HOUSEHOLD WATER INSECURITY EXPERIENCES (HWISE) SCALE: THE PROTOCOL OF CULTURAL ADAPTATION AND STATISTICAL VALIDATION

Navodya C. Selvaratnam[1i],
Naren D. Selvaratnam[2],
A.M.N.A.D.J.S. Nanayakkara[1],
S. Tennakoon[1]
[1]Department of Community Medicine,
Faculty of Medicine,
University of Peradeniya,
Sri Lanka
[2]School of Psychology,
Sri Lanka Institute of Information Technology (SLIIT),
Malabe, Sri Lanka

Abstract:

Annually Sri Lanka experiences substantial rainfall; but, the availability of water within the island nation is unbalanced. Lack of an appropriate system view, suitable national-level policy, and bimodal rainfall have resulted in Sri Lankans experiencing medium to high water stress. Considering the impact individual households have to endure due to the unavailability of water, the following study presents a protocol to culturally adapt and statistically validate the Household Water Insecurity Experiences (HWISE) scale to the Sinhala language. The protocol critically discusses the application of classical test theory (CTT) and item response theory (IRT) and documents all relevant steps for the scale validation process. Through this study protocol, the authors attempt to demonstrate the procedural rigor of the adaptation and validation process of the HWISE scale to the Sinhala language in Sri Lanka. The availability of a valid scale to evaluate household water insecurity experiences will assist public health experts, policymakers, and the government in understanding the nature and severity of water stress and the consequences across health, economic, and psychosocial contexts. Similarly, preventative interventions and key policies could be implemented to further support the initiatives of responsible authorities.

Keywords: household water insecurity, classical test theory, item response theory, psychometrics, graded response model

[i] Correspondence: email navoselva@gmail.com

1. Introduction

Water is vital for the survival of all beings and crucial for human health and well-being. It is fundamental for the sustenance of agriculture, industry, economic productivity, and political stability of a country (Kangmennaang & Elliott, 2023; United Nations (UN), 2023). Water may also hold cultural significance and symbolic meanings in some communities (Jepson *et al.*, 2017; Popkin *et al.*, 2010; Rosinger and Young, 2020). Despite its importance, between 2 billion to 3 billion people around the globe annually experience shortages of water for at least one month (UNESCO, 2023). It is estimated that 2 billion people around the world lack access to safe drinking water while 3.6 billion lack access to safe sanitation (UNESCO, 2023). Shortages persist as only 3% of Earth's surface water is available as freshwater; and about 30% of the available freshwater is appropriate for human use (Hannah *et al.*, 2020). Since the water sources are not distributed equally across the globe, at present, the Middle East, North Africa, and South Asia experience the highest water stress (Kuzma *et al.*, 2023). This may be exacerbated as a result of the growing population, burgeoning demand, climate change, floods, droughts, surge in agricultural interventions, and degradation of water infrastructure (Abbott *et al.*, 2019; Kangmennaang & Elliott, 2021; Miller *et al.*, 2020).

In the region of South Asia, in comparison to some other countries, the island country Sri Lanka is not widely known to experience water scarcity. However, Sri Lanka experiences an unbalanced water availability with water deficits and surfeits across its climatic zones (Figure 1).
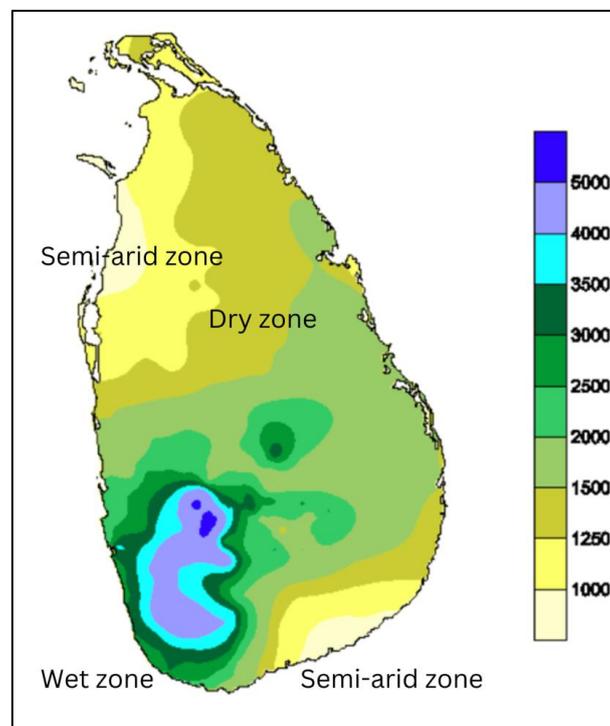

Figure 1: Annual rainfall of Sri Lanka in millimeters
(Department of Meteorology, 2019)

This temporal and spatial variability of water occurs primarily due to the bimodal nature of the country's annual rainfall (Chandrasekara *et al.*, 2021; Gunatilaka, 2008). Additionally, about 50% of the country does not occupy sufficient amounts of groundwater. Sri Lanka's higher population density (353/km^2) with an approximate population of 22.156 million may further add to the identified insufficiency of consumable water. However, the island nation predominantly houses a rural population where only 19% of its populace resides in urban settings (Sri Lanka Department of Census and Statistics, 2021; World Bank, 2022). In Sri Lanka, the majority of water precipitation is through rainfall. The driest parts (dry zone) in the country receive an annual mean rainfall of under 900mm, while the wettest parts (wet zone) receive over 5000mm (Department of Meteorology, 2019; Gunatilaka, 2008). In 2022, only 89% of the total population had at least basic drinking water services, while only 85.35% of the total population had basic handwashing facilities including soap and water (Word Bank, 2022).

Currently, Sri Lanka has a level three or medium to high (20% - 40%) water stress from the six-level indicator for baseline water stress (Kuzma *et al.*, 2023). Calculating water stress (water demand to renewable supply ratio) is particularly important as it aids in measuring the competition in a country for its local water resources. Gunatilaka (2008) also highlighted that the limited groundwater available in Sri Lanka is at risk of overexploitation. Potential driving factors may include the use of well water for daily requirements by a majority of the rural population and the upward trend of groundwater extraction for agricultural purposes. The risk of overexploitation may be aggravated due to no-fee access to irrigation water, double cropping of rice, and high wastage of water (Gunatilaka, 2008). It is important to note that 85% of the country's water is used by the irrigation sector, of which, 70% is used for paddy cultivation. During the primary cultivation season, which spans from September to March, paddy cultivation is practiced at 100%. During the May to August cultivation season, paddy cultivation is practiced again in combination with other crops (Chandrasekara *et al.*, 2021). Besides, Gunatilaka (2008) predicted that the top four highest-populated districts (Colombo, Gampaha, Kurunegala, Kandy), Jaffna, and Puttalam districts would be at the highest risk of depletion of water resources.

