



Development and Validation of a Community-Led Total Sanitation (CLTS) Evaluation Instrument for Diarrhea Intervention

Zaenal Abidin^{1,2}, Adelia Kusuma Prasetyo Putri²

¹ Medical Faculty, Universitas Negeri Semarang, Central Java. Indonesia.

² STIKes Bhakti Husada Mulia, Madiun, East Java. Indonesia.

ARTICLE HISTORY

Received May 24, 2025
Accepted July 20, 2025
Published August 31, 2025

Keyword:

Community-Led Total Sanitation (CLTS); Diarrhea; Instrument Development; Validity; Reliability

Kata kunci:

Sanitasi Total Berbasis Komunitas (CLTS); Diare; Pengembangan Instrumen; Validitas; Reliabilitas



Ruwa Jurai: Jurnal Kesehatan Lingkungan is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

ABSTRACT

Community-Led Total Sanitation (CLTS) is a crucial methodology for preventing diarrheal diseases; however, the availability of valid and reliable assessment tools remains inadequate. The study aims to develop and evaluate the validity and reliability of an instrument for assessing CLTS initiatives. This study employed a mixed-methods framework consisting of two distinct phases: (1) expert validation conducted by five environmental health specialists to assess content validity (*Content Validity Index/CVI*), and (2) reliability assessment involving 30 participants using *Cronbach's Alpha*. The content validity (Scale-Level CVI/S-CVI) for all constructs reached the threshold of ≥ 0.85 , although two items were considered invalid ($I-CVI < 0.78$). The reliability coefficients derived from *Cronbach's Alpha* ranged from 0.61 to 0.79, with three items identified as inconsistent. This instrument is suitable for implementation following the revision of items exhibiting ambiguity. The practical implications highlight the need for enhanced precision in monitoring CLTS programs.

Sanitasi Total Berbasis Komunitas (CLTS) merupakan metodologi penting untuk mencegah penyakit diare; namun, ketersediaan alat penilaian yang valid dan reliabel masih kurang memadai. Studi ini bertujuan untuk mengembangkan dan mengevaluasi validitas dan reliabilitas instrumen untuk menilai inisiatif CLTS. Studi ini menggunakan kerangka kerja metode campuran yang terdiri dari dua fase berbeda: (1) validasi ahli yang dilakukan oleh lima spesialis kesehatan lingkungan untuk menilai validitas isi (Indeks Validitas Isi/CVI), dan (2) penilaian reliabilitas yang melibatkan 30 partisipan menggunakan Alpha Cronbach. Validitas isi (CVI Tingkat Skala/S-CVI) untuk semua konstruk mencapai ambang batas $\geq 0,85$, meskipun dua item dianggap tidak valid ($I-CVI < 0,78$). Koefisien reliabilitas yang diperoleh dari Alpha Cronbach berkisar antara 0,61 hingga 0,79, dengan tiga item diidentifikasi sebagai tidak konsisten. Instrumen ini cocok untuk implementasi setelah revisi item yang menunjukkan ambiguitas. Implikasi praktisnya menyoroti perlunya peningkatan ketelitian dalam memantau program CLTS.

* Corresponding author: Zaenal Abidin
Medical Faculty, Universitas Negeri Semarang, Central Java, Indonesia
Email: zapb17@students.unnes.ac.id

1. Introduction

Diarrhea remains a leading cause of global morbidity and mortality, particularly in low- and middle-income countries. As reported by the World Health Organization, approximately 1.6 million deaths each year are attributable to

diarrhea, with a significant proportion of cases linked to inadequate sanitation, unsafe drinking water, and poor hygiene practices (WHO, 2023). In Indonesia, diarrhea continues to rank among the top ten diseases with the highest incidence, especially among young children (Badan Kebijakan Kesehatan Nasional, 2023).

