Hand Hygiene for All

Monitoring Working Group

Working paper: Mapping and gap analysis of tools designed to collect data on hand hygiene in public spaces

June 2021

WHO UNICEF JMP

TABLE OF CONTENT

List of tables

List of figures

ACCRONYMS

HH4A	Hand Hygiene for All initiative
HHF	Hand hygiene facility
MWG	Monitoring working group (under the HH4A)
UNICEF	United Nations Children's Fund
WASH	water, sanitation and hygiene
WHO	World Health Organization

1 INTRODUCTION

1.1 BACKGROUND

The Hand Hygiene for All Global Initiative (HH4A) is led by WHO and UNICEF. It aims to implement WHO's global recommendations on hand hygiene to prevent and control the COVID-19 pandemic and work to ensure lasting infrastructure and behaviour. The initiative is a call to action for all of society to achieve universal hand hygiene and stopping the spread of COVID-19 and followed the WHO call for mandatory hand hygiene in public spaces. Under HH4A, templates for Country Action Plans are made available to support the development of effective country-wide hand hygiene promotion campaigns. More information on the initiative can be found here https://www.who.int/water_sanitation_health/sanitation-waste/sanitation/hand-hygiene-for-all/en/

The monitoring working group (MWG) under HH4A initiative is conducting a mapping of existing monitoring tools. Such tools can be resources containing targets, standards/requirements, indicators, or monitoring tools (questionnaires, spot-checks, etc.). The purpose of the mapping is understanding to what extent gaps exist in the availability of tools for different settings. Based on the mapping, where necessary, recommendations will be developed. These recommendations can entail a proposition for common indicators, survey methods and/or data collection mechanisms.

Before the development of recommendations, the MWG seeks feedback on the framework that is employed to assess the gaps. A solid review framework will contribute to the development of good recommendations. In addition, the framework can subsequently serve the review of tools for other settings where gaps exist (such as detention centres, work environments).

1.2 THE CATALOGUE

Through the network of the HH4A partners, over 100 monitoring resources have been shared with the MWG and catalogued. These resources cover a wide range of settings, including households, schools, health care facilities, work environments, detention centres and public spaces. Currently, the catalogue is a combination of an excel workbook and a dropbox folder containing copies of all the materials. The catalogue is available here (version 4): https://www.dropbox.com/sh/vcxjwwi5fukpdx4/AACOIrcdmwmECjVe5tfzywWGa?dl=0

The resources in the catalogue include targets, standards/requirements, indicators, monitoring tools, manuals, complete monitoring frameworks. The monitoring tools use diverse sampling strategies, measurement techniques, and data collection methods which collect information on a wide range of hand hygiene parameters. These parameters include elements of the enabling environment, interventions/campaigns, availability of hand hygiene stations and consumables, demand, and psycho-social factors. A wide range of setting is included: health care facilities, schools, households, prisons, work environments, public spaces and camps/humanitarian contexts. The catalogue contains tools developed in 34 different countries and resources which have been developed at global level. Most tools are not specifically designed to monitor hand hygiene in the context of the COVID-19 pandemic.

The mapping showed a lack of available tools for the following settings: public spaces, detention centres, and work environments. Given the importance of hand hygiene in public spaces in response to the COVID-19 pandemic, the MWG wants to prioritise this important issue and analyse the gap.

1.3 OBJECTIVES OF THE ASSESSMENT

The specific objective of this assessment is to better understand the characteristics and scope of existing tools that are used worldwide to monitor hand hygiene in public spaces. This is done as part of the HH4A initiative and it is carried out by the Monitoring Working Group (MWG). The objective of this evaluation is fourfold:

- Map the characteristics of the methodologies of the different monitoring tools (chapter 2)
- Map the scope of the different monitoring tools and identify potential gaps (chapter 2)
- Assess the quality of the measurements of the different monitoring tools (chapter 3)
- Assess the data and indicators of the different monitoring tools (chapter 4)

1.4 METHODOLOGY

To analyse the characteristics and scope of hand hygiene monitoring in public spaces, twelve monitoring tools are assessed which have been designed and used around the world (an overview of these tools is part of chapter 2). These tools are analysed in the following three chapters. Chapters 2 and 3 have their own conceptual frameworks for conduct a specific analysis:

- In chapter 2, the mapping of the methodology of the tools is based on a synthesis of elements commonly used to assess hand hygiene measurement instruments. The identification of the content of the tools is based on commonly used hygiene and water supply frameworks and results-chains.
- In chapter 3, the assessment of the quality of measurements is based on concepts commonly used to assess survey questions and observations (reliability, validity, bias)

Each of these chapters starts with an introduction of the methods by presenting the conceptual framework followed by the results of the analysis and the conclusions.

