of Texas at Austin, Austin, TX, USA (S Vyas MPA, D Spears PhD);  
Population Research Center, University of Texas at Austin, Austin, TX, USA (S Vyas MPA,**

1  
2  
3 D Spears PhD, D Coffey PhD); **LBJ School of Public Affairs, University of Texas at Austin,**  
4 **Austin, TX, USA** (N Srivastav BA); **IFMR Lead, Chennai, India** (D Mary MS, S Srinivasan  
5 MBA, A Tannirkulam PhD); **International Initiative for Impact Evaluation (3ie), New Delhi,**  
6 **India** (N Goel PhD); **Bill & Melinda Gates Foundation** (R Ban PhD); **Indian Statistical**  
7 **Institute, Delhi, India** (D Spears PhD, D Coffey PhD).  
8  
9

10 Correspondence to

11 Sangita Vyas, Economics, University of Texas at Austin, Austin, TX 78712, USA;  
12 sangita.vyas@utexas.edu  
13

14 Funding

15 International Initiative for Impact Evaluation, Bill & Melinda Gates Foundation.  
16  
17

18 Role of funders

19 Radu Ban, who is employed on the WSH program of BMGF, contributed to the analysis of the  
20 results. Neeta Goel contributed to coordinating data collection. Both contributed inputs to  
21 drafting the report.  
22  
23

24 Contributors

25 SV NS DS DC contributed to the study design. SV NS DC DS designed the survey instruments.  
26 NS DM SS AT oversaw data collection. SV NS NG coordinated between the 3ie research teams  
27 and the survey team for this study. SV RB DC DS contributed to analysis. All authors  
28 contributed to redrafting the report.  
29  
30

31 Acknowledgements

32 We are very grateful for the collaboration and contributions of the 3ie research teams in  
33 formulating the individual-level questions. The questions are the product of a deliberative  
34 process between the authors and the research teams. We are also grateful for the research teams'  
35 cooperation on this study. In particular, we thank researchers from Oxford Policy Management,  
36 London School of Hygiene and Tropical Medicine, Indian Institute of Public Health  
37 Gandhinagar, Eawag, and Emory University, and the implementing agencies with which they are  
38 working, including World Vision, Coastal Salinity Prevention Cell, Wateraid, and Rural Welfare  
39 Institute, for sharing their census data with us, coordinating field work, and supporting this  
40 endeavor. This study would not have been possible without their collaboration and assistance.  
41  
42

43 Declaration of interests

44 We have no competing interests.  
45  
46

47 Data sharing

48 Deidentified observation-level data will be made publicly available online upon publication of  
49 this article. The website will be provided at the proofing stage.  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Strengths and limitations

- This is the first study that experimentally tests the difference in open defecation estimated from different survey questions.
- This study provides evidence that the India's Demographic and Health Survey, which asks about the defecation behavior of everyone in the household in one question, substantially underestimates open defecation in India.
- The study shows that measuring open defecation at the individual level is feasible in a large household survey, and finds more open defecation than household-level questions.
- Monitoring open defecation in India is important for understanding progress towards the Sustainable Development Goals.
- Because the households participating in this study are not representative of the rural parts of the states they are in, or rural India, the estimates presented here should not be considered as estimates of open defecation for any of the states, or the country as a whole.

## Introduction

Rural India is home to more than half of the world's open defecation.<sup>1</sup> Because the persistence of open defecation threatens gains in child health, the Sustainable Development Goals (SDGs) call for its elimination by 2030. Progress towards eliminating open defecation in rural India will be essential to meeting this goal. India's most recent Demographic and Health Survey (DHS), conducted between January 2015 and December 2016, estimates that 54% of households in rural India defecated in the open, down from 75% in the 2005-2006 DHS.<sup>2,3</sup> This measure is based on a household-level question that asks about the behavior of everyone in the household in the same question.

Recent evidence from studies carried out in India suggests, however, that it is common for individuals living in households with latrines to nevertheless defecate in the open. In rural parts of five north Indian states, Coffey et al. found that 21% of individuals defecated in the open, despite owning a latrine.<sup>4</sup> In rural Tamil Nadu, Yogananth and Bhatnagar report that 54% of respondents defecated in the open despite having a household latrine.<sup>5</sup> In Odisha, Barnard et al. found that less than half of members of households with latrines reported using their latrines at all times.<sup>6</sup>

Experimental studies of sanitation interventions have found similar results. Clasen et al. report on a sanitation intervention in Odisha and note open defecation among individuals living in households with latrines as a reason for not observing impacts on child health outcomes.<sup>7</sup> Patil et al. conducted a sanitation intervention in Madhya Pradesh and experienced a similar problem: modest increases in latrine coverage, and even more modest reductions in open defecation.<sup>8</sup> These findings suggest that open defecation among latrine-owning households is substantial. Since it is probable that latrine use is the socially desirable response to questions on sanitation behavior, measures based on household-level questions, such as those from the DHS, will likely underestimate true open defecation in rural India, particularly among households with latrines.

Because open defecation is an individual behavior, an individual-level survey question may be able to more accurately measure it compared to a household-level question, particularly among households with latrines. We designed this study to experimentally test this hypothesis in rural India. We aimed to investigate whether a balanced question about latrine use or open defecation for every member of a household finds different levels of open defecation compared to a household-level question.

This is the first study to experimentally vary survey methodology to improve upon the measurement of open defecation currently being used. Jenkins et al. study sanitation survey methods, and report on an index they develop for quantifying household excreta disposal.<sup>9</sup> Their study focuses on developing and piloting a new tool rather than comparing different measures.