1.1 Significance of water insecurity

As Jepson *et al.* (2017) and Young *et al.* (2019b) identify, the condition of water insecurity occurs when the affordability, reliability, adequacy, and safety of water are diminished significantly or become unattainable to potentially threaten the physical-mental well-being of individuals with hindrances or impediments to engagement in social, cultural, and productive activities. In 2022, at least 1.7 billion people were using a drinking water source contaminated with feces globally (World Health Organization (WHO), 2023). When the available water is of poor quality or contaminated microbiologically, the risk of transmission of diarrheal diseases, dysentery, cholera, typhoid, and polio could increase (WHO, 2023). Besides, globally, diarrheal diseases are the second largest

contributor to the mortality of children under 5 years of age (WHO, 2017). Contaminated water, though not microbiologically, may increase exposure to high levels of inorganic arsenic, potentially leading to chronic arsenic poisoning (WHO, 2022). The potential impacts may encompass cancer and skin lesions. If exposed in utero and early childhood, the cognitive development of the child could be affected negatively and could increase deaths in young adults (WHO, 2022).

Similar to having safe water, accessing safe water is equally important. When the available water is further away, water fetching may take extended periods resulting in a loss of time to engage in income-generation activities for families and to attend school for children. Women and children are the most affected due to water scarcity because they tend to hold the responsibility of water fetching (UNICEF, 2020). UNICEF (2020) further identified that girls tend to be at a greater disadvantage due to water shortages, especially in schools. As water fetching limits income generation activities, the purchasing power of households may also diminish. As a result, families may be forced to choose between purchasing water, food, effective ways to irrigate crops, and raising livestock (Hannah *et al.*, 2020). Due to water scarcity, individuals may not be able to engage in social, cultural, and productive activities and fulfill duties. They may also experience feelings of embarrassment and social shame. These hindrances and negative experiences may potentially lead to elevated levels of frustration, anxiety, and depression in individuals (Kimutai *et al.*, 2023). In extreme situations, water insufficiency may elicit rumination and suicidal ideation (Kangmennaang & Elliott, 2021). As Kangmennaang & Elliott (2021) further highlight, the distress due to water scarcity may be more pronounced in women. This may be underpinned by the culturally constructed gender roles that dictate women in the household to engage more in negotiating access to safe water. As a result, it may expose women to harsh weather conditions, quarrels, and accidents, and may also subject them to an increased risk of assault and rape en route to the water source (Kimutai *et al.*, 2023).

1.2 Assessing water insecurity - The Household Water Insecurity Experiences Scale

The experiences of households concerning water scarcity can be measured through the scale, Household Water Insecurity Experiences (HWISE). It is the first scale to equivalently quantify household water insecurity experiences across both middle-income countries and low-income countries. The HWISE scale can be utilized in different geographic regions, water-provisioning settings, and cultures as it can comparably measure key universal household water insecurity experiences despite any contextual diversity. The scale was originally developed by the HWISE Research Coordination Network (RCN) in the English language. Currently, the HWISE scale is available in more than fifty languages. A four-item shortened version of the HWISE scale is also available (Northwestern University Institute for Policy Research, 2019a; Young *et al.*, 2019a).

The scale examines water insecurity experiences of households within the past four weeks across four components, that is, *accessibility, adequacy, reliability,* and *safety*. It requires generally around four minutes to administer the scale. The HWISE scale is

comprised of simple worded questions under the twelve labels, *(1) Worry, (2) Interrupt, (3) Clothes, (4) Plans, (5) Food, (6) Hands, (7) Body, (8) Drink, (9) Angry, (10) Sleep, (11) None,* and *(12) Shame*. The responses to each item include *if 0 times – 'never', if 1 to 2 times – 'rarely', if 3 to 10 times – 'sometimes', if 11 to 20 times – 'often',* and *if more than 20 times – 'always'*. The responses will be coded as *'never'* – 0, *'rarely'* – 1, *'sometimes'* – 2, and *'often/always'* – 3. The coded values for all 12 items will be summed to obtain the scale score for each household, which may range from 0 to 36. A household with a scale score of 12 or above will be considered water-insecure as it is recognized as the provisional cut-point for a water-insecure household (Young *et al.*, 2019a; Young *et al.*, 2019b). The HWISE scale can be administered to individuals aged 16 or 18 years and above who identify themselves to be knowledgeable about water acquisition and use in their households. The scale will be administered in the language preferred by the respondent, subjected to the availability of a culturally adapted and statistically validated HWISE scale (Young *et al.*, 2019a; Young *et al.*, 2019b). As per the findings of Young *et al.* (2019a), the HWISE scale is highly reliable with internal consistency reliability ranging from Cronbach's α .84 to α .93. Despite the scale's consideration of four unique components in the development, it is intended to be used as a unidimensional scale.

The HWISE scale allows both the identification of determinants of water insecurity and the assessment of consequences of household water insecurity across health, economic, and psychosocial contexts, including food insecurity. It further facilitates the identification of vulnerable subpopulations within communities. The HWISE scale can be adopted for large-scale monitoring and evaluation due to the relative easiness of administering. The data generated with the use of the HWISE scale can be utilized in selecting water-related programs, implementing technologies and policies, and evaluating the impact and cost-effectiveness of interventions. Furthermore, the scale would facilitate identifying and monitoring any trends between climate variation, political changes, economic changes, social changes, and water insecurity (HWISE Research Coordination Network, 2019).

1.3 Present study

The HWISE scale, though used in numerous settings, has not been culturally adapted to Sri Lanka. Thus, the present study attempts to culturally adapt the HWISE scale to Sri Lanka by translating it into the Sinhala language and analyzing the psychometric properties of the translation. The scale will be tested for its psychometric properties through both classical test theory (CTT) and item response theory (IRT). Specific statistical tests under CTT include reliability testing and validity testing. Validity for the scale will be established through predictive validity, convergent validity, discriminant validity, and internal structural validity. IRT includes the use of a graded response model (GRM) to assess individual items.

1.4 Significance of the study

Translation of the HWISE scale into the Sinhala language will facilitate the identification of water insecurity experiences of households in Sri Lanka, including its determinants and consequences across health, economic, and psychosocial contexts. It will further assist researchers, public health experts, policymakers, and the government in identifying vulnerable subpopulations within the Sinhala-speaking communities, potential effective interventions to improve household water insecurity, impact and cost-effectiveness of interventions, and mitigate negative consequences. Similarly, the mainstream use of IRT analyses for scale validation in Sri Lanka is scarce. Accordingly, this study will also function as a step-by-step guide for researchers intending to learn both CTT and IRT for scale adaptation and validation.