Community-Led Total Sanitation (CLTS) interventions have been recognised as a practical approach to improving sanitation conditions and reducing diarrhoea incidence. CLTS comprises five fundamental pillars: Stop Open Defecation (ODF), Handwashing with Soap (HWWS), Safe Drinking Water and Food Management, Solid Waste Management, and Household Wastewater Management. CLTS has been adopted as a national strategy and prioritised as an intervention to reduce diarrhoea prevalence through community empowerment (Kementerian Kesehatan RI, 2022). However, the success of CLTS interventions depends heavily on the ability to conduct accurate, systematic evaluations of sanitation conditions at the community level. Existing sanitation evaluation tools often do not fully reflect local characteristics and comprehensive community involvement.

Existing Community-Led Total Sanitation (CLTS) assessment tools often lack relevance to local contexts and fail to meet stringent validity and reliability standards, rendering their results susceptible to bias and inaccuracy and thereby impeding resource allocation and program effectiveness (Folk & Smith, 2002; Ingram et al., 2022). Studies indicate that instruments with insufficient validity and reliability hinder accurate evaluation of sanitation intervention effectiveness (Sinharoy et al., 2022). Therefore, the development of a CLTS evaluation tool encompassing the five pillars, incorporating expert validation and reliability testing, is essential to support effective and sustainable diarrhea interventions. This study assesses not only content validity but also the instrument's internal consistency, addressing previous methodological gaps (Polit & Beck, 2010). A valid and reliable instrument is expected to enhance CLTS program implementation and significantly reduce the incidence of diarrhea among children under five.

2. Methods

1.1. Instrument

The instrument was self-developed with reference to the Five Pillars of CLTS, comprising five constructs/indicators for diarrhea prevention interventions: ODF, HWWS, Safe Drinking Water and Food Management, Solid Waste Management, and Household Wastewater Management, totaling

36 items. Scoring was based on a 4-point Likert scale using favorable statements, where "Always = 4, Often = 3, Sometimes = 2, and Never = 1" and for unfavourable statements, the scoring was reversed, with "Always = 1, Often = 2, Sometimes = 3, and Never = 4".

1.2. Validity Content from Expert

A panel of five experts (n=5) with a minimum background in environmental health as associate environmental health experts, including environmental health specialists with doctoral degrees as educators, professors of environmental science education from universities, and three managers/implementers of environmental health programmes at public health services with diploma qualifications. Item relevance was assessed using a 4-point scale, and the Content Validity Index (CVI) was calculated according to the criterion of I-CVI \geq 0.78 (Polit & Beck, 2021). The Item-Level CVI (I-CVI) is determined by counting the number of experts who assign a score of 3 or 4 to each item, using the following formula:

$$I-CVI = \frac{\text{Number of experts who gave a score of 3/4}}{\text{Total Experts}}$$

The validity criterion for I-CVI is \geq 0.78 (for five experts: at least four experts must agree), while items with an I-CVI $<$ 0.78 are subject to revision or removal. The Scale-Level CVI (S-CVI) is calculated as the average I-CVI of all items within a construct, with an S-CVI \geq 0.9 indicating high content validity.

1.3. Reliability Test

The instrument was pilot-tested on 30 respondents (n=30), consisting of heads of household, mothers, or caregivers in the rural areas of Kebonsari Subdistrict, Madiun Regency. This was conducted because assessing the validity and reliability of the developed tool prior to its widespread use in the main study is a crucial strategic step in developing a community-based total sanitation evaluation instrument. Cronbach's Alpha was used to assess internal consistency, with $\alpha \geq$ 0.7 considered indicative of good reliability (Taber, 2018). Data analysis was carried out using computer software (SPSS) to calculate the Corrected Item-Total Correlation, Cronbach's

Alpha, and descriptive statistics (mean, variance) to evaluate the distribution of responses.

