

**Research Paper****A data-driven approach to optimize a sanitary survey for communities in Maputo, Mozambique**Lee E. Voth-Gaeddert <sup>a,b</sup><sup>a</sup> Biodesign Center for Health Through Microbiomes, Arizona State University, Tempe, AZ, USA<sup>b</sup> Center for Indigenous Health Research, Wuqu' Kawoq | Maya Health Alliance, Tecpan, Guatemala

E-mail: lee.gaeddert@asu.edu

 LEV, 0000-0002-9541-1958**ABSTRACT**

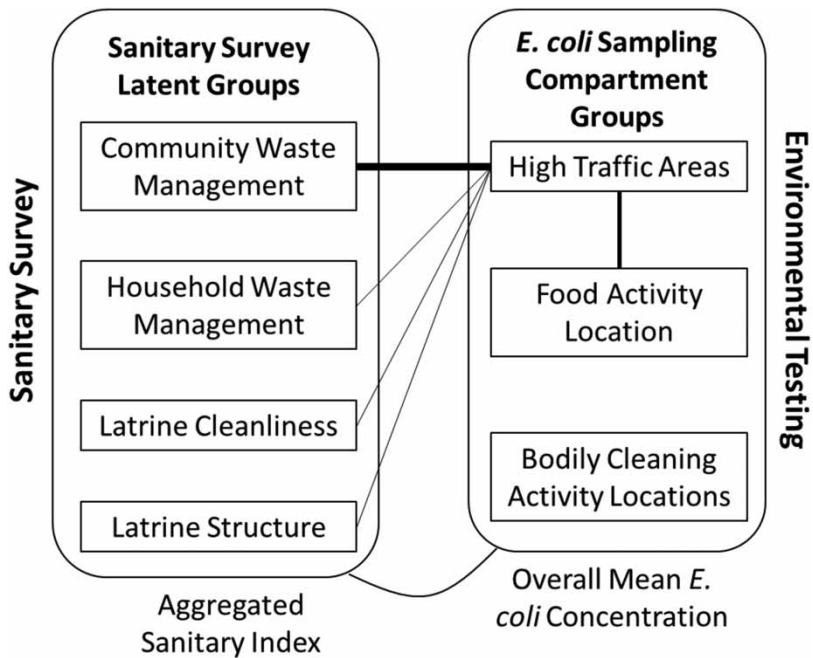
Current approaches to align sanitary surveys to environmental fecal waste indicators have demonstrated limited success, leaving stakeholders focused on fecal waste management with limited and costly tools. We analyzed data from the Maputo Sanitation Project using exploratory factor analysis (EFA) and structural equation modeling (SEM) to enhance the survey's correlation with *Escherichia coli* concentration data. EFA grouped related survey questions and sampling locations into distinct latent factors. SEM was then used to assess the relationship between the survey latent factors and the mean *E. coli* concentrations from grouped locations. The results suggested four survey question subgroups: latrine structure, latrine cleanliness, compound/household waste management, and community waste management. In addition, three sampling location subgroups were identified: high-traffic areas (HTAs), food activity locations, and bodily cleaning areas. The largest significant effect size identified suggested that for every 1-unit increase in community waste management, there was a decrease of 1.94 in log<sub>10</sub> *E. coli* per gram of soil in HTAs ( $p = 0.03$ ), a substantial improvement from the initial 0.50 decrease reported for an expert-weighted metric of all survey questions. These results underscore the importance of community-level waste management and demonstrate the use of a data-driven approach in enhancing environmental health assessments and planning interventions.

**Key words:** community sanitation, environmental health, exploratory factor analysis, fecal indicator bacteria, structural equation modeling

**HIGHLIGHTS**

- Enhanced sanitary surveys using exploratory factor analysis to correlate survey data with environmental *E. coli* concentrations.
- Identified significant latent factors impacting *E. coli* concentrations in high-traffic areas.
- Demonstrated that improved community waste management practices are significantly associated with *E. coli* concentrations in high-traffic areas.

## GRAPHICAL ABSTRACT



## INTRODUCTION

Sanitary surveys are an important tool for local environmental health specialists in monitoring various aspects of physical systems that are used to control contaminants in the environment and reduce human exposure (Gallego-Ayala *et al.* 2018). Monitoring typically focuses on the pathways of exposure or the fecal-oral transmission pathways primarily captured in the '5-F diagram' of floors, foods, fluids, fingers, and flies (Wagner & Lanoix 1958). These tools are often used at a community or regional scale and guide decisions on (1) adjustments to current water or sanitation systems or services or (2) future planning for water and sanitation services (Jenkins *et al.* 2014). In addition, sanitary surveys and general water, sanitation, and hygiene (WASH) surveys embedded in national surveys (e.g., demographic and health surveys) have been used to help evaluate sanitary conditions at a household- or community-level as environmental testing can be expensive (Voth-Gaedert *et al.* 2018a,b). In addition, concepts of the source-pathway-receptor model of environmental health can help guide decisions around control interventions (Mbae *et al.* 2024). Several attempts have been made to validate observational metrics with environmental testing (often using fecal indicator bacteria (FIB) as a proxy for overall contamination) (Jenkins *et al.* 2014; Campos *et al.* 2015; Ercumen *et al.* 2017; Kelly *et al.* 2021), however, with limited success. This can be for several reasons, not limited to observations of infrastructure not capturing true microbial contamination or the dynamism of environmental contamination and the limited proxies used to measure it. Recently, Capone *et al.* (2019) conducted a robust evaluation of a sanitary survey adapted from the World Bank's Urban Sanitation Status Index and corresponding soil and surface swab samples testing for *Escherichia coli* among compounds in Maputo, Mozambique (a set of 3–5 households) (Capone *et al.* 2019). A localized sanitation survey index (LSSI) generated based on expert weights and responses from a sanitary survey was correlated with *E. coli* concentrations in soil. However, a 10% increase in the LSSI only corresponded to a 0.05 log<sub>10</sub> reduction in *E. coli*/gram of dry soil. The authors state, 'Overall, the LSSI may be associated with fecal contamination in compound soil; however, the differences detected may not be meaningful in terms of public health hazards.' While disappointing, a shift to an exploratory, data-driven approach could help increase this reduction and optimize the survey to more closely align with environmental testing results.