2. Methods

2.1 Study design

The present study aims to culturally adapt and statistically validate the HWISE scale to the Sinhala-speaking community in Sri Lanka by (1) *Translating the HWISE scale into the Sinhala language* and (2) *Analyzing the psychometric properties of the translation.* The process of cross-cultural adaptation is guided by Gronier (2022) and Scholz *et al.* (2002).

2.2 Translation of scale
2.2.1 Preparation of Version 1

The permission to translate the original HWISE scale into the Sinhala language was obtained from Dr. Sera L. Young on behalf of the developers of the HWISE scale. Since permission was granted, the original English version of the HWISE scale will be translated into the Sinhala language (forward translation) by the researchers N.C.S and J.S.N independently. Both researchers are native speakers of Sinhala and are fluent in both Sinhala and English languages. The translation will be performed to reach conceptual equivalence between the original English version and the Sinhala translation of the HWISE scale. Once the forward translation is completed, the research team will examine the two Sinhala translations for any discrepancies. Any discrepancies will be resolved by discussion and consensus-based decisions. The initial Sinhala translation of the HWISE scale will be finalized and will be named *Sinhala HWISE Version 1*.

2.2.2 Preparation of Version 2

The *Sinhala HWISE Version 1* will be translated back into the English language (back-translation) by two translators fluent in both Sinhala and English languages independently. The translators will not be familiar with the original HWISE scale or the concepts involved with the scale. Back-translation will aim to assess whether the contents of the items in the Sinhala translation reflect the original English version. Once the back-translation is completed, the research team will examine the two back-translations against the original English version. Back translation is especially important to meet

requirements for content validation. In all scale adaptation processes, finding the ideal equivalent word for an English word in the destination language is not possible (i.e., words such as 'coping,' 'stigma,' etc. are often difficult to translate accurately). In such instances, the next most appropriate word will be chosen. The same steps may apply for certain phrases (i.e., 'how frequently') of the present scale. However, any required amendments to *Sinhala HWISE Version 1* will be made through discussion and consensus-based decisions. The resulting Sinhala translation of the HWISE scale will be named *Sinhala HWISE Version 2.*

2.2.3 Preparation of Version 3

A panel of experts in the areas of public health, psychometrics, linguistics, cross-cultural adaptation of scales, and community-based research will be assembled. The experts in the panel will independently review each item of the *Sinhala HWISE Version 2* against the corresponding item of the original English scale. Experts will rate each item of the *Sinhala HWISE Version 2* on a 1 – 10 Likert scale to facilitate the assessment of content and consensual validity of all items. To enable this step in the scale adaptation, a content validation technique named Delphi will be used (de Zoysa *et al.*, 2007). In the Delphi process, each item (Table 1) will be assessed by the chosen group of experts under five criteria to evaluate content and consensual validity. If more than 70% of the aggregate scores of all experts for any of the criteria for a given item fall below 3, the items will be subjected to revisions. Changes to the Sinhala *HWISE Version 2* at this stage will be discussed in a panel discussion and *Sinhala HWISE Version 3* will be finalized through consensus-based decisions.

2.3 Pre-test of Sinhala HWISE Version 3

A pre-test will be conducted to assess the appropriateness of *Sinhala HWISE Version 3* in facilitating the investigation of household water insecurity experiences as originally intended by the English HWISE scale. The pre-test will further assist authors in assessing the difficulties research participants may encounter while completing the 12-item scale. Common problems may include incomprehensible words which could result due to limited context items may provide to some individuals. In such instances, the researchers will seek the assistance of the expert panel and a consensus-based decision will be taken to fine-tune the scale further.

Table 1: Delphi process to establish content validity

Items in HWISE	Content-related validation		Consensual-related validation		
	Appropriateness of language used	Assessment of the concept	Retains the conceptual meaning	Appropriateness with individuals of 18 years and above	Cultural relevance
1. Item 1	(1-10)	(1-10)	(1-10)	(1-10)	(1-10)

Note: Only 1 item is displayed in the above example. In the original validation process, all 12 items will be listed in a similar table to proceed with content validation.

2.4 Final version of Sinhala HWISE scale and demographic questions

Upon finalizing the scale, a demographic questionnaire will also be constructed to gather other relevant data from all informants. The demographic questionnaire will contain a series of qualifier questions that will help researchers filter participants to fit the identified sampling frame (Section 2.4.1.). After the qualifiers, the participants will answer questions about their gender, age, profession, area of residence, availability of water resources, time to the nearest water resource, difficult experiences associated with securing usable water to the household, and any injuries participants have encountered while fetching water. The questions to include in the demographic questionnaire will further be decided based on the original HWISE scale validation of Young *et al.* (2019a) and other pertinent contemporary research. Some of the questions will contain a fixed number of polytomous responses while a few may remain open-ended. The answers to open-ended questions will be assessed and quantified based on commonality. All demographic questions will aid in descriptive statistical analyses and validity testing.

2.4.1 Sample

The *Sinhala HWISE Version 3* will be administered among individuals of age 18 years and above, who have knowledge about water acquisition and use in their households. The qualifier questions of the demographic questionnaire will be based on the inclusion criteria specified in the preceding sentence. In addition to the above, to improve the efficiency and the accuracy of the sampling process, Azam *et al.* (2021) recommend identifying the target population to recognize the sampling frame which includes elements within the sample that will be selected for the study. Further information related to the sampling strategy is given in Table 2.

As per the report by the Sri Lanka Department of Census and Statistics (2021), approximately 13 million Sri Lankans are between the ages of 18 and 69. Considering the approximate average number of people within a house is 4, the estimated number of houses could be more or less in the vicinity of 3 – 5 million. Since the majority of Sri Lankans live in rural settings, the data will be collected mostly in rural areas. Considering the highest rainfall is reported in the wet zone, the data will mostly be collected from households in the dry zone and the semi-arid zones which also have much less population density compared to the wet zone. Accordingly, the sample size for this study will be considered based on the sample-to-item ratio suggested by Castello & Osborne (2005), Gorsuch (1983), and Memon *et al.* (2020). The recommended minimum sample required per item as per the above sample-to-item ratio is 5:1. However, in most instances' researchers consider 10:1 and 20:1 ratios as well. As per Costello & Osborne (2005), the chances of obtaining the appropriate factor structures are 70% at 20:1 whereas it is 60% for 10:1 and 40% for 5:1. This sample size calculation ratio is considered for studies with exploratory factor analysis (EFA) and principal component analysis (PCA). In that sense, at 20:1, a minimum of 240 participants will be required for this study. However, to increase the reliability of the study, the researchers of the present study will seek a sample of 300.