LIMITATIONS

This assessment has certain **limitations**. The tools are assessed based solely on the content of the available documents, without using feedback from survey teams who have used the tools in the context for which they have been designed. In addition, for most tools limited supporting documents are available to the MWG. The lack of context makes it difficult to judge the relevance or quality of certain questions or observations. Although information gathered by some tools might not seem actionable to an outsider, they may nonetheless have relevance to programme managers. These limitations have to be considered when interpretating the conclusions.

1.5 PUBLIC SPACES

For the purpose of this assessment, public spaces are defined using the recommendations of the WHO and additional comments by the special rapporteur on the right to water and sanitation. WHO recommends that universal access to hand hygiene facilities should be provided in front of all public buildings and transport hubs – such as markets, shops, places of worship, schools, healthcare facilities and train or bus stations. In addition, WHO recommends that functioning handwashing facilities with water and soap should be available within 5 m of all toilets, both public and private (WHO, 2020; WHO / UNICEF, 2020). The Special Rapporteur on the human rights to safe water and sanitation recently underlined the importance of WASH services in parks, streets, markets and transport hubs (OHCHR, 2019).

For the purpose of this assessment, tools which are designed to monitor hand hygiene in such spaces can be included. Tools that have been designed to assess the WASH services inside buildings or institutions (such as schools, health care facilities, prisons) and inside work environments are not part of this review.

The complexity of public spaces

The management and monitoring of hand hygiene in public spaces can be complex. The advancement of hand hygiene in public spaces depends on both the availability of hand washing facilities, their proper functioning and maintenance as well as their widespread and correct use by individuals. While the former depend on the implementation and maintenance of the necessary hand washing facilities, the latter greatly relies on a behavioural change within society.

Adding to the above it is important to recognize that some spaces are open day and night, others have specific opening hours or are used once a week or once a month. Estimating the acceptable number of hand hygiene facilities requires knowledge of the floating population and the moment of crowding. At times it may be challenging to assign clear accountability for the hand hygiene facility, or to find a respondent when interview or survey data is required. In addition, there are spaces without clearly defined limits and without a clear entrance. There is a large variety of public spaces (see previous paragraph) with sizes that vary considerably. In many contexts there can be a lack of comprehensive lists and registers of public spaces which can be monitored.

The Special Rapporteur highlighted an additional complexity of public spaces, related to the adequate WASH coverage on each site in a 2019 report:

> *The precise quantitative features of the provision will depend upon the circumstances on the ground, including the type of space being considered and its size, the usage of that space, in terms of the numbers of people present, the demographics of users, and the reasons for their attendance therein, as well as the length of time people typically remain within a particular space. Having considered these and other pertinent factors, it is essential that States ensure the construction and maintenance of water and sanitation facilities in those spaces in line with the needs of all users and in sufficient quantities to ensure their ease of access.* (OHCHR, 2019)

Ideally tools would explain how the complexity of monitoring in public spaces is addressed.

WHO's recommendations on hand hygiene in public spaces

The WHO call for mandatory hand hygiene in public spaces has implications for monitoring. It includes statements about various elements, such as the availability of hand hygiene stations, its location, the accessibility, accountability, supervision, operation, maintenance and concerning regulation. This is the list of recommended measures by WHO to help prevent the transmission of the COVID-19 virus in public spaces:

1. One or several hand hygiene stations (either for handwashing with soap and water (a) or for hand rubbing with an alcohol-based hand rub)(b) should be placed in front of the entrance of every public (including schools and healthcare facilities) or private commercial building, to allow everyone to practice hand hygiene before entering and when leaving it.
    - *Note (a). Where alcohol-based hand rub or bar soap is not feasible, a liquid soap solution, mixing detergent with water, can be used. The ratio of detergent to water will depend on types and strengths of locally available product.*
    - *Note (b). Chlorine hand washing solutions are not recommended because of potential harm to users and those making the solutions, as well as degradation of chlorine exposed to sunlight or heat. Soap is generally cheap and easy to find, and liquid soap solutions can also be used.*
2. Facilities should be provided at all transport locations, and especially at major bus and train stations, airports, and seaports.
3. The quantity and usability of the hand hygiene stations should be adapted to the type (e.g. young children, elderly, those with limited mobility) and number of users to better encourage use and reduce waiting time.
4. The installation, supervision, and regular refilling of the equipment should be the overall responsibility of public health authorities and delegated to building managers. Private sector and civil society initiatives to support the commodities, maintenance, and effective use are welcome.
5. The use of public hand hygiene stations should be obligatory before passing the threshold of the entrance to any building and to any means of public transport during the COVID-19 pandemic. Repeated hand hygiene whenever outside private homes can in this way become part of the routine of everyday life in all countries.

(WHO, 2020)

Other recommendations or standards exist, which are less comprehensive than WHO 2020[1].

2 THE FOCUS AND METHODOLOGIES OF THE MONITORING TOOLS

This chapter looks at the characteristics of the methodologies and the focus of the tools that are used to monitoring hand hygiene in public spaces. The tools are mapped against key characteristics, a results-chain and a behaviour change framework. This chapter also describes the sampling strategies of the tools.

Each tool is designed for a specific purpose. Comparing the content of the tools with the different frameworks will highlight potential gaps in monitoring hand hygiene in public settings in general. This does not mean that individual tools contain gaps or are insufficient.

2.1 METHODS

The characteristics of the monitoring methodologies are described using a framework that is based on a synthesis of seven recent reviews[2] of measurement instruments in the domain of hand hygiene. Annex A provides an overview of the recent reviews and these characteristics. Table 1 presents the resulting framework which contains 9 elements to characterise the methodologies **and 2 methods for mapping the content of the tools.**

It's important to note that the assessment of the characteristics of the 12 tools is based on a review of the information that is available to the MWG. Institutions who have designed the tools may provide more background information on these tools which can be used to update the description of the characteristics in the future.

Table 1. The framework used to map the content and the methodologies of the monitoring tools

Element	Description	The 12 monitoring tools analysed in this report have the following characteristics

Data collection method	The method of data collection refers to the process used to collect data: who collects data and how? A distinction is made between data that is collected in-person vs. remote data collection.
	In-person means the data is collected by an official surveyor who is present at given the
		– **In-person**
		  – Official surveyors
		– **Remote**
		  – Public mobile app-based survey:
		  – With volunteers/citizens
		  – HHF caretakers
		  – Online survey
		  – Aggregation of NGO activity data

[1] For example:

- Wash stations and hand sanitizers are compulsory in all public spaces (Government of Rwanda 2020)

- Handwashing with soap or alcohol-based sanitizer at critical times, particularly after coughing or sneezing, after visiting of public spaces (public transport, markets, places of worship, etc.), after touching any surfaces outside the house, and before and after visiting/caring for at-risk or sick people, before and after handling a mask. (UNICEF 2020b)

- Public handwashing stations must be established in a place that makes handwashing convenient during critical times (within the communal areas). Both soap and water for handwashing must be present at the designated location (World Vision International 2020)

[2] (Oliveira and Paula, 2011; Vindigni, Riley and Jhung, 2011; Ram, 2013; Valim et al., 2014; Van Remoortel et al., 2017; Jeanes et al., 2019; Rutter et al., 2019)