3. Results

3.1. Content Validity

The *Content Validity Index* (CVI) is a quantitative method used to assess the extent to which items in an instrument are considered relevant and representative of the construct being measured,

based on expert panel evaluations. In the past two years, several international studies have emphasized that CVI remains the gold standard for instrument development and validation, ensuring content validity objectively and systematically (Hair et al., 2022; Polit & Beck, 2010). The widespread use of CVI is recognized for its ability to provide a precise measure of expert agreement regarding the relevance of each item in the developed instrument.

Table 1. Item-Level CVI (I-CVI) in CLTS Instruments Based on Panel Expert

Sec	Number	Statement/Declaration	1	2	3	4	5	I-CVI
I		Stop Open Defecation (ODF)						
	5	<i>Toilets/latrines are dirty and smelly (Unfavourable)</i>	4	3	4	2	4	0,8
	9	<i>Toilet/latrine where insects live</i>	3	4	1	3	2	0,6
II		Handwashing with Soap (HWWS)						
	4	<i>Cloth/Tissue at the hand washing station</i>	4	3	3	1	2	0,6
IV		Safe Management of Household Solid Waste						
	6	<i>Regular management of waste disposal</i>	3	2	3	1	3	0,6

Based on the assessment of five experts as presented in Table 1, the CLTS evaluation instrument demonstrated adequate content validity, with an average Scale-Level CVI (S-CVI) of 0.85 across all constructs (Five pillars of CLTS). This value exceeds the minimum threshold of 0.78 (Polit

& Beck, 2010), indicating that the instrument is generally relevant for measuring community-based sanitation practices. However, there were four items (11%) that were deemed invalid (I-CVI < 0.78), as shown in Table 2 below:

Table 2. Scale-Level CVI (S-CVI) on CLTS Instruments

Construct/Pilar CBTS	Σ Item	S-CVI
Stop Open Defecation (ODF)	10	0,85
Handwashing with Soap (HWWS)	9	0,88
Safe Food and Drinking Water Management	6	0,92
Safe Solid Waste Management	6	0,89
Household Wastewater Management	5	0,91

3.2. Reliability

Reliability is a key indicator of measurement instrument quality, reflecting the degree of consistency in results when the instrument is administered repeatedly under the same

conditions. Instruments with high reliability produce stable, trustworthy data, making reliability essential for accurate research findings (Hair et al., 2022; Polit & Beck, 2010).

Table 3. Cronbach's Alpha Value on CLST Pillar Instrument Items

Construct/Pilar CLTS	Σ Item	α	Unreliable Item ($r < 0,3$)
Stop Open Defecation (ODF)	10	0,75	Item 5, 9
Handwashing with Soap (HWWS)	9	0,68	Item 4
Safe Food and Drinking Water Management	6	0,72	-
Safe Solid Waste Management	6	0,61	Item 6 c
Household Wastewater Management	5	0,79	-

The results of the reliability test using Cronbach's Alpha, as presented in Table 2, indicate high values ($\alpha > 0.7$) for ODF, Safe Management of Household Food and Drinking Water, and Safe Management of Household Liquid Waste. In contrast, relatively lower values ($\alpha < 0.7$) were observed for HWWS and Safe Management of Household Solid Waste. This suggests that the complexity of handwashing and waste management practices is difficult to capture accurately with a Likert scale.

4. Discussion

Community-Led Total Sanitation (CLTS) is a key strategy for diarrhea prevention, particularly in developing countries such as Indonesia, where the incidence of diarrheal diseases remains high (Kementerian Kesehatan RI, 2022; WHO, 2023). Practical evaluation of CLTS programs requires scientifically validated instruments to support accurate decision-making (Benham & Hawley, 2014; Stuart et al., 2021). Instruments that lack validity or reliability can introduce bias into evaluation results, lead to incorrect identification of intervention needs, and ultimately result in program failure. Valid and reliable evaluation instruments are essential for assessing intervention effectiveness, identifying implementation barriers, and supporting evidence-based decision-making (Benham & Hawley, 2014; Panulo et al., 2024). However, many instruments currently in use remain generic and fail to adequately account for local context and sociocultural dynamics (Elwood & Murphy, 2015).