Table 2: Sampling strategy for the statistical validation of HWISE scale

Component	Explanation
Target population	Sri Lankans 18 years and above who live with their families and share substantial responsibilities with them in their respective households will be considered for this study. Individuals who may live in hostels, residence halls, etc. will not be considered. Both mother and father of a nuclear family, and adults of an extended family who share key responsibilities in the household also qualify to take part in this study.
Sampling unit	Households (not more than 1 qualifying person from a single household).
Sampling frame	A list of registered households in Sri Lanka from government records that are accessible to researchers.
Sampling technique	Systematic sampling will used to identify a few locations representing the three major climatic zones of Sri Lanka. Within each zone, households will be randomly selected to meet individuals who may qualify to take part in the present study.
Sample size calculation	Sample-to-item ratio (20:1) as suggested by Costello & Osborne (2005), Gorsuch (1983), and Memon *et al.* (2020).

2.4.2 Data collection

Each potential respondent will be provided with study information and if willing to participate, each respondent will be provided with a consent form. The consent form will include a space to write the first letter of the participant's first name, the first letter of the participant's last name, the birth month, and the birth year. The intention is to generate a unique code for each participant in an attempt to isolate the answer script if the respondent wishes to withdraw their data from the study. Participants will be able to withdraw from the study at any time within a time frame of two weeks from the day of submission of responses to the demographic questionnaire by requesting the research team and providing their unique code to the research team. Further, participants will be informed to refrain from including any personal information such as name, surname, and so forth to maintain anonymity. Only upon receiving informed consent, the participant will be provided with *Sinhala HWISE Version 3*. The scale will be self-administered. Once the participant completes their responses, the participant will be provided with a large envelope where the answer script could stay along with the answer scripts of other participants. The study will also administer *Sinhala HWISE Version 3* online using Google Forms to gather data from areas that the authors of the present study find difficult to travel. In instances of online data collection, all previously described good practices will be implemented.

2.4.3 Data storage

All envelopes containing answer scripts will be placed securely in a locked drawer in a locked room, to which only the research team will have access. All digitized data (including data from Google Forms) will be placed in a password-protected folder in a password-protected computer, to which only the research team will have access. All participants will be provided with the contact details of the researcher N.C.S to contact if they wish to withdraw their data from the study. Data from participants who may wish

to withdraw from the study will be destroyed securely. Answer scripts will not be shared with anyone and will be kept confidential.

2.5 Data analysis

The collected data will be first entered into Microsoft Excel. Upon adding data to Microsoft Excel, the researchers will first conduct a thorough data cleaning process to remove entries with missing data. The answers of each participant will be further checked for unusual answer combinations. In instances where the data is collected via Google Forms, the time taken by respondents to complete the scale will also be assessed to potentially exclude any entries that do not at least have utilized a reasonable amount of time (i.e., not at least 45 seconds). Similarly, researchers will assess qualifier questions to eliminate respondents who may not qualify to take part in the study at this stage (i.e., those who live in hostels and resident halls). After this careful data-cleaning process, authors will transfer the data to R Studio where the descriptive analyses and other inferential statistical measures will be performed.

2.5.1 Descriptive statistical analyses

After transferring the Excel file into R Studio, the authors will conduct a series of descriptive analyses using demographic questions listed in the questionnaire. Once this process is completed, summary statistics will be obtained to assess the HWISE scale scores for the sample. For this purpose, a new variable with the aggregated score from each HWISE scale item for every individual household will be created. The summary statistics function helps authors evaluate the overall HWISE scale scores of the sample against measures of central tendency, dispersion, and normality. The following analyses will be performed using the 'pacman' package of R Studio, which is a package that offers a variety of descriptive analyses for researchers (Selvaratnam, 2023).

2.5.2 Classical test theory (CTT)
2.5.2.1 Principal component analysis (PCA)

Scales can be validated using CTT and item response theory (IRT). CTT includes reliability analysis, exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and other forms of validity documenting (i.e., predictive validity, convergent validity, etc.). To proceed with the identified analyses, researchers will consider 'pacman', 'dplyr', 'psych' packages on R Studio. The EFA process will start with the Kaiser-Meyer Olkin (KMO) Index, which evaluates the adequacy of the sample to undergo factor analysis (Field, 2000; Shrestha, 2021). The KMO value ranges from 0 – 1, and adequacy is indicated if the resulting statistic exceeds 0.6.

After completing the assessment on sampling adequacy, the next step in the statistical validation process will be generating a correlation matrix for all 12 items of the scale. The primary objective of creating a correlation matrix is to observe possible signs of multicollinearity, which is when items indicate strong intercorrelations. Generally, correlations above .8 among two items could be considered a sign of multicollinearity

(Bergqvist *et al.*, 2020). Since the original scale is unidimensional, researchers will also look for any possible negative correlations or poor correlations (correlations below .2). Either of negative or poor correlations could imply errors in the translation process. In such instances, the scale's content validation process will be revisited. However, more information on individual items can be further studied through reliability analysis (see section 2.5.2.2). If the correlation matrix fails to indicate any problems, Bartlett's test of sphericity will be performed. Bartlett's test of sphericity checks the suitability of conducting a factor analysis by comparing the current correlation matrix to an identity matrix (Field, 2000). An identity matrix is when items considered for analysis are unrelated. As a result, the sphericity test has to be statistically significant indicating that the current correlation matrix significantly diverges from an identity matrix and is suitable to undergo factor analysis. Both the KMO index and Bartlett's test of sphericity are statistical measures to assess the factorability of the collected data (Shrestha, 2021).

After the KMO index and sphericity test computations, the next step is to observe the internal factor structure of the scale. Although the original scale of Young *et al.* (2019a) has yielded a unidimensional solution, the Sinhala HWISE scale still requires a factor structure exploration due to possible cross-cultural effects. To evaluate the internal structure, principal component analysis (PCA) will be employed. PCA is a form of dimension reduction that helps to identify the number of dimensions, components, or latent factors in a scale (Santos *et al.* 2019; Yong & Pearce, 2013). The significance of each component is assessed through the Kaiser-Guttman eigenvalue criterion (Yong & Pearce, 2013, p. 85). According to the Kaiser-Guttman eigenvalue criterion, components that generate an eigenvalue above 1.0 will be considered significant components. The significant components can also be visualized using a scree plot. The PCA will be performed according to the methodology used by Scholz *et al.* (2002) study on observing psychometric properties of the generalized self-efficacy scale. Any significant component that fails to retain at least 3-factor loadings, and is incapable of explaining a minimum of 5% of the total variance of the observed variables will not be considered a significant component as per the guidelines of Scholz *et al.* (2002). In the present study, authors expect a unidimensional solution. However, in the instances of a multi-factor solution, an oblique rotation will be performed to further look for unidimensionality (Figure 2).

2.5.2.2 Reliability and item-total correlation

Followed by PCA, Cronbach's alpha (α) will be computed to find the internal consistency reliability of the HWISE scale. Generally, α above 0.7 is considered acceptable (Taber, 2018). The raw reliability score of the scale will be further interpreted using Feldt and Duhachek's reliability estimates, Guttman Lambda 6 (G6), signal-to-noise (S/N) ratio, and average standard error (ASE). The purpose of each of the above interpretations is listed in the table below (Table 3).