To improve alignment or optimize the sanitary survey to environmental testing data, it is important to understand (1) which survey questions are associated with which sampling locations and (2) which groups of questions and sampling locations are related to each other. Both the survey questions as well as the environmental sampling approaches are often proxies for underlying 'latent' factors that cannot be directly measured. For surveys, questions are often a proxy for the infrastructure

functionality or use and adherence to WASH recommendations while, for environmental testing, *E. coli* (or qPCR targets) is often used as a general proxy for microbial pathogen contamination or an indicator of fecal waste contamination. To better understand which sets of questions and which sets of sampling locations represent the same underlying latent variables, exploratory factor analysis (EFA) can be used (Fabrigar & Wegener 2011). EFA provides a method to identify sets of variables that may all be affected by the same underlying latent factor. In addition, structural equation modeling (SEM) can then be used to evaluate the magnitude and direction of the relationship between these underlying latent variables and observable variables (Kline 2011; Voth-Gaeddert *et al.* 2018c).

In this study, a secondary data analysis was conducted on data collected in the Maputo Sanitation Project to build on previous work from Capone *et al.* (2019). First, EFA was applied to explore the clustering of similar sanitary survey questions and similar environmental sampling locations in compounds. Second, the identified clusters of survey questions (and their estimated latent variable) were compared to mean levels of *E. coli* across identified clusters of sampling locations. The aim was to identify which subgroups of survey questions provided the largest effect size on concentrations of *E. coli*.

## METHODS

### Location

The Maputo Sanitation Project was conducted in urban, low-income communities of Maputo, Mozambique. Maputo has a complex sanitary infrastructure and service system as well as a diverse socio-economic landscape. It is a city characterized by a mix of formal and informal settlements, with variations in sanitation facilities ranging from modern sewage systems to basic latrines (Capone *et al.* 2019). The city's coastal location, coupled with its dense population, presents unique challenges for sanitation management, making it an ideal location for studying the impact of sanitation practices on *E. coli* contamination. The study area consisted of compounds, which were groups of 3–5 homes or nuclear families.

### Sanitary survey and data collection

The secondary data were retrieved from the Center for Open Science's OSF archive (<https://osf.io/d967q/>) and Capone *et al.* (2019) provide an extensive overview of the survey development and sample collection. Briefly, the sanitary survey was a questionnaire based on the World Bank's Urban Sanitation Status Index (Gallego-Ayala *et al.* 2018). It covered various aspects of sanitation, including the type of facility, its shared use, structural features, and waste management practices (see Capone *et al.* (2019) for further details). Trained field staff conducted the surveys through interviews and direct observations within the study area. The survey was a nested cross-sectional sub-study of a larger before-after sanitary control trial.  $N = 80$  compounds were surveyed and were selected based on an evenly spaced geographic distribution. Survey data collected were ordinal variables from 0 to 1 with intermediate values utilized where necessary, divided evenly based on a number of ordinal levels. Capone *et al.* (2019) also created a weighted LSSI from the 20 survey questions and a panel of experts providing weights.

Environmental sampling included soil and surface swabs collected from pre-defined locations or compartments within each compound, such as entrances to compounds, central yard areas, areas for food preparation, and latrine entrances. These locations were identified by an adult member of the compound. Sampling and testing techniques followed previously established protocols for FIB (Campos *et al.* 2015). The final dataset generated by Capone *et al.* (2019) with complete data from both survey data and environmental sampling data included  $n = 75$  compounds.

### Data analysis

Both the survey data and the environmental testing results (i.e., *E. coli* data) were evaluated using EFA. EFA is a statistical method used for identifying underlying latent variables or factors that explain the observed correlations among a subset of measured variables (Fabrigar & Wegener 2011). EFA is similar yet distinct from principal component analysis (PCA) in that EFA aims to identify underlying latent factors explaining observed correlations by modeling shared variance among variables, while PCA focuses on maximizing total variance by transforming variables into new uncorrelated principal components.