Content validity, assessed using the Content Validity Index (CVI) by a panel of five experts, yielded an average Scale-Level CVI (S-CVI) of 0.85 across all CLTS constructs, surpassing the minimum acceptable value of 0.78 (Zakianis et al., 2020). This high S-CVI score underscores the importance of expert assessment in developing sanitation evaluation tools, indicating that the instrument is generally relevant for measuring community sanitation practices. Nevertheless, four items (11%) had I-CVI scores below 0.78, including "Toilet/latrine where insects live" and "Cloth/Tissue at the hand washing station", highlighting ambiguities that need to be addressed. Recent studies emphasise the need to involve multidisciplinary expert panels, including public

health specialists, anthropologists, and local practitioners, to evaluate the relevance and clarity of items (Spoto et al., 2025).

The qualitative methods, such as cognitive interviewing, prior to quantitative testing is effective in identifying ambiguities and improving item wording (Grand-Guillaume-Perrenoud et al., 2023). This supports earlier findings by Nawab et al., (2006) regarding how cultural variations and unclear phrasing can reduce item validity. Therefore, revising ambiguous terms, such as replacing "cloth/tissue" with "clean tissue/towel," is essential to enhance clarity.

The application of cognitive interviewing to assess respondent comprehension (Hair et al., 2022) and the inclusion of health anthropologists in the expert panel are also recommended, particularly for evaluating the contextual relevance of terms such as "cloth/tissue." The use of a glossary is equally important to ensure consistent interpretation among respondents (Morris, 2018), especially for terms such as "garbage thrown carelessly," which can be clarified as "organic/inorganic waste." Ultimately, content validity is a fundamental aspect of instrument development, ensuring that each item accurately represents the intended construct (Polit & Beck, 2010). Recent studies further emphasize the need for multidisciplinary expert panels, including public health specialists, anthropologists, and local practitioners, to assess the relevance and clarity of items (Nawab et al., 2006). Moreover, employing qualitative methods, such as cognitive interviewing, prior to quantitative testing has proven effective for identifying ambiguities and refining item wording (Grand-Guillaume-Perrenoud et al., 2023; Zamanzadeh et al., 2015).

The reliability analysis revealed a range of Cronbach's Alpha values across the CLTS constructs, from 0.61 (Solid Waste Management) to 0.79 (Household Wastewater Management). This analysis indicates that specific constructs, particularly Solid Waste Management, exhibit suboptimal internal consistency. This is evident in the item "garbage thrown carelessly" (Item 6c), which was less effective in measurement because prevailing social norms consider such practices unacceptable. This finding is consistent with the literature, as waste management measurements often face stability challenges due to the

complexity of field practices and the influence of social norms. Additionally, some items were found to be insufficiently specific, such as "Toilet/latrine where insects live" (Item 9), which may have led to diverse interpretations among respondents (Zamanzadeh et al., 2015). According to Hair et al., (2022), the a Cronbach's Alpha value below 0.7 indicates the need to revise items or add new items to enhance reliability. In this context, three items (5, 9, and 6c) with item-total correlations below 0.3 require revision or elimination. For a more comprehensive reliability assessment, particularly for complex constructs such as Solid Waste Management, the Rasch Model approach may be considered as an alternative (Taber, 2018).

Reliability, particularly internal consistency as measured by Cronbach's Alpha, is a crucial indicator for assessing the stability of an instrument. Contemporary research indicates that while an ideal Cronbach's Alpha value is above 0.7, for newly developed instruments or complex constructs, values between 0.6 and 0.7 may still be acceptable provided that additional reliability tests, such as Composite Reliability and Rasch Analysis, are conducted (Taber, 2018). To further enhance reliability, especially in aspects influenced by social norms such as household solid waste management, refinement or addition of items with low consistency is necessary (Laurent-Vannier, 2022; Zamanzadeh et al., 2015).