1  
2  
3 Sinha et al. compare answers to survey questions on latrine use behavior to measures of actual  
4 behavior generated from passive latrine use monitors that were set up in the latrines of  
5 respondents, and find poor to moderate agreement between the two measures.<sup>10</sup> Our study  
6 contributes to this literature by comparing estimates of open defecation obtained from questions  
7 that can be administered in a large household survey, and highlighting potential sources of error  
8 in open defecation measurement.  
9  
10

11  
12 Four years ago, the Government of India launched the Swachh Bharat Mission (SBM), a national  
13 sanitation campaign, which aims to eliminate open defecation in India by 2019. Many latrines  
14 have been constructed in rural India as a result of this campaign. Yet, the effect the SBM has had  
15 on reducing open defecation is still unknown. Because large reductions in open defecation in  
16 India are essential to meeting the SDGs, and because exposure to open defecation has serious  
17 consequences for child mortality, health, and human capital development, it is essential to  
18 monitor its progress as accurately as possible.  
19  
20  
21  
22

## 23 Methods

24  
25  
26 The study received ethical approval for research involving human subjects from the Institute for  
27 Financial Management and Research's Institutional Review Board in India, Approval #  
28 IRB00007107, and is registered in the Registry for International Development Impact  
29 Evaluations (RIDIE), number 5b55458ca54d1.  
30  
31  
32

### 33 Sample: mostly latrine-owning households in rural parts of four states

34  
35  
36 This study uses as its sampling frame the study areas of 3ie's Promoting Latrine Use in Rural  
37 India Thematic Window. This Window has funded four independent research teams to conduct  
38 randomized control trials of distinct behavioral campaigns to promote the use of pit latrines in  
39 rural parts of Bihar, Gujarat, Karnata, and Odisha. The study areas are spread across India,  
40 representing different contexts and varying levels of rural open defecation.  
41  
42  
43

44  
45 Because these trials focus on behavioral strategies rather than latrine construction, they are being  
46 carried out in villages that had high levels of coverage of pit latrines at baseline, relative to other  
47 rural parts of the same states. The households that comprise the sampling frame for this study are  
48 those that were identified as having a functional latrine in a census conducted by the research  
49 teams in the villages in which they were working. In all states except for Odisha, only  
50 households that had been excluded from the research teams' samples could be selected for this  
51 study. In Odisha, the sample selected for this study overlaps with the research teams' sample. We  
52 aimed to survey households that own latrines because we expect that an important source of  
53 misreporting of open defecation comes from individuals who do not use the functional latrines  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 that their households own, and that other members of their households use.  
4  
5

6 Figure 1 describes the sample selection. The villages visited in each state were randomly selected  
7 from the full set of villages included in the 3ie research teams' studies. The full set of villages  
8 were selected by the research teams in collaboration with the implementation agencies they were  
9 working with. The research team led by Oxford Policy Management worked with World Vision  
10 in Bihar, the team led by London School of Hygiene and Tropical Medicine worked with Coastal  
11 Salinity Prevention Cell (CSPC) in Gujarat, Eawag worked with Wateraid in Karnataka, and  
12 Emory University worked with the Rural Welfare Institute (RWI) in Odisha. Data for our study  
13 were collected in 22 to 25 villages in each of the four study areas. In most areas, we sampled  
14 more villages than we actually visited in order to facilitate coordination with the research teams.  
15 95 villages were visited in total.  
16  
17  
18  
19

20 Up to 40 households in each village were randomly assigned to receive the household or  
21 individual questions. In some villages, fewer than 40 households were assigned because fewer  
22 than 40 households met the eligibility criteria. The survey team visited as many assigned  
23 households as it could in these villages, given time constraints, and availability of household  
24 members. On average, the survey team interviewed 25 households per village. Data collection  
25 took place between March and July 2018.  
26  
27  
28  
29

30 In each state, data collection took place after the 3ie research teams had conducted their censuses  
31 and baselines, but before they had started their interventions. Since in all states, the households  
32 visited in this study were also visited at the time of the census, response bias may be a concern.  
33 This would not, however, impact the internal validity of this study since randomization generates  
34 equal response bias, in expectation, across treatment arms.  
35  
36  
37

### 38 Randomization and masking: random variation in latrine use questions at the 39 household level 40 41 42

43 We randomly assigned the type of latrine use question administered in the survey at the  
44 household level. Roughly half of the households were assigned individual-level questions on  
45 latrine use. The other half were assigned a household-level question. One of the authors who was  
46 not involved in data collection carried out the randomization using a random number generator in  
47 Stata. Because of the nature of the study, it was not possible to blind the respondents or  
48 surveyors to the type of survey question administered in the survey. However, in the interest of  
49 data quality, respondents were not explicitly told that the primary purpose of the survey, which  
50 took approximately 25 minutes to complete, was to measure open defecation. Additionally,  
51 surveyors did not know which survey question had been assigned to a household until starting  
52 the survey with the household. This was facilitated through SurveyCTO, the mobile data  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 collection platform used in the study, which was programmed to store the randomization  
4 assignment for each household ID prior to the commencement of data collection. In the field,  
5 surveyors were only given a list of households to interview. When a surveyor had correctly  
6 identified a household and was ready to start the survey, she would enter the household ID into  
7 SurveyCTO, and SurveyCTO would automatically start the questionnaire type assigned to the  
8 household.  
9  
10