First, EFA was utilized to identify latent variables from (1) the sanitary survey data and (2) *E. coli* concentrations from compound compartments, grouping related sampling compartments and survey questions into distinct latent factors. The resulting rotated component matrix and individual factor loadings were evaluated and used to identify which subgroup the (1) survey question or (2) sampling compartment should belong to. Based on previous EFA studies and prior hypotheses, three- and four-group models were evaluated for the survey question EFA while two- and three-group models were evaluated for the

sampling compartment EFA. Final models were selected based on eigenvalues, model fit indices, and previous literature. After selection, several variables in the survey question EFA did not have one clear high factor loading, and, in this situation, previous literature and expert judgement were used to place the variable into the most relevant group. Specifically, placement was based on the functional characteristics of the survey question and its alignment with identified EFA groups. As the goal of EFA is to inform downstream analysis of the latent factor groupings, expert judgement can aid this initial step when the signal-to-noise ratio is too low (Fabrigar *et al.* 1999; Costello & Osborne 2005; Brown 2015).

Second, SEM was used to evaluate the magnitude and direction of the associations between each of the subgroups of factors identified among the sanitary survey variables and *E. coli* concentrations from the subgroups of compartments. SEM is a multivariable statistical analysis technique that allows for the estimation of hypothesized causal relationships among observed and latent variables using a combination of linear regression and factor analysis (Kline 2011; Voth-Gaeddert & Oerther 2014). The arithmetic mean values of *E. coli* concentrations were estimated for each of the subgroups identified via EFA (given the evenly distributed geospatial sampling approach, no weighing beyond the original LSSI weighting was necessary). Each of these subgroups was individually compared to the latent variable estimated for each of the subgroups for the survey questions identified in the EFA. SEM estimates *p*-values and effect sizes which were used to identify the subgroup of survey questions that had the largest effect on the mean concentration of *E. coli* from a certain subgroup of compartments. In R, the *fa()* function in the psych package (Revelle 2023) was used for EFA (*nfactors* = \*, *method* = oblique, *fm* = ml) while the *sem()* function in the LAVAAN package (Rosseel *et al.* 2023) was used for SEM (*estimator* = 'WLSMV'). All SEM models controlled for the presence of ducks, chickens, and sunlight as well as the socio-economic status via a compound wealth score. Further details are provided in the Supplementary material and in Capone *et al.* (2019).

## RESULTS

### EFA of *E. coli* concentrations across sampling compartments

Capone *et al.* (2019) and results from our secondary analysis suggested that there were varying levels of *E. coli* concentrations across different sampling compartments within the compounds. However, EFA was able to identify distinct patterns within subgroups of these compartments (see Table 1 and Figure 1). Upon inspection, the compartment subgroups could be placed into three categories based on factor loadings (see the Supplementary material for further output details). High-traffic areas (HTAs), food activity locations (FALs), and bodily cleaning activity locations (BCALs) exhibited distinct patterns of *E. coli* contamination. The mean *E. coli* concentration for these three areas is shown in Table 2, with BCALs having the highest concentration (3.11 log<sub>10</sub> *E. coli*/gram soil) and HTAs having the lowest (2.71 log<sub>10</sub> *E. coli*/gram soil).