The evaluation instrument for CLTS in diarrhea interventions demonstrates feasibility for use following the revision of ambiguous and unreliable items. However, several limitations require attention to improve the instrument's quality and applicability. These limitations encompass four main aspects: a limited sample size ($n=30$), restricted geographic coverage to a single area, ambiguity in the terminology used, low reliability across several constructs, and a constrained approach to expert validation.

For the future advancement of the instrument, it is essential to consider several critical factors, including the local contextual framework, community sociocultural dynamics, and the systematic implementation of validity and reliability assessments (Hair et al., 2022; Tiberghien et al., 2011). For future advancement of the instrument, it is essential to consider several critical factors, including the local contextual framework,

community sociocultural dynamics, as well as the systematic implementation of validity and reliability assessments (Hair et al., 2022; Zamanzadeh et al., 2015).

5. Conclusions

The development and validation of the CLTS evaluation instrument for diarrhea intervention yielded several important findings. Firstly, the instrument exhibited acceptable validity, with an S-CVI score of 0.85, surpassing the minimum threshold of 0.78. Nonetheless, certain items need refinement because their I-CVI values fall below 0.78, especially those involving ambiguous terminology. Secondly, the reliability analysis revealed variability in Cronbach's Alpha values ranging from 0.61 to 0.79 across the five CLTS pillars. Three constructs of Open Defecation Free (ODF), Safe Drinking Water and Food Management, and Household Wastewater Management, demonstrated satisfactory reliability with alpha values above 0.7. Conversely, the remaining two constructs, Handwashing with Soap (HWWS) and Solid Waste Management, showed lower reliability and thus require further improvement. Lastly, the instrument is suitable for evaluating CLTS programs, provided that problematic items, particularly those related to waste and wastewater management, are revised accordingly.

These refinements are expected to enhance the accuracy of program evaluation and the effectiveness of diarrhea interventions. Future development should include pilot testing with larger sample sizes and broader geographic coverage to improve the instrument's generalizability. The inclusion of experts from various disciplines in the expert panel is also recommended to enhance the instrument's contextual validity.

References

- Badan Kebijakan Kesehatan Nasional. (2023). Laporan Survei Kesehatan Indonesia Dalam Angka Tahun 2023. In *Kemendes RI* (Ed 1). Kemendes. https://drive.google.com/file/d/1rjNDG_f8xG6-Y9wmhJUnXhJ-vUFevVJC/view
- Benham, B., & Hawley, D. (2014). The effectiveness of tools used to evaluate successful critical decision