	public place. Remote data collection refers to a range of different data collection methods where no official surveyor is needed. Public mobile app-based surveys, online surveys and the aggregation of NGO activity data are examples of remote data collection methods	
Measurement technique	The measurement technique refers to the method used to assess the different aspects of hand hygiene. Examples of common techniques are self-reports, structured observations, spot-checks, product consumption, visual inspection of hand cleanliness etc. The tools that are assessed use a total of 5 measurement techniques. The strengths and weaknesses of these techniques is presented in A key characteristic of tools is the measurement technique. Existing reviews on hand hygiene have provided an overview of the strengths and limitations of these different measurement techniques (see for example (Ram, 2013; Rutter *et al.*, 2019; GHP, 2020)). **Erreur ! Référence non valide pour un signet.** presents an overview of these strengths and limitations of the techniques used by the 12 tools. Table 3.	– Self-reports (on tasks/responsibilities) – Spot checks on the presence of infrastructure or consumables – Observations of hand hygiene practices (comprehensive) – Observations of hand hygiene practices (limited) – Indirect reports by caretakers of HHF
Respondent	When a tool includes questions, the interviewee is the respondent. In the available tools, interviewees can be HHF caretakers, certain local committees or members of the general population. When a tool includes observations or spot-checks, a situation is interpreted by e.g. a surveyor, who then de-facto becomes the respondent. There are also cases where an NGO is the respondent.	– Official surveyors – WASH committees – Volunteers/citizens – HHF caretakers – Population participating in an online survey – NGOs (activity reports)
Scale	Some tools are specifically designed for collecting data at	– National – Sub-national

	national scale. Others are used at sub-national level, such as project sites or in cities. NB: Tools which are currently used at sub-national level could at a different stage or in a different context be used at national level.	
Sampling strategy	Sampling strategies are important to obtain information on the coverage of hand hygiene in public spaces. Information on the sample size and sampling process are known for a limited number of tools. In most contexts there can be a lack of comprehensive lists and registers of public spaces which could/should be covered by monitoring programmes. Not all tools are expected to have a sampling strategy (such as activity reports of NGOS).	– Stratification and random sampling, (quasi representative) – Purposeful sampling – Random sampling by mail through probability, address-based sampling
Setting details	This refers to the type of public space for which the monitoring tool is designed, for example, markets, transport hubs, public toilets. Not all tools are designed for specific settings but rather for public spaces in general.	– Markets and transport hubs – Public/Communal Toilets – Various public spaces (in front of schools, health facilities, religious places, transport stations, public transportation, traditional markets, shopping centres, community halls)
Handwashing moment	Although in the available documentation most tools are not explicit about the purpose of the monitoring, the specific hand hygiene moment of interest can be inferred from the content of the tools. The overview of tools indicates what specific hand hygiene moment is monitored.	– After using a public toilet – Before entering a public space – After using a public toilet/Before eating in a restaurant – Not Specified
Costs/resource requirements	This assessment does not use cost data of the various tools, but the relative costs of data collection methods that are used can be compared. Information on the relative costs of these data collection methods is available in existing reviews (Ram, 2013; Rutter et al., 2019; GHP, 2020).	
Data availability	A share of the tools have been implemented and the data has been made public. For some	– Data available – Data not available

	tools data is currently not available (or not shared yet for the purposes of this assessment).	
Mapping the content of the tools		
Content of the tools 1: results-chain	The tools monitor specific elements of hand hygiene in public spaces. These elements can be mapped on a results-chain for hand hygiene programme implementation. A simplified results-chain is developed for this purpose using the elements of existing hand hygiene and water supply frameworks. Here six[3] [4] frameworks have been selected that together constitute a comprehensive list of elements that together consititute the results-chain in Table 2. • The WHO recommendations on hand hygiene in public spaces (WHO, 2020) • The elements under the human rights to water and sanitation (OHCHR, 2014) • The elements under the hand hygiene surveillance in schools (WHO EUROPE, 2019) • The Rural Water Metrics Global Framework (World Bank Group, 2017) • Technical Guide for handwashing facilities in public places and buildings (WaterAID, 2020) • The elements of the HH4A results-chain (HH4A, 2020)	See Table 2 for a detailed description of the results-chain.
Content of the tools: hand hygiene determinants	The content of the tools is also compared to a behaviour change framework. Although the tools are not necessary used to inform hand hygiene promotion programmes, to map them against such frameworks can show if important areas for behaviour change are omitted. The Behaviour Centred Design framework (BCD) is used for this additional mapping, which contains a comprehensive set of behavioural determinants (White *et al.*, 2020). **Erreur ! Source du renvoi introuvable.**Figure 1 provides definitions of each BCD determinant adapted for hand hygiene. The BCD covers five categories of determinants that influence hand hygiene practices: brain, body, behaviour setting, environment, and external context. Additional (non-determinant) categories include interventions (inputs), behaviours (outcomes) and changes to the state-of-the-world (impact). In addition to a comprehensive set of psycho-social factors, the BCD framework gives importance to the "behaviour setting", such as infrastructure, consumables (called props) but also the "stage". The stage includes *the design and set up of the specific physical spaces where handwashing behaviour takes place*. The	See Figure 1 for an overview of the BCD determinants. Figure 2 maps the determinants of the BCD framework on the results-chain that is presented in Table 2