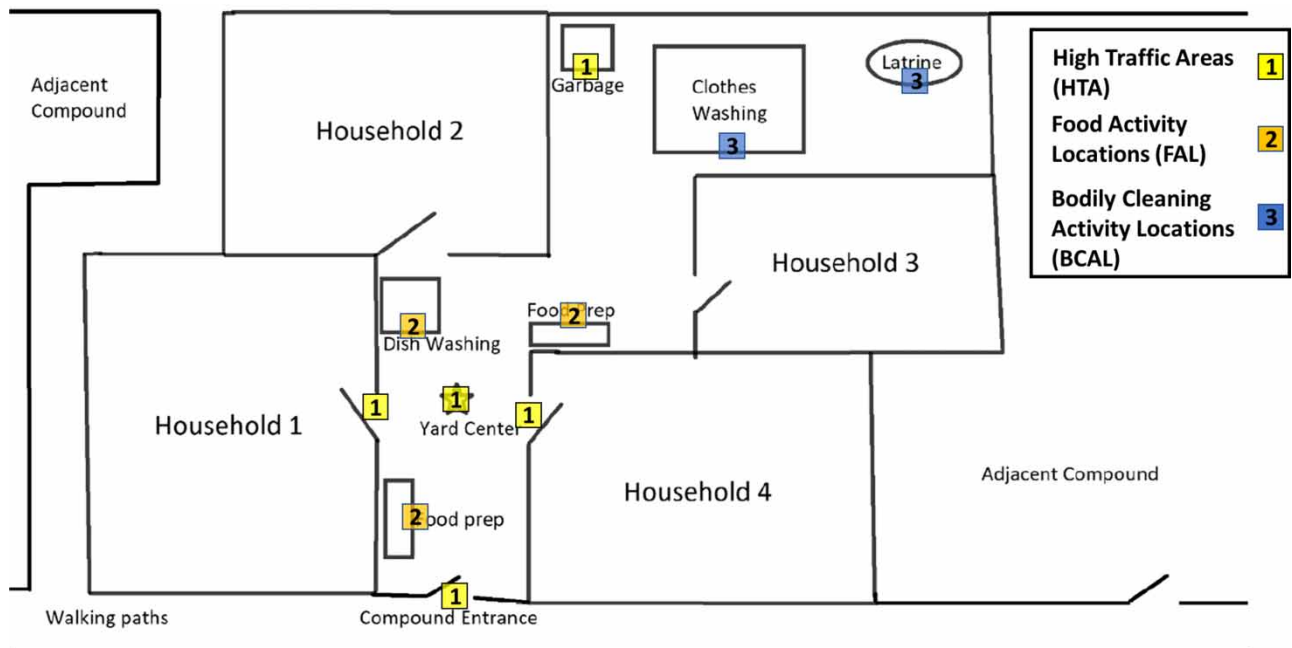
### EFA of sanitary survey questions

The EFA identified latent factor subgroupings of sanitary survey questions as well, however, some loadings of specific variables did not provide a clear categorization. Expert judgement was then used to place them into the most relevant subgroup. Based on this exercise, four primary groupings were identified: latrine structure, latrine cleanliness, compound or household

**Table 1** | EFA factor loadings on the nine compartments

Variable	HTA	FAL	BCAL
Compound entrance	0.65	0.02	0.07
Household entrance intervention <sup>a</sup>	0.62	-0.15	-0.02
Household entrance control <sup>a</sup>	0.48	0.32	-0.02
Yard center	0.43	-0.01	0.02
Garbage	0.46	0.27	0.06
Clothes washing	-0.08	0.36	0.44
Latrine entrance	0.02	-0.03	1
Dish washing	0.02	0.99	-0.01
Food prep.	0.1	0.36	0.2

<sup>a</sup>The Maputo Sanitation Project included an improved sanitation facility intervention where each compound had an intervention and control household. HTA, high-traffic area; FAL, food activity location; BCAL, bodily cleaning activity location. Shaded cells indicate which subgroup the variable best belongs to.



**Figure 1** | Depiction of where the sampling compartments are located in a typical compound. Base layout from Capone *et al.* (2019).

**Table 2** | Mean *E. coli* concentrations across latent subgroups identified in the EFA

Compartment	Mean <i>E. coli</i> concentrations
HTAs	2.71 (SD: 0.69)
FALs	2.98 (SD: 0.90)
BCALs	3.11 (SD: 0.85)

SD, standard deviation.

waste management, and community waste management. Table 3 presents the factor loadings and Figure 2 depicts the locations within the compound. Further detail on each survey question can be found in the supplementary material of Capone *et al.* (2019).

### Associations between sanitary survey question clusters and *E. coli* concentrations

Next, SEM was used to compare each combination of the survey question latent factor groupings ( $n = 4$ ) to the mean *E. coli* concentrations of the latent factor groupings ( $n = 3$ ). Of the 12 total comparisons, four had significant associations. Each of the survey question subgroupings (latrine cleanliness, latrine structure, compound waste, community waste) was significantly associated with HTA ( $p$ -value  $< 0.05$ ) suggesting this subgroup of sampling compartments is a key contamination location. The highest significant effect size identified was between the community waste latent factor subgrouping and the HTA at  $-1.94$  ( $p$ -value = 0.03; the three other effect sizes were all greater than  $-0.60$ ). This suggests that a 1-unit increase in the community waste subgroup of survey questions is associated with a 1.94 reduction of  $\log_{10}$  *E. coli*/gram of soil in high-trafficked areas. A 1-unit increase would involve improving community waste piles and dumping from present to absent and improving greywater and stormwater management (see Figure 3).

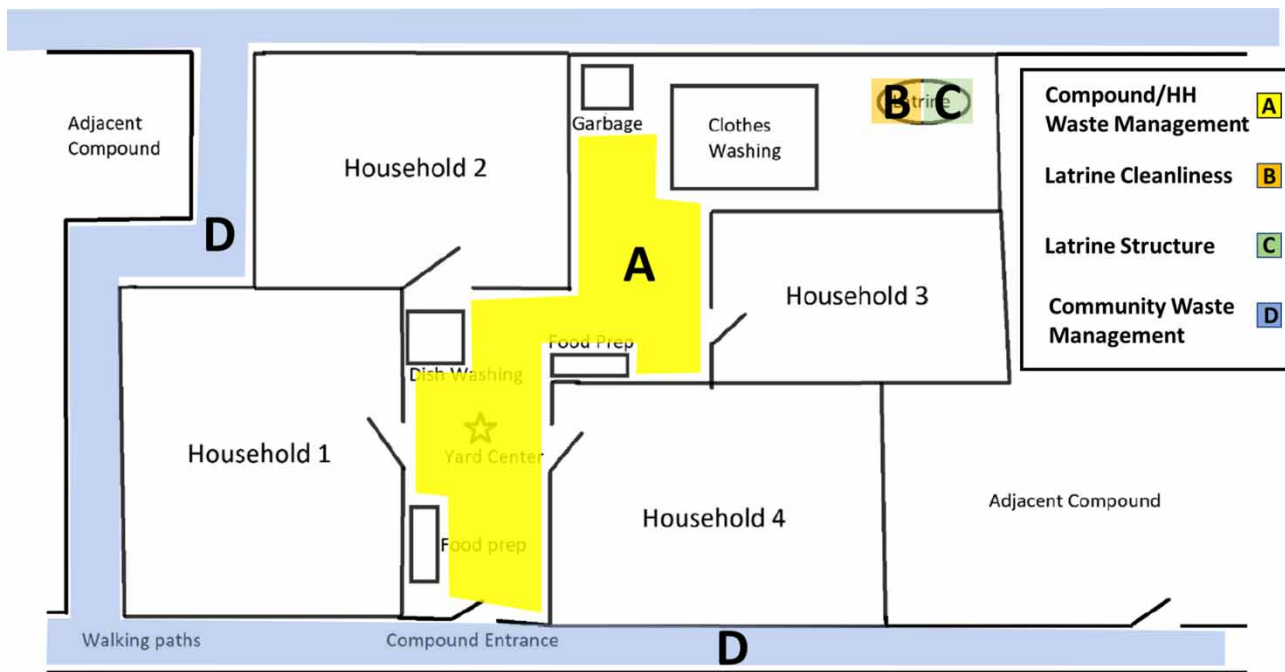
In addition, we also reevaluated the expert-weighted LSSI metric evaluated by Capone *et al.* (2019) and compared this to the three *E. coli* compartment subgroupings. Again, the HTAs subgroup was the only significant association ( $p < 0.05$ ) but produced an effect size of 0.82 (compared to the 0.50 found in Capone *et al.* 2019). This supports the previous finding that the HTA may play an important role in overall contamination. Finally, the only variables or factors significantly associated with the FAL *E. coli* compartment subgroup were the HTA and the presence of ducks (also associated with HTAs; see Supplementary material, Figure S1). No factors were associated with BCAL. Figure 4 depicts the integrated results of the analysis.