11  
12 The individual-level questions asked for every household member age five or older whether the  
13 individual defecated in the open or used the latrine. The preface to this series of questions was: “I  
14 have seen that some people defecate in the open, and some people use the latrine. Now I want to  
15 ask about where you and your family members defecate.” Then, the surveyor asked the following  
16 question for each individual in the household, and coded the answer in a household roster: “The  
17 last time [*name of household member*] defecated, did [*name of household member*] defecate in  
18 the open or use the latrine?” The answer options included latrine, open, and somewhere else.  
19 Surveyors used the last option, which meant that the household member defecated in a bedpan,  
20 cloth, or other place, in less than 0.5% of cases. Because the priming statement and the behavior  
21 question include both open defecation and latrine use, they are balanced between the two  
22 different behaviors and could reduce social desirability bias. The surveyor asked household  
23 members who were participating in the interview directly about their behavior, and asked the  
24 main respondent, in most cases an adult female member of the household, to report on the  
25 behavior of their family members who were not participating in the interview.  
26  
27  
28  
29  
30  
31

32 The rest of the households were assigned the household-level question used in India’s DHS:  
33 “What kind of toilet facility do members of your household usually use?”<sup>11</sup> The answer codes  
34 were also the same as those used in India’s DHS: flush to piped sewer system, flush to septic  
35 tank, flush to pit latrine, flush to somewhere else, flush to don’t know where, ventilated  
36 improved pit or biogas latrine, pit latrine with slab, pit latrine without slab or open pit, twin pit or  
37 composting toilet, dry toilet, and no facility or uses open space or field. We also included an  
38 individual-level question on mobile ownership or preferring vegetarian food versus non-  
39 vegetarian food in the surveys that asked the household-level question so that both types of  
40 surveys would take approximately the same amount of time to complete.  
41  
42  
43  
44

45 There are three main factors that differ between the two types of latrine use survey modules: the  
46 level of aggregation, the reference period, and the presence of a priming statement. Therefore,  
47 the differences in reported open defecation that we observe reflect the fact that the two sets of  
48 questions vary on all of these factors combined.  
49  
50  
51

## 52 Statistical analyses

53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 The primary outcome of interest is reported open defecation. For the household-level questions,  
4 we created a dummy variable that is equal to one if the response was “no facility or uses open  
5 space or field,” and zero otherwise. The unit of observation for households assigned the  
6 individual-level question is the individual, while the unit of observation for those assigned the  
7 household-level question is the household. Therefore, in order to directly compare and test the  
8 significance of differences in reported open defecation between the two question types, we  
9 construct estimates that use the same unit of observation. We impute individual-level open  
10 defecation from responses to the household-level questions, and household-level open defecation  
11 rates from responses to the individual-level questions. To construct individual-level open  
12 defecation using the household-level questions, we assign the answer from the household  
13 question to each individual in the household. Similarly, to construct household-level open  
14 defecation using the individual-level questions, we average the responses among individuals in  
15 the household, and assign this average as the household value. Our main analysis tests  
16 differences in measured open defecation by question type. We show pooled results, as well as  
17 results by study area.  
18  
19  
20  
21  
22  
23

24 We also conduct sub-group analyses. These analyses investigate differences in the same primary  
25 outcome measure, reported open defecation, but look at differences by question type among  
26 different subgroups. First, we investigate whether the difference in reported open defecation by  
27 question type depends on whether the latrine was constructed privately, or with assistance from  
28 the government or an NGO. In practice, assistance to construct latrines often comes from the  
29 government, but sometimes NGOs get involved in facilitating the implementation of the  
30 government program. As part of the SBM, the Government of India assists rural households to  
31 construct latrines either by providing financial assistance directly to households so they can  
32 construct their own latrines, or by local government officials constructing latrines for  
33 households. In the discussion that follows, we will describe a household as having “received  
34 help” if it received financial assistance or a partially or completely constructed latrine from the  
35 government or an NGO.  
36  
37  
38  
39  
40  
41

42 The Indian government promotes and constructs latrines with pits that are approximately 60  
43 cubic feet.<sup>12</sup> However, many rural Indians aspire to construct latrines with pits that are much  
44 larger, so that they can avoid emptying the pit, a task that is associated with ritual pollution.<sup>13–16</sup>  
45 Compared to latrines constructed privately, those constructed with government help are less  
46 likely to be used due to concerns over purity and pit emptying. Since a large fraction of rural  
47 households are likely to receive latrines with help from the government as a result of the SBM, it  
48 is important to explore how much open defecation different types of latrine use questions  
49 measure, based on having received help to construct the latrine.  
50  
51  
52

53 The second subgroup analysis investigates whether the difference in measured open defecation  
54 between the two question types is statistically different for males compared to females. Sex  
55  
56  
57  
58  
59  
60

1  
2  
3 differences are an important aspect to explore because observational studies have found  
4 consistently higher open defecation among latrine owners for males compared to females.<sup>4</sup> This  
5 observation could reflect greater demand for latrine use among females due to, for instance,  
6 greater psychosocial stress experienced when defecating in the open,<sup>17</sup> or it could be because of  
7 cultural norms that keep females in their reproductive years inside the home.  
8  
9

10  
11 Means and differences in reported open defecation, by question type, are calculated using  
12 ordinary least squares regression with cluster robust standard errors, clustered by village.  
13 Statistical analyses were conducted using Stata (version 11).  
14  
15

## 16 17 Public involvement

18  
19 The individual-level questions used in this study are the product of a deliberative process  
20 between the authors and research teams from Oxford Policy Management, London School of  
21 Hygiene and Tropical Medicine, Indian Institute of Public Health Gandhinagar, Eawag, and  
22 Emory University. The findings of this study contribute to our understanding of the scale of an  
23 important public health problem.  
24  
25  
26