[3] WHO 2020 has been selected because it forms the reference for the promotion for hand hygiene in public spaces; OHCHR 2014 because of its important for WASH programming; WHO 2019 contains the most comprehensive set of hand hygiene criteria in institutions that is currently part of the catalogue of monitoring tools; WSP 2017 is a synthesis of water supply indicators; HH4A 2020 is a results-chain for the development of country level action plans for achieving hand hygiene for all.

[4] The elements under each of the 7 frameworks is presented in annex D

	reliability of the water supply, the location of the facility within the public space, its accessibility and the management of the hand hygiene facility all make up the stage. The stage can be particularly important for public spaces, especially today as hand hygiene facilities are often installed where hand hygiene was not actively encouraged before. There is evidence that an optimal behaviour setting constitutes one of the most important determinants for hand hygiene behaviour (White *et al.*, 2020).	

Table 2 presents the simplified results-chain based on the 6 hand hygiene and water supply frameworks[5] (see Table 1, methods for mapping the content of the tools). This results-chain will be used to map the tools and to conduct a gap analysis. It's important to note that there is no causal relationship between the three catagories of outputs. Outputs have been divided into three categories for the purpose of the mapping. In addition, the elements in the results-chain have complex relationships and some might actually overlap to some degree (see (Kayser et al., 2013)). Nevertheless, this comprehenisve list is helpful for a general gap analysis

A visual representation of the elements of the results-chain is shown in Figure 2.

Table 2. A simplified results-chain for hygiene programming that will be used to map the content of the 12 tools and to conduct the gap analysis

Element of the results-chain	Description
The enabling environment	
The HH4A enabling environment	Implies measurements that assess the political leadership, champions across all levels of government and society in addition to the areas under Sanitation and Water for All (SWA), namely, policies & strategies, institutional arrangements, financing, planning, monitoring & reporting, capacity development
The HH4A accelerators (based on the SDG 6 Global Acceleration Framework[6])	Implies measurements that assess governance and multi-stakeholder engagement, financing, data information & monitoring, Capacity development, Innovation, evidence and learning
Inputs: Activities: Supply of services and promotion of behaviours	
Activities related to hand hygiene in public spaces	Implies measurements that assess the completion of activities related to the supply of services and promotion of behaviours in public spaces (e.g. installation of HHFs, the promotion of hand hygiene, capacity building, etc.)
Outputs 1: Psycho-social factors and capabilities	
Psycho-social factors	Implies measurements that assess the state of psycho-social processes that can produce behaviours in the target population, for example, knowledge, motives, norms, automated behaviour, perceived difficulties, etc
Hand hygiene capabilities	Implies measurements that assess to what extent a target population has the capacity to perform a correct hand hygiene technique

[5] Affordability is a common element in water supply frameworks and relates to the height of user fees. That element is not included in the results-chain. User fees are not a common when using hand hygiene facilities in public spaces. In contrast, the costs of the HHF is part of the results-chain.

[6] The SDG 6 Global Acceleration Framework is a new, unifying initiative that aims to deliver fast results at an increased scale. It is part of the UN Secretary-General's Decade of Action to deliver the SDGs by 2030. See https://www.unwater.org/publications/the-sdg-6-global-acceleration-framework/