**Table 3** | EFA factor loadings on the 20 sanitary survey questions

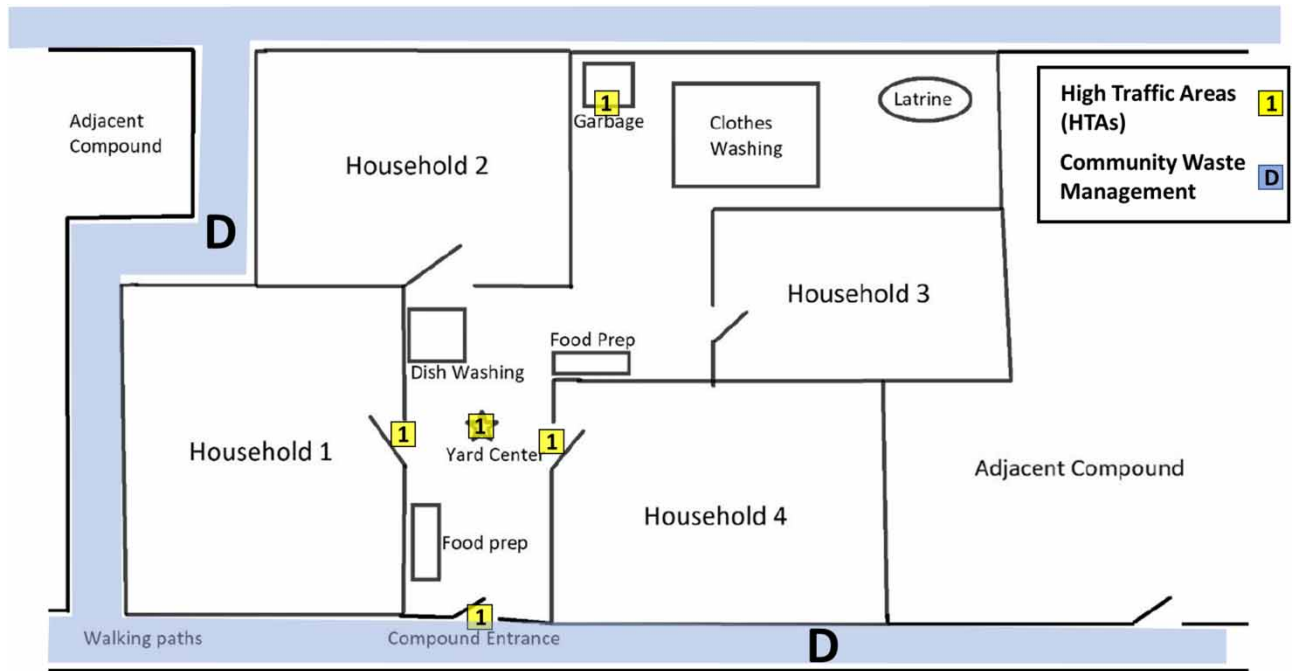
Variables	Latrine cleanliness	Latrine structure	Compound waste	Community waste
Sanitation type	0.46	0.1	0.23	0.09
Sanitation sharing	-0.42	-0.14	0.06	-0.17
Sanitation roof	0.95	-0.06	0.13	0.07
Sanitation containment effectiveness	0.48	0.03	-0.32	-0.18
Sanitation hygiene	0.42	0.33	0.01	-0.06
Sanitation pit/hole lid	0.53	0.29	-0.22	0
Water availability for flushing/cleaning <sup>a</sup>	0.22	-0.22	0.2	-0.03
Sanitation structural stability	0.03	0.91	0.1	0.08
Sanitation lining	-0.01	0.65	0.01	0.06
Sanitation walls	-0.02	0.51	-0.24	-0.21
Compound waste emptying method	0.06	-0.01	0.98	0.07
Compound waste dumping	0	0.17	0.77	-0.15
Quality of compound waste disposal	0.16	-0.12	0.51	-0.33
Compound waste piles <sup>a</sup>	0.24	-0.27	-0.2	-0.06
Hand soap at sanitation facility <sup>a</sup>	0.15	-0.31	0.26	-0.2
Community waste piles	0.1	0.11	-0.02	0.95
Greywater management	-0.23	-0.28	-0.02	0.57
Compound and community standing storm water	0.09	-0.08	0.02	0.38
Community waste dumping	-0.18	-0.12	0.05	-0.24

<sup>a</sup>Variables had unclear loadings, additional expert judgement was used to make the final decision on how to group them. Ground water level was removed due to lack of relevance and unclear factor loadings.