## 27 28 Results

29  
30  
31 Table 1 shows summary statistics for households assigned the two types of latrine use questions.  
32 The total sample consisted of 2,431 households, which were approximately equally divided  
33 across question type in each of the study areas. There were no significant differences on  
34 measures relevant for latrine use between households assigned different types of latrine use  
35 questions. Households in both groups had approximately the same number of household  
36 members, fraction female, fraction Hindu, educational attainment of the household head, and  
37 asset ownership of 13 assets, including mobile phone, electricity, radio, television, fan, mosquito  
38 net, bicycle, motorcycle, car, chair, gas stove, pressure cooker, and shoes for everyone in the  
39 family. It is important for the validity of the results that the sample is balanced on religious  
40 composition, since studies have documented an association between household religion and  
41 latrine use.<sup>18,19</sup>  
42  
43  
44  
45  
46

47 As the study design intended, most households in both groups had a latrine. Conditional on  
48 having a latrine, 64% of households had gotten help from the government or an NGO to build the  
49 latrine, and the average pit size was 180 cubic feet. It is important that the sample is balanced on  
50 these two characteristics because, compared to latrines constructed privately, those constructed  
51 with government help are less likely to be used because of concerns over purity and pit  
52 emptying.<sup>13-16</sup> Finally, among households with latrines, approximately 80% of them appeared to  
53 the surveyor to be in use upon observation.  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Figure 2 presents the main results of the study; it shows means and 95% confidence intervals from the individual-level (shown as dashed red bars) and the household-level (shown as solid blue bars) questions, for the full dataset and for different sub-samples. Observations are individuals for the individual-level estimates, and households for the household-level estimates. In the full sample, and in all sub-samples analyzed in this figure, the individual questions find more open defecation. Moreover, the size of the difference in reported open defecation between question types is large and consistent. The first set of estimates shown in the figure uses the full sample. The second, third, fourth, and fifth sets of estimates break the sample up by project area. The sixth set of estimates uses only households with latrines. No matter how the data are broken up, the individual-level, balanced latrine use questions find significantly higher rates of open defecation than the household-level question.

Table 2 shows actual and imputed open defecation rates, measured at the individual and household levels. Columns 1 and 3 are the estimates shown in Figure 2. Columns 2 and 4 show imputed open defecation rates, at the individual and household level, respectively. Imputed values are calculated based on the method described in the Statistical Analyses section. Column 5 shows the difference in the measured rate of open defecation between the individual and household questions when observations are individuals, and column 6 shows the same difference when observations are households.

In the full sample, the individual-level, balanced questions find 21 (95% CI 16 to 25) percentage points more open defecation than the household-level question when observations are individuals, and 20 (95% CI 16 to 25) percentage points more open defecation when observations are households. Notably, the individual-level questions measure consistently higher levels of open defecation in the full sample and in all subsamples, irrespective of how the difference is calculated. All differences are significant at the one percent level.

The seventh set of estimates shown in Figure 2 show reported open defecation from individual- and household-level questions among households that received help to construct their latrines. Comparing these estimates to the sixth set of estimates in the figure, which include all households with latrines, suggests that the household-level question underestimates open defecation by more among those that received help to construct their latrines, compared to those that did not. The seventh row in Table 2 shows that, using different methods for computing differences, individual-level questions measure 28 to 29 (95% CI 22 to 35 for both estimates) percentage points more open defecation than the household-level question, among those that received help to construct their latrines, compared to those that did not.

The first two columns in Table 3 test whether the difference in measured open defecation between the two question types is statistically different among households that received help to

1  
2  
3 construct their latrines compared to households that did not receive help to construct their  
4 latrines. The first column in Table 3 uses individuals as observations, and the second column  
5 uses households. The coefficients in the third row represent the difference-in-differences  
6 estimate.  
7  
8  
9

10 The difference in measured open defecation between the two question types is 18 (95% CI 11 to  
11 25) to 19 (95% CI 11 to 26) percentage points larger among households that received help to  
12 construct their latrines, compared to households that did not, depending on the method used to  
13 compute the difference. Among households that did not receive help to construct their latrines,  
14 the individual-level questions find 10 (95% CI 5 to 15) percentage points more open defecation  
15 than the household-level question.  
16  
17  
18

19 The last column in Table 3 investigates whether the difference in measured open defecation  
20 between the two question types is statistically different among males compared to females. Using  
21 the full sample, individual questions find 4 (95% CI 2 to 7) percentage points more open  
22 defecation than the household question among males compared to females. Breaking the sample  
23 up by sex, the individual questions find 23 (95% CI 18 to 28) percentage points more open  
24 defecation than the household question among males, and 19 (95% CI 14 to 24) percentage  
25 points more among females. The household question underestimates open defecation by more  
26 among males compared to females.  
27  
28  
29  
30

## 31 Discussion

32  
33

34 Our findings show that in our sample, individual-level, balanced questions find 20 to 21 (95%  
35 confidence intervals 16 to 25 for both estimates) percentage points more open defecation than the  
36 household-level question. This is both a statistically significant and practically important  
37 difference. This study presents compelling evidence that India's DHS, which provides the most  
38 recent nationally representative estimates of open defecation for rural India, and other surveys  
39 that ask household-level questions, greatly underestimate open defecation among households  
40 with latrines.  
41  
42  
43  
44