Outputs 2: Hand Hygiene Facilities: Infrastructure & Consumables	
Availability of HHFs	Implies measurements that assess if enough HHF facilities are available at a given public space. These measurements can also assess what type of HHF that is available (type of taps, drainage system, reservoir, cleansing agent, etc).
Accessibility of HHF	Implies measurements that assess if hand hygiene facilities in these places are available for use by all people at all times, and that the facilities are suitably located within the public space and designed so as to can be accessed safely and securely (adapted from (OHCHR, 2019))
Quality/Acceptability	Implies measurements that assess to what degree the quality of the HHFs impacts the acceptability for the users: the attractiveness of the HHF, the convenience and ease-of-use of the HHF (for all), and if the HHF facilitates effective hand hygiene. Measurements can also assess the cleanliness and the technical safety of the facility.
Functionality	Implies measurements that assess to what extent a HHF is operational. The measurements can include the functionality of the tap, problems with the soap dispenser or the water reservoir, the water flow, the drainage system of the HHF, etc.
Continuity	Implies measurements that assess the ongoing stability of the service provided by the HHF. Measurements can assess interruptions caused by a lack of water, consumables, defects of the HHFs, opening hours, etc.
Outputs 3: Hand Hygiene Facilities - Management	
Supervision of the HHF	Implies measurements that assess in what respect the HHF and its use are supervised. The measurements can also assess if the supervision includes steering people towards the HHF or obliging people to use the HHF before entering the public space
Accountability for HHF	Implies measurements that assess the presence and the role of an individual or institution, the “duty-bearer”, responsible for the installation, O&M, quality/acceptability and functionality of the hand hygiene service in a particular public space. Measurements can also assess the process of dealing with complaints.
Costs of hand hygiene services + financing	Implies measurements that assess how much resources are required for managing the HHF and how these are financed: costs can include the supervision of the HHF, operation and maintenance costs, capital maintenance costs, expenditure on direct and indirect support.
Operation and Maintenance	Implies measurements that assess the fulfilment of activities related to the replacement of consumables or water, minor and capital repairs to the HHF, the cleanliness of HHFs, and the replacement of HHF
Adherence to other COVID measures around HHF	Implies measurements that assess adherence to other COVID-19 hygiene measures at the HHF in a public space such as physical distancing (between taps and between people in a queue), inter-user contamination by using the station (fomites))
Sustainability of hand hygiene services	Implies measurements that assess the level of continued provision of a hand hygiene service in public spaces. Aside from the functionality of HHFs over time, measurements can include agreed characteristics of hand hygiene service (e.g. the quality/service levels).
Outcome: Hand hygiene practices	
Hand hygiene practices	Implies measurements that assess to what extent the individuals in a specific public space use the hand hygiene facilities. Measurements can include the assessment of moment of using the hand hygiene facility (before entering, after using the public restrooms etc) and the hand hygiene technique (use of soap, duration of hand washing practice, hand drying technique, etc)

Long term adherence to regulations related to hand hygiene practices in public spaces	Implies measurements that assess to what extent a population adheres to specific recommendations or regulations on hand hygiene practices in public spaces over a longer period of time (e.g. 1 year)
Impact: Health	
Reduction in hand hygiene related morbidity	Implies measurements that assess to what extent programmes contribute to a reduction in hand hygiene related morbidity (gastro-intestinal and respiratory illnesses)