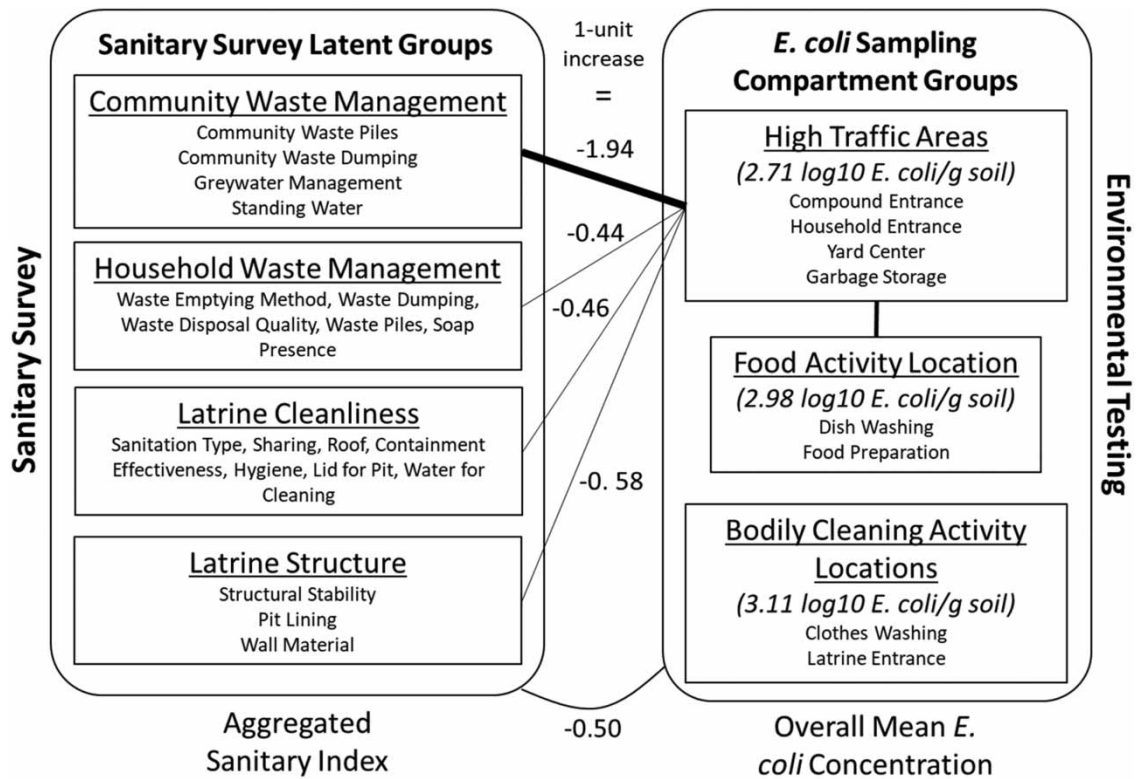
Shaded cells indicate which subgroup the variable best belongs to.



**Figure 2** | Depiction of where the sanitary survey questions are located in a typical compound.



**Figure 3** | Depiction of the locations of the two factors with the largest effect size.



**Figure 4** | Integrated results of survey question subgroups and *E. coli* sampling compartments. 1-unit increase refers to all variables of a latent group going from 0 to 1. Values are effect sizes of the relationship between the sanitary survey latent group and the *E. coli* sampling compartment.

## DISCUSSION AND CONCLUSION

### The efficacy of a data-driven, exploratory approach in sanitary survey optimization

Our study demonstrates how a data-driven, exploratory approach can significantly optimize tools like sanitary surveys. By systematically analyzing *E. coli* concentrations across different compartments and correlating these with various sanitation practices and infrastructure conditions, we identified key factors potentially contributing to fecal contamination. This approach could allow for more targeted interventions, given validation of initial findings, focusing on the most impactful areas and practices. While research studies often use an expert or hypothesis-driven approach, reframing the problem as a 'tool to be optimized' allows for a potentially more efficient approach to developing an effective and data-driven sanitary survey. Such an approach can be important to refine sanitary surveys, ensuring they are more focused, efficient, and contextually relevant. This strategy of optimization aligns with the evolving needs of urban sanitation management, where varied and complex factors interplay to influence public health outcomes.