45 We also found that the difference in reported open defecation between the two question types is  
46 significantly greater for households that received help to construct their latrines compared to  
47 households that did not. Among households that received help to construct their latrines, the  
48 individual-level questions find 28 to 29 (95% confidence intervals 22 to 35 for both estimates)  
49 percentage points more open defecation than the household-level question. This suggests that as  
50 more and more households receive government assistance for a latrine through the Swachh  
51 Bharat Mission, household-level questions will become even less accurate at estimating open  
52 defecation.  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 The larger difference in measured open defecation between the two question types among  
4 households that received help compared to those that did not is likely arising from higher rates of  
5 open defecation among households that received help to construct their latrines. There are  
6 several reasons that could explain why households receiving assistance may be less likely to use  
7 their latrines. First, these households are likely to have lower demand for latrine use, compared  
8 to households that built latrines on their own. Second, households that received help have latrines  
9 with pits that are on average 150 cubic feet smaller than the pits of latrines in households that did  
10 not receive help. Because of concerns over ritual purity, rural Indians are less likely to use  
11 latrines with pits that need to be emptied manually every few years, like the latrines that are  
12 promoted and constructed by the government.<sup>13-16</sup> Whether only one, both, or other factors are  
13 leading to more open defecation among households that received help, the individual-level  
14 questions are better able to capture this open defecation than the household-level question.  
15  
16  
17  
18  
19

20 We also find a statistically significant difference in reported open defecation between the two  
21 question types for males compared to females. The difference between the individual-level,  
22 balanced questions and the household-level question is 4 (95% CI 2 to 7) percentage points more  
23 for males compared to females. This supports evidence that, conditional on latrine ownership,  
24 males are more likely to defecate in the open compared to females. Individual-level questions  
25 understate the difference in open defecation between the two sexes by less than the household-  
26 level questions. Compared to the difference in reported open defecation by receiving help to  
27 construct the latrine, differences by sex are not as large.  
28  
29  
30  
31

32 Measuring open defecation at the individual level is feasible. Our survey team's experience  
33 suggests that adding the balanced, individual-level questions on use to a survey that already  
34 contains a household roster increases survey time by about 2 minutes, on average. Of course, the  
35 amount of time required to ask the individual-level questions depends on the number of  
36 individuals in the household.  
37  
38  
39

40 The household-level question asked in the DHS also collects information on the types of latrines  
41 that households own, data that is still of great interest to researchers and practitioners. Therefore,  
42 individual questions on use, combined with a separate question on the types of latrines that  
43 households own, would satisfy both goals: evaluating latrine infrastructure, and measuring open  
44 defecation as accurately as possible.  
45  
46  
47

48 A limitation of our study is that the samples from these project areas are not representative of the  
49 rural parts of the states they are in, nor are they collectively representative of rural India. The  
50 households in this study are much more likely to have a latrine than the average rural Indian  
51 household, and therefore, the individuals in this study are more likely to use a latrine. For this  
52 reason, the estimates presented here should not be considered as estimates of open defecation for  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 any of the states, or the country as a whole. Rather, they show a large and significant difference  
4 in reported open defecation based on the type of question asked.  
5  
6

7 Measuring open defecation at the individual-level in a large household survey is doable and will  
8 provide a more accurate estimate of open defecation in rural India. Since reducing open  
9 defecation in India is important for meeting the SDGs, and since open defecation is an important  
10 factor contributing to poor health among children in India, it is important to measure its progress  
11 as accurately as possible.  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## References

- 1 WHO and Unicef. *Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines*. Geneva: World Health Organization (WHO) and the United Nations Children's Fund (UNICEF), 2017.
- 2 International Institute for Population Sciences (IIPS) and ICF. National Family Health Survey (NFHS-4), 2015-16 [Dataset]. IAHR74DT.DTA. India. Mumbai: IIPS [Producers]. ICF [Distributor], 2017.
- 3 International Institute for Population Sciences (IIPS) and Macro International. National Family Health Survey (NFHS-3), 2005-06 [Dataset]. IAHR52DT.DTA. India. Mumbai: IIPS [Producers]. ICF [Distributor], 2007.
- 4 Coffey, D, A Gupta, P Hathi, et al. Revealed preference for open defecation. *Economic & Political Weekly* 2014, 49(38), 43-55.
- 5 Yogananth, N, & T Bhatnagar. Prevalence of open defecation among households with toilets and associated factors in rural south India: an analytical cross-sectional study. *Transactions of The Royal Society of Tropical Medicine and Hygiene* 2018, 112(7), 349-360.
- 6 Barnard, S, P Routray, F Majorin, et al. Impact of Indian Total Sanitation Campaign on latrine coverage and use: a cross-sectional study in Orissa three years following programme implementation. *PloS One* 2013, 8(8), e71438.
- 7 Clasen, T, S Boisson, P Routray, et al. Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial. *The Lancet Global Health* 2014, 2(11), e645-e653.
- 8 Patil, SR, BF Arnold, AL Salvatore, et al. The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: a cluster randomized controlled trial. *PLoS Medicine* 2014, 11(8), e1001709.
- 9 Jenkins, MW, MC Freeman, & P Routray. Measuring the safety of excreta disposal behavior in India with the new Safe San Index: Reliability, validity and utility. *International Journal of Environmental Research and Public Health* 2014, 11(8), 8319-8346.
- 10 Sinha, A, CL Nagel, E Thomas, WP Schmidt, B Torondel, S Boisson, & TF Clasen. Assessing latrine use in rural India: a cross-sectional study comparing reported use and passive latrine use monitors. *The American Journal of Tropical Medicine and Hygiene* 2016, 95(3), 720-727.