- making skills for applicants to healthcare graduate educational programs: a systematic review protocol. *JBI Database of Systematic Reviews and Implementation Reports*, 12(3), 39–50. <https://doi.org/10.11124/jbisrir-2014-1392>
- Elwood, J., & Murphy, P. (2015). Assessment systems as cultural scripts: a sociocultural theoretical lens on assessment practice and products. *Assessment in Education: Principles, Policy & Practice*, 22(2), 182–192. <https://doi.org/10.1080/0969594X.2015.1021568>
- Folk, V. G., & Smith, R. L. (2002). *Models for Delivery of CBTs* (1st Ed). Routledge. <https://doi.org/10.4324/9781410612250-3>
- Grand-Guillaume-Perrenoud, J. A., Geese, F., Uhlmann, K., Blasimann, A., Wagner, F. L., Neubauer, F. B., Huwendiek, S., Hahn, S., & Schmitt, K.-U. (2023). Mixed methods instrument validation: Evaluation procedures for practitioners developed from the validation of the Swiss Instrument for Evaluating Interprofessional Collaboration. *BMC Health Services Research*, 23(1), 83. <https://doi.org/10.1186/s12913-023-09040-3>
- Hair, J. F., Babin, B. J., & Anderson, R. E. (2022). *Multivariate Data Analysis*. Cengage Learning.
- Ingram, C., Chen, Y., Buggy, C., Downey, V., Archibald, M., Rachwal, N., Roe, M., Drummond, A., & Perrotta, C. (2022). Development and validation of a multi-lingual online questionnaire for surveying the COVID-19 prevention and control measures used in global workplaces. *BMC Public Health*, 22(1), 74. <https://doi.org/10.1186/s12889-022-12500-w>
- Kementerian Kesehatan RI. (2022). *Strategi Nasional STBM 2022-2026*.
- Laurent-Vannier, A. (2022). Shaken Baby Syndrome (SBS) or Pediatric Abusive Head Trauma from Shaking: Guidelines for Interventions During the Perinatal Period from the French National College Of Midwives. *Journal of Midwifery & Women's Health*, 67(S1). <https://doi.org/10.1111/jmwh.13427>
- Morris, K. A. (2018). Measurement equivalence: a glossary for comparative population health research. *Journal of Epidemiology and Community Health*, 72(7), 559–563. <https://doi.org/10.1136/jech-2017-209962>
- Nawab, B., Nyborg, I. L. P., Esser, K. B., & Jenssen, P. D. (2006). Cultural preferences in designing ecological sanitation systems in North West Frontier Province, Pakistan. *Journal of Environmental Psychology*, 26(3), 236–246. <https://doi.org/10.1016/j.jenvp.2006.07.005>
- Panulo, M., Chidziwisano, K., MacLeod, C., Kapazga, T., Dreibelbis, R., Beattie, T. K., & Morse, T. (2024). *CLTS implementation in Malawi: process evaluation of a sanitation and hygiene intervention*. <https://doi.org/10.31219/osf.io/9bfqd>
- Polit, D. F., & Beck, C. T. (2010). Essentials of Nursing Research: Appraising Evidence for Nursing Practice. In *Lippincott Williams & Wilkins* (Tenth, Nor, Vol. 95, Issue 2). Lippincott Williams & Wilkins.
- Sinharoy, S. S., Conrad, A., Patrick, M., McManus, S., & Caruso, B. A. (2022). Protocol for development and validation of instruments to measure women's empowerment in urban sanitation across countries in South Asia and Sub-Saharan Africa: the Agency, Resources and Institutional Structures for Sanitation-related Empowerment (ARISE). *BMJ Open*, 12(2), e053104. <https://doi.org/10.1136/bmjopen-2021-053104>
- Spoto, A., Nucci, M., Prunetti, E., & Vicovaro, M. (2025). Improving content validity evaluation of assessment instruments through formal content validity analysis. *Psychological Methods*, 30(2), 203–222. <https://doi.org/10.1037/met0000545>
- Stuart, K., Peletz, R., Albert, J., Khush, R., & Delaire, C. (2021). Where Does CLTS Work Best? Quantifying Predictors of CLTS Performance in Four Countries. *Environmental Science & Technology*, 55(6), 4064–4076. <https://doi.org/10.1021/acs.est.0c05733>
- Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Tiberghien, J. E., Robbins, P. T., & Tyrrel, S. F. (2011). Reflexive assessment of practical and holistic sanitation development tools using the rural and peri-urban case of Mexico. *Journal of Environmental Management*, 92(3), 457–471. <https://doi.org/10.1016/j.jenvman.2010.08.032>
- WHO. (2023). *Diarrhea Disease*. <https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease>
- Zakianis, Sabarinah, Djaja, I. M., & Agustina, H. (2020). Validity and Reliability of Instruments Assessing Household Satisfaction With 3R-Based Waste Management (Tempat Pengolahan Sampah) and Waste Bank Services in Depok, Indonesia. *European Journal of Molecular and Clinical Medicine*, 7(10), 2347–2360.

<https://www.ejmcm.com/archives/volume-7/issue-10/9566>

Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A.-R. (2015). Design and Implementation Content Validity Study: Development of an instrument for

measuring Patient-Centered Communication. *Journal of Caring Sciences*, 4(2), 165–178. <https://doi.org/10.15171/jcs.2015.017>