Table 3: Reliability statistics and interpretations

Type of reliability statistic	Purpose
Feldt and Duhachek estimates	Felt and Duhachek estimates provide a lower and upper bound for α with 95% confidence. The main reason to check this is due context-dependent variability of α. For example, a scale with $\alpha = .83$ may vary between .81 (lower bound) and .85 (upper bound) based on its standard error. Since the lowest bound is .81, this also means that the measure could get contaminated up to 19% (Iacobucci & Duhachek, 2003).
Guttman Lambda (G6)	G6 is an alternative measure of reliability. Higher scores indicate higher reliability (Isernia *et al.*, 2022).
Signal-to-Noise (S/N) ratio	S/N ratio assesses the true score variance to the error variance. Resulting scores between 3 – 4 are generally considered acceptable as lower scores (closer to zero) suggest more error (Cronbach & Gleser, 1964).
Average Standard Error (ASE)	ASE provides an estimate of the precision of the reliability estimate. Lower values closer to 0 are considered acceptable (Duhachek & Iacobucci, 2004). This also helps in determining the confidence estimates of alpha (Iacobucci & Duhachek, 2003).

Further, the raw correlation between each item and the total score will also be inspected. This could further indicate the strength and the direction of the linear relationship between each item and the scale. This would also support the quality of the items in the scale, and it would assist in comprehending how much influence each item possesses in identifying individuals who associate the most with the latent trait (Wang *et al.*, 2017).
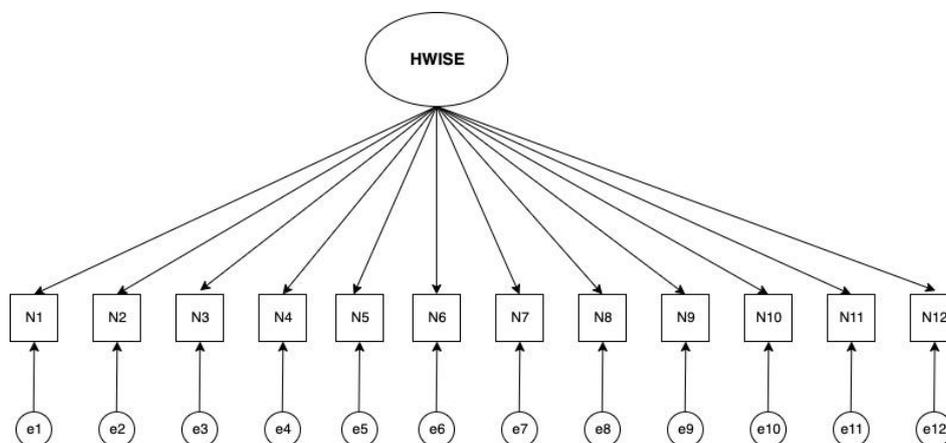

Figure 2: Path analysis of the proposed HWISE scale displaying unidimensionality (created by the authors using draw.io)

2.5.2.3 Confirmatory factor analysis (CFA)

After EFA, the authors will conduct CFA to assess the model fit for the obtained data. Figure 2 displays the model for the CFA of the Sinhala HWISE scale. Compared to EFA, CFA is a restricted factor model and it cannot be rotated. As a result, the model will be tested using the comparative fit index (CFI), root mean square error of approximation (RMSEA), and the normed chi-square (χ^2). The values should be around 2.0 or less for the normed chi-square, less than .08 for RMSEA, and greater than .90 for CFI (Wickrama *et al.* 2023).

2.5.2.4 Construct validity

To display predictive validity, the HWISE scale score will be correlated with the satisfaction with life scale by Ed Diener (Espejo *et al.*, 2021). Convergent validity will be tested by analyzing the association between HWISE scale score and distance to a water source. Similarly, discriminant validity will be evaluated by comparing the HWISE scale score for individuals who have experienced and who have not experienced injuries while fetching water. The simple mean comparison can be achieved through a between-groups t-test. The mean comparison could display individuals with higher water insecurity experiences to encounter more injuries compared to those who do not have exposure to water insecurity. If the authors of the present research could demonstrate a unidimensional solution with predictive, convergent, and discriminant validity satisfactorily, the scale will meet construct validity.

2.5.3 Item response theory and latent trait modeling (LTM)

Once the CTT procedures are completed, the researchers will assess the Sinhala HWISE scale's psychometric properties through an IRT model. The original research by Young *et al.* (2019a) used a Rasch model for IRT analysis. Generally, the Rasch model is used for items with dichotomous answer options. Rasch model considers two parameters, probability and ability of an item. Rasch model can be written as follows.

$$\Pr(X_{ij} = 1) = \frac{e^{(\eta_i - \alpha_j)}}{1 + e^{(\eta_i - \alpha_j)}}$$

In the Rasch model, $\Pr(X_{ij} = 1)$ represents the probability an individual i will respond correctly to item j. Here, η_i is the ability of the person i while α_j is the difficulty parameter specific to item j (Selvaratnam *et al.*, 2018). The Rasch model offers insights into the scoring accuracy of items but lacks a discrimination parameter, preventing differentiation between items in terms of their ability to gauge individuals' positions on the latent trait (An & Yung, 2014). The discrimination parameter present in the two-parameter logistic (2PL) model indicates an item's capacity to distinguish individuals who best align with the latent trait being investigated. The 2PL model also aids in understanding the relative effectiveness of items in discriminating between individuals with varying abilities. The 2PL model can be written as follows.

$$\Pr(X_{ij} = 1) = \frac{e^{\lambda_j(\eta_i - \alpha_j)}}{1 + e^{\lambda_j(\eta_i - \alpha_j)}}$$

Within the 2PL model, λ_j represents the discrimination parameter specific to item j. Similar to the Rasch model, the 2PL model is suitable for measuring items with binary response options. However, to accommodate items with polytomous data, the graded response model (GRM) offers a solution. GRM incorporates ability, difficulty, and discrimination parameters to evaluate items with multiple response categories,

providing a framework specifically tailored for analyzing polytomous data (Thorpe & Favia, 2012). GRM can be written as follows.

$$\Pr\left(X_{ij} \leq k\right) = \frac{e^{\{\lambda_j(\eta_i - \beta_{ik})\}}}{1 + e^{\{\lambda_j(\eta_i - \beta_{ik})\}}}$$

In the above formula, $\Pr\left(X_{ij} \leq k\right)$ signifies the probability that the response to item j by the person i is less than or equal to category k. Within this formula, η_i represents person i's ability on the latent trait, while β_{ik} stands for the threshold parameter specific to item j and category k. The discrimination parameter λ_j signifies the item-specific discrimination ability. The variable k denotes the category of the response for item j, reflecting the cumulative probability of responding up to that category. Accordingly, the authors will evaluate the quality of the Sinhala HWISE scale by running a GRM upon obtaining data.