- 1  
2  
3 11 International Institute for Population Sciences (IIPS) and ICF. National Family Health  
4 Survey (NFHS-4), 2015-16. Household Questionnaire. India. Mumbai: IIPS, 2017.  
5  
6  
7 12 Ministry of Drinking Water and Sanitation. Guidelines for Swachh Bharat Mission  
8 (Gramin). Government of India, New Delhi, India, 2017.  
9  
10 13 Coffey, D, A Gupta, P Hathi, et al. Understanding open defecation in rural India:  
11 Untouchability, pollution, and latrine pits. *Economic & Political Weekly* 2017, 52(1), 59-  
12 66.  
13  
14 14 O'Reilly, K, R Dhanju, & E Louis. Subjected to sanitation: Caste relations and sanitation  
15 adoption in rural Tamil Nadu. *The Journal of Development Studies* 2017, 53(11), 1915-  
16 1928.  
17  
18 15 Routray, P., WP Schmidt, S Boisson, T Clasen, & MW Jenkins. Socio-cultural and  
19 behavioural factors constraining latrine adoption in rural coastal Odisha: an exploratory  
20 qualitative study. *BMC Public Health* 2015, 15(1), 880.  
21  
22  
23  
24 16 Coffey, D, & D Spears. *Where India goes: Abandoned toilets, stunted development and*  
25 *the costs of caste*. Harper Collins, India, 2017.  
26  
27  
28  
29 17 Sahoo, KC, KR Hulland, BA Caruso, et al. Sanitation-related psychosocial stress: a  
30 grounded theory study of women across the life-course in Odisha, India. *Social Science*  
31 *& Medicine* 2015, 139, 80-89.  
32  
33  
34 18 Vyas, S, & D Spears. Sanitation and Religion in South Asia: What Accounts for  
35 Differences across Countries? *The Journal of Development Studies* 2018, 54(11), 1-17.  
36  
37 19 Geruso, M, & D Spears. Neighborhood sanitation and infant mortality. *American*  
38 *Economic Journal: Applied Economics* 2018, 10(2), 125-62.  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



**Table 1: Randomization balance: No significant differences in observed means between households assigned different latrine use questions**

	individual (1)	household (2)	difference (3)
Number of households	1,215	1,216	-1
Number of households by project			
World Vision (in Bihar)	316	313	3
Coastal Salinity Prevention Cell (in Gujarat)	309	319	-10
Wateraid (in Karnataka)	297	296	1
Rural Welfare Institute (in Odisha)	293	288	5
Household members	5·685 (0·0950)	5·604 (0·101)	0·0803 (-0·117 to 0·278)
Female	0·491 (0·00485)	0·489 (0·00490)	0·00190 (-0·0115 to 0·0153)
Hindu	0·967 (0·00882)	0·960 (0·0102)	0·00737 (-0·00383 to 0·0186)
Household head completed at least eight years of schooling	0·288 (0·0192)	0·319 (0·0199)	-0·0310 (-0·0675 to 0·00552)
Count of assets (max 13)	8·202 (0·111)	8·234 (0·117)	-0·0327 (-0·196 to 0·130)
Has latrine	0·943 (0·00795)	0·946 (0·00781)	-0·00251 (-0·0215 to 0·0165)
Got help from government or NGO to build toilet (given has toilet)	0·625 (0·0304)	0·656 (0·0274)	-0·0309 (-0·0692 to 0·00747)
Pit size (cubic feet, given has toilet)	179·0 (14·33)	180·7 (14·27)	1·666 (-25·94 to 22·61)
Toilet looks used (given has toilet)	0·805 (0·0234)	0·794 (0·0233)	0·0103 (-0·0243 to 0·0450)

Cluster robust standard errors, clustered by village, under means in columns 1 and 2. 95% CI under differences in column 3. \*\* $p < 0.01$ , \*  $p < 0.05$ .

**Table 2: Individual-level, balanced latrine use questions find significantly higher rates of open defecation than household level questions**

unit of observation: question type:	individuals		households		difference (1)-(2) (5)	difference (4)-(3) (6)
	individual (1)	household (imputed) (2)	household (3)	individual (imputed) (4)		
1. Full sample						
Estimates	0.324 (0.0288)	0.115 (0.0150)	0.122 (0.0158)	0.326 (0.0289)	0.209** (0.163 to 0.254)	0.204** (0.160 to 0.248)
n (individuals or households)	13,070	13,070	2,431	2,431		
2. World Vision sample (in Bihar)						
Estimates	0.153 (0.0361)	0.0184 (0.00796)	0.0256 (0.0112)	0.154 (0.0321)	0.134** (0.0659 to 0.202)	0.128** (0.0719 to 0.184)
n (individuals or households)	3,675	3,675	629	629		
3. Coastal Salinity Prevention Cell sample (in Gujarat)						
Estimates	0.192 (0.0358)	0.0848 (0.0212)	0.0721 (0.0174)	0.169 (0.0299)	0.108** (0.0355 to 0.180)	0.0973** (0.0417 to 0.153)
n (individuals or households)	3,340	3,340	628	628		
4. Wateraid sample (in Karnataka)						
Estimates	0.558 (0.0580)	0.184 (0.0365)	0.193 (0.0406)	0.550 (0.0575)	0.374** (0.262 to 0.486)	0.358** (0.247 to 0.468)
n (individuals or households)	3,112	3,112	593	593		
5. Rural Welfare Institute sample (in Odisha)						
Estimates	0.434 (0.0484)	0.204 (0.0284)	0.208 (0.0294)	0.450 (0.0463)	0.230** (0.150 to 0.309)	0.242** (0.170 to 0.313)
n (individuals or households)	2,943	2,943	581	581		
6. Households with toilets only						
Estimates	0.289 (0.0284)	0.0734 (0.0137)	0.0791 (0.0150)	0.291 (0.0287)	0.215** (0.166 to 0.265)	0.211** (0.162 to 0.261)
n (individuals or households)	12,366	12,366	2,296	2,296		
7. Households that received help						
Estimates	0.375 (0.0370)	0.0895 (0.0181)	0.0968 (0.0193)	0.378 (0.0368)	0.285** (0.220 to 0.351)	0.281** (0.217 to 0.346)
n (individuals or households)	7,958	7,958	1,470	1,470		

Cluster robust standard errors, clustered by village, under means in columns 1 through 4. 95% CI under differences in columns 5 and 6. \*\* $p < 0.01$ , \* $p < 0.05$ . Household imputed refers to estimates of individual open defecation imputed from answers to the household-level question. Each individual in the household is given the same answer as the household-level answer. Individual imputed refers to estimates of household open defecation imputed from answers to the individual level questions. The household estimates are constructed by averaging open defecation among individuals in the household.