### The role of community-level sanitation in household-level contamination

Our findings underscore the significance of the broader neighborhood or community context in influencing household-level contamination (Mombert *et al.* 2020). The association between community waste management practices and *E. coli* concentrations in HTAs highlights the interconnectedness of individual and community sanitation practices. This observation aligns with research indicating the importance of understanding community-level sanitation in interpreting individual health outcomes (Wolf *et al.* 2019) and its impact on child health (Fuller *et al.* 2016). Wolf *et al.* (2019) demonstrate why a sanitation or WASH intervention at an individual level may not have an effect on individual health outcomes due to community-level fecal waste contamination. In addition, Fuller *et al.* (2016) provide an excellent demonstration of how focusing on community-level fecal waste management issues can provide herd protection benefiting individuals. These studies emphasize that improvements in community sanitation infrastructure can have far-reaching effects on reducing household-level contamination risks. However, scrutiny must be given to identifying the critical components of community sanitation (e.g., waste piles, drainage) and identifying sustainable service models to alleviate these issues (Armitage & Rooseboom 2000). These studies and our results reemphasize the necessary shift back to the role of the sanitation service provider and other community stakeholders in improving fecal waste management practices, ultimately improving public health outcomes.

### Key aspects of fecal waste management in low-resource communities

In the context of low-resource communities, the management of fecal waste presents unique challenges and opportunities. In addition to the need for a focus on community-level fecal waste, there may be options for households to reduce their individual exposure profile. Interventions such as improved flooring can significantly reduce soil-based pathogen transmission and exposure (EarthEnable 2023). Furthermore, the implementation of protective barriers or habits, like grass cover, shoe removal before home entry, foot washing or shoe cleaning before home entry, or proper drainage systems, can also mitigate transmission and contamination risks. These measures, coupled with enhanced community waste management practices, can potentially decrease transmission and incidence of fecal pathogen-related diseases. These interventions are particularly crucial in densely populated urban areas, where the risk of contamination is heightened due to the close proximity of living spaces and the often-inadequate sanitation infrastructure or services. Effective fecal waste management in these settings requires a holistic approach, incorporating both household-level improvements and broader community-based strategies.

### Limitations of exploratory analysis and the need for confirmatory work

While the data-driven approach adopted in this study provides valuable insights, it is not without limitations. One key limitation is the inherent uncertainty in drawing definitive conclusions from exploratory analyses. These analyses are designed to generate hypotheses and identify potential associations rather than to confirm causal relationships. Therefore, there is a need for subsequent confirmatory studies to validate the findings and establish stronger causal links (Voth-Gaeddert *et al.* 2019). Such follow-up research could employ larger sample sizes and more granular environmental sampling to provide more definitive evidence. Furthermore, additional studies in varied geographical and cultural contexts would be beneficial to ascertain the generalizability of an optimized sanitary survey. Despite these limitations, our exploratory study provides a framework for optimizing sanitary surveys to aid low-resource communities and stakeholders in improving fecal waste management.



## CONCLUSIONS

This study aimed to address the limitations of sanitary surveys in urban, low-resource settings by employing a data-driven approach to identifying key related components of surveys and environmental fecal contamination. Using EFA and SEM, the research identified key latent factors within survey data and their relationship to *E. coli* concentrations in specific sampling compartments. The exploratory results highlighted the potentially significant impact of community-level waste management practices on reducing fecal contamination in HTAs among compounds. This approach underscores the importance of targeted interventions and provides a framework for improving environmental health assessments and planning effective sanitation interventions in low-resource urban settings.

## ACKNOWLEDGEMENTS

The author is grateful to the households, field teams, practitioners, and researchers involved in the original MapSan project.

## DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

## CONFLICT OF INTEREST

The authors declare there is no conflict.