The GRM offers both constrained and unconstrained versions. In the unconstrained variant, the discrimination parameter varies across items, enabling researchers to assess each item's quality and its alignment with the latent trait. Generally, items with higher discrimination parameters are more effective in measuring the latent trait. Similarly, interpretation of the GRM often involves item information curves (IICs), item category characteristic curves (ICCCs), and the test information function (TIF), aiding researchers in understanding individual item performance and overall test information. After a careful assessment, researchers could test the model fit for constrained vs. unconstrained for further interpretation of the items. Model fit can be interpreted using log-likelihood, Akaike information criterion (AIC), and Bayesian information criterion (BIC) values (Min & Aryadoust, 2021). In the present study, authors will use the R package 'ltm' to conduct the mentioned GRM. The section, 2.5.3.1 summarizes the basic purposes of ICCC to facilitate easier comprehension of the IRT for the readers of this protocol.

2.5.3.1 Item category characteristic curves (ICCCs)

In R Studio, through a GRM, researchers can generate ICCCs for all 12 items (Figure 3). Obtaining ICCCs is part of the GRM in IRT. Each figure (ICCC) will display information about the degree to which each response in the item can assess the ability of an individual. Here, ability (θ) is the latent variable. Each curved line represents an answer option and researchers could interpret which response options are more likely to be selected at a particular level of (θ). In the sample ICCCs given below (Figure 3), answer options 1 and 2 are more likely to be selected even at lower (θ) as opposed to items 3 and 4 which are more difficult choices.

For instance, hypothetically assume that the above ICCCs belong to item 1 of the HWISE scale. Accordingly, a person who has encountered more water insecurity experiences (θ) is more likely to select answer option 4 as opposed to the rest of the items. Additionally, the discriminability of item choices can be further assessed by observing

the steepness of each curve. Higher steepness is associated with higher discriminability, and higher discriminability indicates the item's relatedness to the latent trait similar to loadings to a factor in factor analysis (Hays, Morales, & Reise, 2000). In the sample ICCCs (Figure 3) given below, discriminability is kept constant. Similar to ICCC, IIC provides insights about which items are capable of providing more information about the latent variable and TIF provides how well the overall scale provides information about the latent variable (Sethar *et al.* 2022). Thus, in addition to CTT, researchers can learn more about the overall quality of the items in the HWISE scale using IRT.
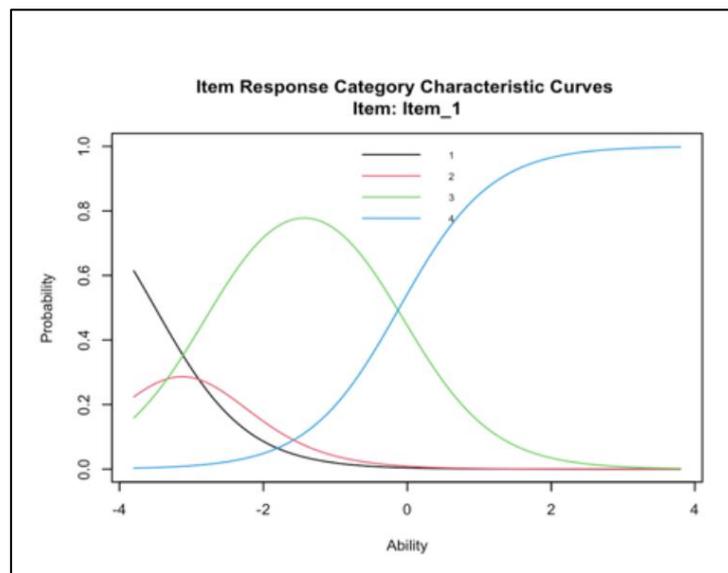
Figure 3: Sample ICCCs for an item with polytomous responses

2.6 Ethics approval

The ethics for the present study was obtained from the Ethics Review Committee of the Faculty of Medicine, University of Peradeniya, Sri Lanka. This study is designed as per the prescribed research guidelines of the American Psychological Association's (APA) and the British Psychological Society's (BPS) code of ethics.

2.7 Timeline of the study

The project is expected to be completed within the year 2024. However, if multiple revisions are required to finalize the scale, the project may require more time in the following year as well.

3. Discussion and conclusion

The present study is expected to generate a culturally adapted and statistically validated scale to assess household water insecurity experiences in Sri Lanka. Considering the psychometric properties of the previous studies on the HWISE scale construction, the present study also expects to generate a unidimensional factor solution with satisfactory reliability. The protocol paper is carefully designed utilizing the insights of Scholz *et al.*

(2002) and Young *et al.* (2019a). Similarly, the study will combine both CTT and IRT demonstrating a rigorous evaluation of psychometric properties of the Sinhala HWISE scale. While these processes enhance the overall quality of the scale and the procedural rigor of the study, they also render a noticeable set of benefits to the scientific research community. Firstly, the availability of the scale in the Sinhala language will enable responsible authorities to identify water insecurity experiences among households in multiple parts of Sri Lanka. Such identification will have a direct influence on the health and economic policies of Sri Lanka. For instance, enacting the right policies by understanding water stress and insecurity could help authorities mitigate potential negative impacts on health and the economy. Moreover, the scale will enable other researchers to conduct further research about water conservation and distribution by steering through the community experiences of household water insecurity. Such initiatives and research could have a direct impact on the health and economic policy development of Sri Lanka.

Secondly, understanding more about individual experiences related to water insecurities also carries a psychological component. Individuals with water insecurities may experience psychological disturbance which should be addressed by implementing evidence-based psychotherapeutic interventions. Accordingly, by understanding more about the nature of water insecurity Sri Lankans face, more policies could be implemented to facilitate health promotion. As Chandrasekara *et al.* (2021) and Wickrama *et al.* (2023) reason, the majority of policies Sri Lankan authorities implement for water-related problems (i.e., rainfall management) are not effective. In that case, the availability of a scale enables researchers to assess the possibility of adopting more nomothetic approaches to assist vulnerable populations in circumnavigating water stress. Furthermore, future research could focus more on understanding the experiences of people by considering qualitative and mixed methods approaches to explore the depth of household water problems in Sri Lanka.

Conflict of Interest Statement
The authors declare no conflicts of interest.

About the Author(s)
Navodya C. Selvaratnam earned her Post-Master Degree in Public Health and is presently attached to the University of Peradeniya, Sri Lanka in the capacity of a research assistant.
Naren D. Selvaratnam is a Visiting Lecturer of Psychology at the Sri Lanka Institute of Information Technology (SLIIT), and Sri Lanka Technological Campus (SLTC).