**Table 3: Sub-group analyses**

unit of observation:	individuals (1)	households (2)	individuals (3)
Individual-level question	0.100** (0.0548 to 0.146)	0.0992** (0.0566 to 0.142)	0.229*** (0.182 to 0.275)
Received help for construction	0.0470* (0.00962 to 0.0844)	0.0514** (0.0134 to 0.0893)	
Individual-level question X Received help	0.185** (0.113 to 0.257)	0.182** (0.114 to 0.251)	
Female			0.00134 (-0.0118 to 0.0144)
Individual-level X Female			-0.0409*** (-0.0646 to -0.0172)
Constant	0.0425** (0.0180 to 0.0670)	0.0455** (0.0201 to 0.0708)	0.115*** (0.0846 to 0.145)
n (individuals or households)	12,366	2,296	13,070

95% CI under coefficients, calculated using cluster robust standard errors, clustered by village. \*\* $p < 0.01$ , \* $p < 0.05$ .

1  
2  
3 *Figure 1 title: Sample selection*

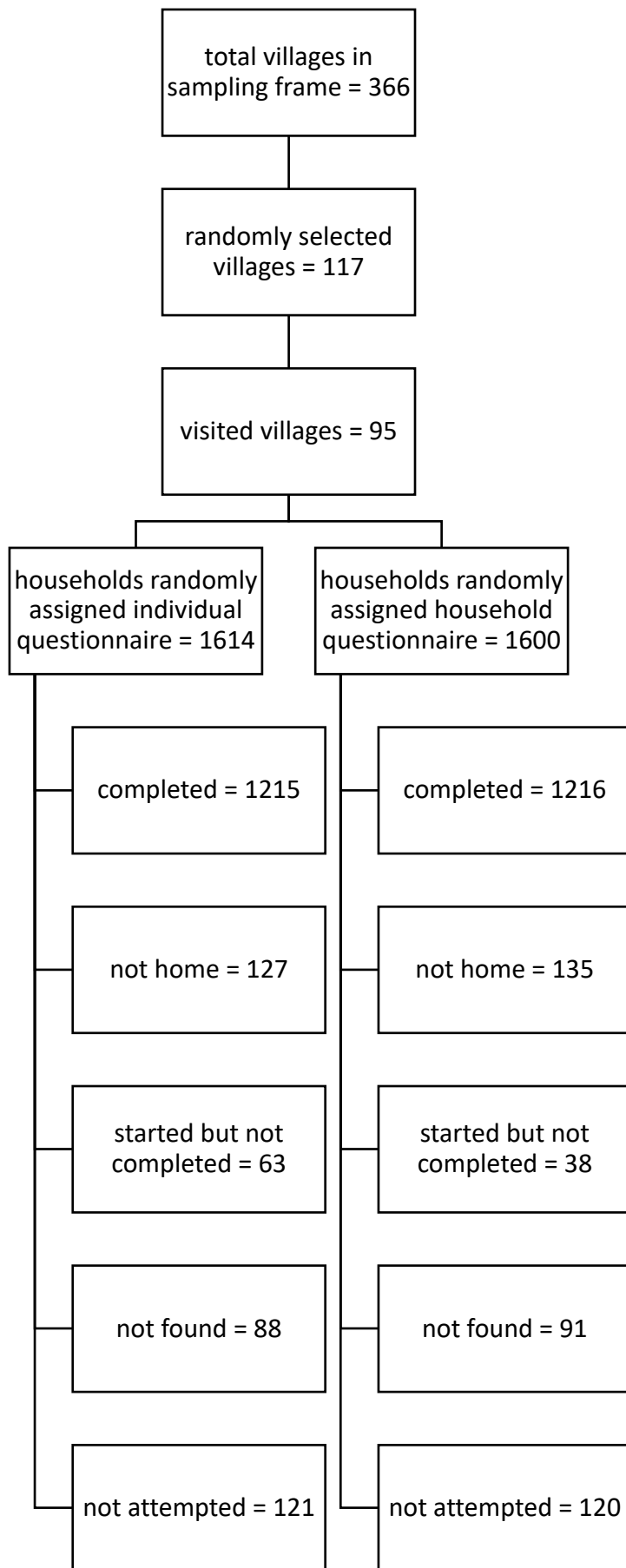
4 *Figure 1 caption: Figure details sample selection. Households that were started but not completed refers to*  
5 *households that refused at the beginning or part of the way through and households in which a suitable respondent*  
6 *was not available.*

7  
8 *Figure 2 title: Individual-level, balanced latrine use questions find significantly higher rates of open defecation than*  
9 *the household-level question*

10 *Figure 2 caption: Figure shows means and confidence intervals. Red, dashed lines indicate responses to individual-*  
11 *level, balanced latrine use questions, and blue solid lines indicate responses to the household-level question. Unit of*  
12 *observation is individuals for individual-level questions and households for household-level questions. Confidence*  
13 *intervals are computed using standard errors clustered by village.*