## REFERENCES

- Armitage, N. & Rooseboom, A. 2000 The removal of urban litter from stormwater conduits and streams: The quantities involved and catchment litter management options. *Water SA* **26** (2), 1-1 and 13-5.
- Brown, T. A. 2015 *Confirmatory Factor Analysis for Applied Research* 2nd edn. Guilford Publications, New York.
- Campos, L. C., Ross, P., Nasir, Z. A., Taylor, H. & Parkinson, J. 2015 Development and application of a methodology to assess sanitary risks in Maputo, Mozambique. *Environment and Urbanization* **27** (2), 371–388. <https://doi.org/10.1177/0956247815595784>.
- Capone, D., Adriano, Z., Berendes, D., Cumming, O., Dreibelbis, R., Holcomb, D. A., Knee, J., Ross, I. & Brown, J. 2019 A localized sanitation status index as a proxy for fecal contamination in Urban Maputo, Mozambique. *PLoS One* **14** (10), e0224333. <https://doi.org/10.1371/journal.pone.0224333>.
- Costello, A. B. & Osborne, J. 2005 Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation* **10**. <https://doi.org/10.7275/JYJ1-4868>.
- EarthEnable 2023 *EarthEnable – Low Cost Flooring Solutions in Rwanda*. Available from: <https://earthenable.org/> (accessed 27-12-2023).
- Ercumen, A., Naser, A. M., Arnold, B. F., Unicomb, L., Colford, J. M. & Luby, S. P. 2017 Can sanitary inspection surveys predict risk of microbiological contamination of groundwater sources? Evidence from shallow tubewells in Rural Bangladesh. *The American Journal of Tropical Medicine and Hygiene* **96** (3), 561–568. <https://doi.org/10.4269/ajtmh.16-0489>.
- Fabrigar, L. R. & Wegener, D. T. 2011 *Exploratory Factor Analysis*. Oxford University Press, New York.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. & Strahan, E. J. 1999 Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods* **4** (3), 272–299. <https://doi.org/10.1037/1082-989X.4.3.272>.
- Fuller, J. A., Villamor, E., Cevallos, W., Trostle, J. & Eisenberg, J. N. S. 2016 I Get height with a little help from my friends: Herd protection from sanitation on child growth in rural Ecuador. *International Journal of Epidemiology* **45** (2), 460–469. <https://doi.org/10.1093/ije/dyv368>.
- Gallego-Ayala, J., Muximpua, O. & Hawkins, P. 2018 *Urban Sanitation Status Index: A Tool for Prioritizing Interventions at City Wide Level*. The World Bank Group, Washington, DC, p. 16.
- Jenkins, M. W., Freeman, M. C. & Routray, P. 2014 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* **11** (8), 8319–8346. <https://doi.org/10.3390/ijerph110808319>.
- Kelly, E., Cronk, R., Fisher, M. & Bartram, J. 2021 Sanitary inspection, microbial water quality analysis, and water safety in handpumps in Rural Sub-Saharan Africa. *NPJ Clean Water* **4** (1), 1–7. <https://doi.org/10.1038/s41545-020-00093-z>.
- Kline, R. B. 2011 *Principles and Practice of Structural Equation Modeling*. New York: The Guilford Press.
- Mbae, M., Hansen, P., Way, C., Mills, F., Willetts, J., Foster, T. & Evans, B. 2024 Onsite sanitation systems and contamination of groundwater: A systematic review of the evidence for risk using the source-pathway-receptor model. *PLOS Water* **3** (7), e0000167. <https://doi.org/10.1371/journal.pwat.0000167>.
- Momberg, D. J., Voth-Gaeddert, L. E., Ngandu, B. C., Richter, L., May, J., Norris, S. A. & Said-Mohamed, R. 2020 Water, sanitation, and hygiene (WASH) factors associated with growth between birth and 1 year of Age in children in Soweto, South Africa: Results from the Soweto Baby WASH study. *Journal of Water and Health* **18** (5). <https://doi.org/10.2166/wh.2020.085>.

- Revelle, W. 2023 *Psych: Procedures for Psychological, Psychometric, and Personality Research*. Available from: <https://cran.r-project.org/web/packages/psych/index.html> (accessed 28-12-2023).
- Rosseel, Y., Jorgensen, T. D., Wilde, L. D., Oberski, D., Byrnes, J., Vanbrabant, L., Savalei, V., Merkle, E., Hallquist, M., Rhemtulla, M., Katsikatsou, M., Barendse, M., Rockwood, N., Scharf, F., Du, H. & Jamil, H. 2023 *Lavaan: Latent Variable Analysis*. Available from: <https://cran.r-project.org/web/packages/lavaan/index.html> (accessed 18-12-2023).
- Voth-Gaeddert, L. E. & Oerther, D. B. 2014 **Utilizing structural equation modeling in the development of a standardized intervention assessment tool**. *Procedia Engineering* **78**, 218–223. <https://doi.org/10.1016/j.proeng.2014.07.059>.
- Voth-Gaeddert, L. E., Cudney, E. A. & Oerther, D. B. 2018a **Primary factors statistically associated with diarrheal occurrences**. *Environmental Engineering Science*. <https://doi.org/10.1089/ees.2017.0338>.
- Voth-Gaeddert, L. E., Stoker, M., Cornell, D. & Oerther, D. B. 2018b **What causes childhood stunting among children of San Vicente, Guatemala: Employing complimentary, system-analysis approaches**. *International Journal of Hygiene and Environmental Health*. <https://doi.org/10.1016/j.ijheh.2018.01.001>.
- Voth-Gaeddert, L. E., Stoker, M., Torres, O. & Oerther, D. B. 2018c **Association of aflatoxin exposure and height-for-age among young children in Guatemala**. *International Journal of Environmental Health Research* **28** (3), 280–292. <https://doi.org/10.1080/09603123.2018.1468424>.
- Voth-Gaeddert, L. E., Stoker, M., Torres, O. R. & Oerther, D. B. 2019 **The influence of local market and household factors on aflatoxin presence in maize and symptoms of its exposure to children in Guatemala**. *International Journal of Environmental Health Research* 1–15. <https://doi.org/10.1080/09603123.2019.1594721>.
- Wagner, E. G. & Lanoix, J. N. 1958 *Excreta Disposal for Rural Areas and Small Communities*. Monograph Series, No. 39. World Health Organization, Geneva, Switzerland.
- Wolf, J., Johnston, R., Hunter, P. R., Gordon, B., Medlicott, K. & Prüss-Ustün, A. 2019 **A faecal contamination index for interpreting heterogeneous diarrhoea impacts of water, sanitation and hygiene interventions and overall, regional and country estimates of community sanitation coverage with a focus on low- and middle-income countries**. *International Journal of Hygiene and Environmental Health* **222** (2), 270–282. <https://doi.org/10.1016/j.ijheh.2018.11.005>.

First received 17 June 2024; accepted in revised form 29 August 2024. Available online 9 September 2024