Janani S. Nanayakkara is an M.Phil. degree candidate at the Faculty of Medicine, University of Peradeniya, Sri Lanka.

Sampath Tennakoon is a Professor of Community Medicine at the Department of Community Medicine, Faculty of Medicine, University of Peradeniya, Sri Lanka.

References

Abbott, B. W., Bishop, K., Zarnetske, J. P., Minaudo, C., Chapin, F. S., Krause, S., Hannah, D. M., Conner, L., Ellison, D., Godsey, S. E., Plont, S., Marçais, J., Kolbe, T., Huebner, A., Frei, R. J., Hampton, T., Gu, S., Buhman, M., Sara Sayedi, S., & Pinay, G. (2019). Human domination of the global water cycle is absent from depictions and perceptions. *Nature Geoscience*, *12*(7), 533-540. https://doi.org/10.1038/s41561-019-0374-y

An, X., & Yung, Y. (2014). Item Response Theory: What it is and How You Can Use the IRT Procedure to Apply it. SAS Institute Inc.

Azam, S. M. F., Yajid, M. S. A., Tham, J., Hamid, J. A., Khatibi, A., Johar, M. G. M., & Arrifin, I. A. (2021). Research methodology: Building research skills, McGraw Hill Education (Malaysia) Sdn Bhd, Kuala Lumpur: Malaysia

Bergqvist, E., Tossavainen, T., & Johansson, M. (2020). An Analysis of High and Low Intercorrelations between Mathematics Self-Efficacy, Anxiety, and Achievement Variables: A Prerequisite for a Reliable Factor Analysis. *Education Research International*, 1–11. https://doi.org/10.1155/2020/8878607

Chandrasekara, S. S. K., Chandrasekara, S. K., Gamini, P. H. S., Obeysekera, J., Manthrithilake, H., Kwon, H. H., & Vithanage, M. (2021). A review on water governance in Sri Lanka: the lessons learnt for future water policy formulation. *Water Policy*, *23*(2), 255-273. https://doi.org/10.2166/wp.2021.152

Cronbach, L. J., & Gleser, G. C. (1964). The signal/noise ratio in the comparison of reliability coefficients, *Educational and Psychological Measurement, XXIV*(3), 467-480

Costello, A., & Osborne, J. (2005). Best Practices in Exploratory Factor analysis: Four Best Practices in Exploratory Factor analysis: Four Recommendations for Getting the Most from Your Analysis Recommendations for Getting the Most from Your Analysis. *Research, and Evaluation Practical Assessment, Research, and Evaluation*, *10*(7), 7. https://doi.org/10.7275/jyj1-4868

Department of Meteorology. (2019). *Climate of Sri Lanka.* https://www.meteo.gov.lk/index.php?option=com_content&view=article&id=94&Itemid=310&lang=en&lang=en#2-southwest-monsoon-season-may-september

De Zoysa, P., Rajapakse, L., & Newcomb, P., A. (2007). Adaptation and Validation of the Personality Assessment Questionnaire on 12-Year-Old Children in Sri Lanka. In New Psychological Tests and Testing Research. ed by. Boyar, L, S. New York: Nova Science Publishers, Inc. 185-202.

Duhachek, A., & Iacobucci, D. (2004). Alpha's Standard Error (ASE): An accurate and precise confidence interval estimate. *Journal of Applied Psychology, 89*(5), 792-808. Doi:10.1037/0021-9010.89.5.792

Espejo, B., Martin-Carbonell, M., Checa, I., Paternina, Y., Fernandez-Daza, M., Higuita, J. D., Albarracin, A., & Cerquera, A. (2021). Psychometric properties of the Diener satisfaction with life scale with five response options applied to the Colombian population, *Frontiers Public Health, 9*, https://doi.org/10.3389/fpubh.2021.767534

Field, A. (2000). *Discovering Statistics Using IBM SPSS Statistics*. SAGE Publications Inc.

Gronier, G. (2022). Psychometric analyses in the transcultural adaptation of psychological scales. In S. Misciagna (Ed), *Psychometrics – New insight in the diagnosis of mental disorders*. IntechOpen. https://doi.org/10.5772/intechopen.105841

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Psychology Press. https://doi.org/10.4324/9780203781098

Gunatilaka, A. (2008). Water security and related issues in Sri Lanka: the need for integrated water resource management (IWRM). *Journal of the National Science Foundation of Sri Lanka, 36*, 3. https://doi.org/10.4038/jnsfsr.v36i0.8045

Hannah, D. M., Lynch, I., Mao, F., Miller, J. D., Young, S. L., & Krause, S. (2020). Water and sanitation for all in a pandemic. *Nature Sustainability, 3*(10), 773-775. https://doi.org/10.1038/s41893-020-0593-7

Hays, R. D., Morales, L. S., & Reise, S. P. (2000). Item response theory and health outcomes measurement in the 21st century. *Medical Care, 38*(9 Suppl), II28-42. https://doi.org/10.1097/00005650-200009002-00007

HWISE Research Coordination Network. (2019). *User Manual: Household Water Insecurity Experiences Scale*. https://doi.org/10.21985/n2-anbv-v974

Iacobucci, D., & Duhachek, A. (2003). Advancing alpha: Measuring reliability with confidence, *Journal of Consumer Psychology, 13*(4), 478-487.

Isernia, S., Rossetto, F., Shamay-Tsoory, S., Marchetti, A. & Baglio, F. (2023). Standardization and normative data of the 48-item Yoni short version for the assessment of theory of mind in typical and atypical conditions. *Front. Aging Neurosci.* 14:1048599. https://doi.org/10.3389/fnagi.2022.1048599

Jepson, W., Wutich, A., Colllins, S., Boateng, G., & Young, S. (2017). Progress in household water insecurity metrics: a cross-disciplinary approach. *WIREs Water, 4*(3): e1214. https://doi.org/10.1002/wat2.1214

Kangmennaang, J., & Elliott, S. J. (2021). Linking water (in)security and wellbeing in low- and middle-income countries. *Water Security, 13*, 100089. https://doi.org/10.1016/j.wasec.2021.100089

Kimutai, J. J., Lund, C., Moturi, W. N., Shewangizaw, S., Feyasa, M. B., & Hanlon, C. (2023). Evidence on the links between water insecurity, inadequate sanitation, and mental health: A systematic review and meta-analysis. *PLOS ONE, 18*(5), e0286146–e0286146. https://doi.org/10.1371/journal.pone.0286146

Kuzma, S., Saccoccia, L., & Chertock, M. (2023, August 16). *25 Countries, Housing One-quarter of the Population, Face Extremely High Water Stress*. World Resources

Institute. https://www.wri.org/insights/highest-water-stressed-countries#:~:text=The%20most%20water%2Dstressed%20regions,%2C%20where%2074%25%20is%20exposed