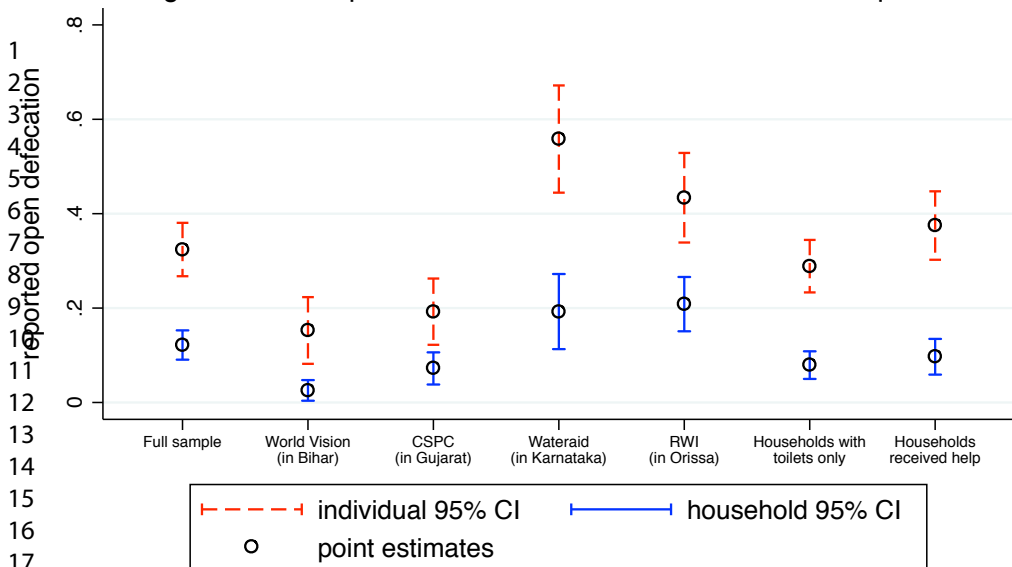
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For peer review only



Note: Figure details sample selection. Households that were started but not completed refers to households that refused at the beginning or part of the way through and households in which a suitable respondent was not available.

Figure 2: Individual-level, balanced latrine use questions find significantly higher rates of open defecation than the household-level question



Note: Figure shows means and confidence intervals. Red, dashed lines indicate responses to individual-level, balanced latrine use questions, and blue solid lines indicate responses to the household-level question. Unit of observation is individuals for individual-level questions and households for household-level questions. Confidence intervals are computed using standard errors clustered by village.



## CONSORT 2010 checklist of information to include when reporting a randomised trial\*

Section/Topic	Item No	Checklist item	Reported on page No
<b>Title and abstract</b>			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	1
<b>Introduction</b>			
Background and objectives	2a	Scientific background and explanation of rationale	4
	2b	Specific objectives or hypotheses	4
<b>Methods</b>			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	6
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	n/a
Participants	4a	Eligibility criteria for participants	5
	4b	Settings and locations where the data were collected	5
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	6-7
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	7
	6b	Any changes to trial outcomes after the trial commenced, with reasons	n/a
Sample size	7a	How sample size was determined	6
	7b	When applicable, explanation of any interim analyses and stopping guidelines	n/a
<b>Randomisation:</b>			
Sequence generation	8a	Method used to generate the random allocation sequence	6
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	6
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	6
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	6
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those	n/a

1			
2		assessing outcomes) and how	
3			
4		11b If relevant, description of the similarity of interventions	<u>7</u>
5	Statistical methods	12a Statistical methods used to compare groups for primary and secondary outcomes	<u>7-8</u>
6		12b Methods for additional analyses, such as subgroup analyses and adjusted analyses	<u>8</u>
7			
8	<b>Results</b>		
9	Participant flow (a	13a For each group, the numbers of participants who were randomly assigned, received intended treatment, and	
10	diagram is strongly	were analysed for the primary outcome	<u>figure 1</u>
11	recommended)	13b For each group, losses and exclusions after randomisation, together with reasons	<u>figure 1, page 6</u>
12	Recruitment	14a Dates defining the periods of recruitment and follow-up	<u>6</u>
13		14b Why the trial ended or was stopped	<u>n/a</u>
14	Baseline data	15 A table showing baseline demographic and clinical characteristics for each group	<u>16</u>
15	Numbers analysed	16 For each group, number of participants (denominator) included in each analysis and whether the analysis was	
16		by original assigned groups	<u>16</u>
17			
18	Outcomes and	17a For each primary and secondary outcome, results for each group, and the estimated effect size and its	
19	estimation	precision (such as 95% confidence interval)	<u>17</u>
20		17b For binary outcomes, presentation of both absolute and relative effect sizes is recommended	<u>17</u>
21	Ancillary analyses	18 Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing	
22		pre-specified from exploratory	<u>18</u>
23			
24	Harms	19 All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	<u>n/a</u>
25			
26	<b>Discussion</b>		
27	Limitations	20 Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	<u>12</u>
28	Generalisability	21 Generalisability (external validity, applicability) of the trial findings	<u>12</u>
29	Interpretation	22 Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	<u>10-12</u>
30			
31	<b>Other information</b>		
32	Registration	23 Registration number and name of trial registry	<u>5</u>
33	Protocol	24 Where the full trial protocol can be accessed, if available	<u>n/a</u>
34	Funding	25 Sources of funding and other support (such as supply of drugs), role of funders	<u>2</u>
35			
36			

37 \*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also  
 38 recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials.  
 39 Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see [www.consort-statement.org](http://www.consort-statement.org).