Memon, A. M., Ting, H., Cheah, J.-H., Thurasamy, R., Chuah, F., & Cham, T. (2020). Sample size for survey research: Review and recommendations. *Journal of Applied Structural Equation Modeling*, *4*(2), i–xx. https://jasemjournal.com/wp-content/uploads/2020/08/Memon-et-al_JASEM_-Editorial_V4_Iss2_June2020.pdf

Miller, J. D., Vonk, J., Staddon, C., & Young, S. L. (2020). Is household water insecurity a link between water governance and well-being? A multi-site analysis. *Journal of Water, Sanitation and Hygiene for Development*, *10*(2), 320-334. https://doi.org/10.2166/washdev.2020.165

Min, S., & Aryadoust, V. (2021). A systematic review of item response theory in language assessment: Implications for the dimensionality of language ability. *Studies in Educational Evaluation*, *68*(68), 100963. https://doi.org/10.1016/j.stueduc.2020.100963

Northwestern University Institute for Policy Research. (2019a). *Water Insecurity Experiences (WISE) Scale: Implementation materials*. https://www.ipr.northwestern.edu/wise-scales/measure-water-insecurity/implementation-materials/

Popkin, B., D'Anci, K., & Rosenberg, I. (2010). Water, hydration, and health. *Nutrition Reviews*, *68*(8), 439-458. https://doi.org/10.1111/j.1753-4887.2010.00304.x

Rosinger, A., & Young, S. (2020). The toll of household water insecurity on health and human biology: Current understandings and future directions. *WIREs Water*, *7*(6): e1468. https://doi.org/10.1002/wat2.1468

Santos, R. de O., Gorgulho, B. M., Castro, M. A. de, Fisberg, R. M., Marchioni, D. M., & Baltar, V. T. (2019). Principal Component Analysis and Factor Analysis: differences and similarities in Nutritional Epidemiology application. *Revista Brasileira de Epidemiologia*, *22*. https://doi.org/10.1590/1980-549720190041

Scholz, U., Dona, B, G., Sud, S., & Schwarze, R. (2002) Is General Self-Efficacy a Universal Construct? Psychometric Findings from 25 Countries. European Journal of Psychological Assessments, 18(3), 242-251.

Selvaratnam, N. D., Hettiarachchi, D., Dantanarayana, D., Bergen, S., Selvaratnam, S., & Ponnamperuma, L. (2018). Investigating the psychometric properties, and IRT analysis of the Sinhala generalized self-efficacy scale (S-GSES), *GARI International Journal of Multidisciplinary Research*, *4*(2), 73-95.

Selvaratnam, N. D. (2023). *How to Run Essential Analyses in R Studio: For First Year Undergraduate Students with No Background in Programming*. DHC Publishers: Colombo, Sri Lanka. https://doi.org/10.61495/ejcr.0001.23

Sethar, W. A., Pitafi, A., Bhutto, A., Nassani, A. A., Haffar, M., & Kamran, S. M. (2022). Application of Item Response Theory (IRT)-Graded Response Model (GRM) to Entrepreneurial Ecosystem Scale. *Sustainability*, *14*(9), 5532. https://doi.org/10.3390/su14095532

Shrestha, N. (2021). Factor Analysis as a Tool for Survey Analysis. *American Journal of Applied Mathematics and Statistics*, *9*(1), 4–11. https://doi.org/10.12691/ajams-9-1-2

Sri Lanka Department of Census and Statistics (2021). *Population* http://www.statistics.gov.lk

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

Thorpe, G., & Favia, A. (2012). *Data Analysis Using Item Response Theory Methodology: An Introduction to Selected Programs and Applications*. https://digitalcommons.library.umaine.edu/cgi/viewcontent.cgi?article=1019&context=psy_facpub

UNESCO. (2023). *Imminent risk of a global water crisis, warns the UN World Water Development Report 2023*. https://www.unesco.org/en/articles/imminent-risk-global-water-crisis-warns-un-world-water-development-report-2023

UNICEF. (2020). *Water scarcity: Addressing the growing lack of available water to meet children's needs*. https://www.unicef.org/wash/water-scarcity

United Nations (UN). (2023). *World Water Day 22 March*. https://www.un.org/en/observances/water-day

Wang, M., Batt, K., Kessler, C., Neff, A., Iyer, N. N., Cooper, D. L., & Kempton, C. L. (2017). Internal consistency and item-total correlation of patient-reported outcome instruments and hemophilia joint health score v2.1 in US adult people with hemophilia: results from the Pain, Functional Impairment, and Quality of life (P-FiQ) study. *Patient Preference and Adherence*, *Volume 11*, 1831–1839. https://doi.org/10.2147/ppa.s141391

World Health Organization (WHO). (2017). *Diarrhoeal disease*. https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease

WHO. (2022). *Arsenic*. https://www.who.int/news-room/fact-sheets/detail/arsenic#:~:text=Long%2Dterm%20exposure%20to%20inorganic,are%20the%20most%20characteristic%20effects

WHO. (2023). *Drinking-water*. https://www.who.int/news-room/fact-sheets/detail/drinking-water

Wickrama, K. A., Wickrama, K. A. S., Merten, M. J., Thanigaseelan, S., Gunasingham, S., & Sathiyavasan, S. (2023). Social, family context, and prevention of mosquito-borne diseases in developing countries: A case of dengue in Sri Lanka, *Eagan Journal of Contemporary Research*, *2*(1), 9-22. https://doi.org/10.61495/ejcr.v2.1.2

WorldBank. (2022). *Data: Sri Lanka*. https://data.worldbank.org/country/sri-lanka

Yong, A. G., & Pearce, S. (2013). A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. *Tutorials in Quantitative Methods for Psychology*, *9*(2), 79–94. https://doi.org/10.20982/tqmp.09.2.p079

Young, S., Boateng, G., Jamaluddine, Z., Miller, J., Frongillo, E., Neilands, T., Collins, S., Wutich, A., Jepson, W., & Stoler, J. (2019a). The Household Water InSecurity Experiences (HWISE) Scale: development and validation of a household water

insecurity measure for low-income and middle-income countries. *BMJ Global Health*, *4*(5): e001750. http://dx.doi.org/10.1136/bmjgh-2019-001750

Young, S., Collins, S., Boateng, G., Neilands, T., Jamaluddine, Z., Miller, J., Brewis, A., Frongillo, E., Jepson, W., Melgar-Quiñonez, H., Schuster, R., Stoler, J., & Wutich, A. (2019b). Development and validation protocol for an instrument to measure household water insecurity across cultures and ecologies: The Household Water InSecurity Experiences (HWISE) Scale. *BMJ Open*, *9*(1): e023558. http://dx.doi.org/10.1136/bmjopen-2018-023558.