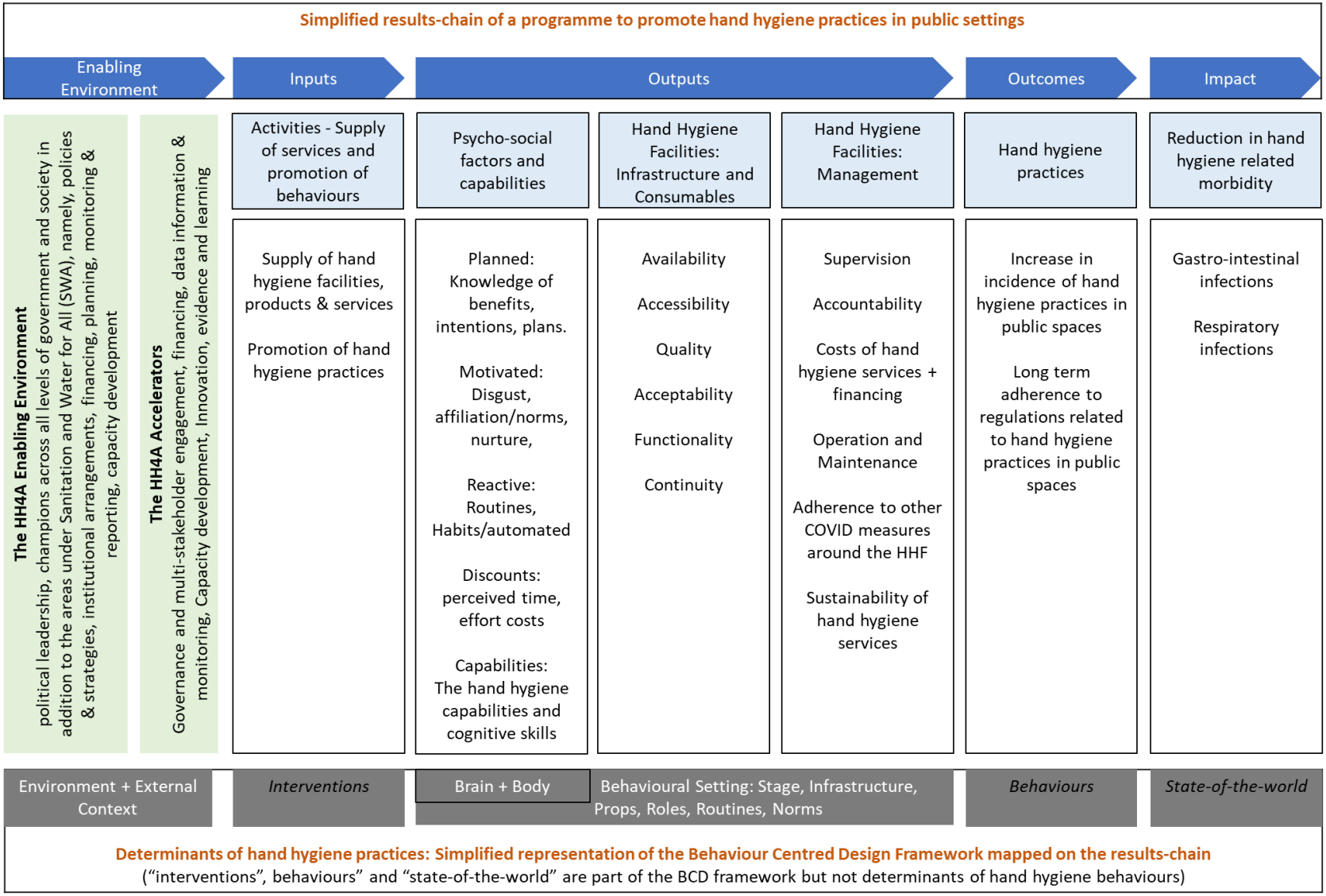
Figure 1. Definitions of the determinants for hand hygiene behaviour (BCD framework)

(authors have adapted the definitions from on the BCD checklist of determinants). Source (White et al., 2020)

Behavioural determinants defined by the BCD framework		Definitions of each determinant adapted to handwashing
Brain	Executive Brain	The extent to which knowledge of handwashing behaviour and its benefits affects handwashing intentions and plans, and eventually performance of the behaviour.
	Motivated Brain	The goal-related drivers of behaviour. Motives for handwashing can include (but is not limited to) disgust (the desire to avoid cues to sources of infection), affiliation (the desire to fit in with others) and nurture (the desire to care for your child).
	Reactive Brain	The extent to which handwashing can be automatically triggered based on past experience and repetition.
	Discounts	The perceived time, effort and costs of washing hands with soap as compared to other courses of action.
Body	Characteristics	Socio-demographic characteristics that may affect handwashing, including gender, wealth, age, education and employment.
	Senses	The sensory perceptions that may cue handwashing behaviour or be experienced during or after handwashing.
	Capabilities	Whether an individual has the skills required to wash their hands with soap. Whether an individual perceives themselves to be able and willing to actually wash their hands at the times required.
Behaviour settings	Stage	The design and set up of the specific physical spaces where handwashing behaviour takes place.
	Infrastructure	Durable infrastructure associated with handwashing such as water supply systems, sanitation, kitchen facilities and handwashing facilities.
	Props	The value, characteristics, usability, ownership and accessibility of soap and other objects used for handwashing.
	Roles	The ways in which an individual's role, identity or responsibilities influence their handwashing practices.
	Routine	The sequence of behaviours regularly performed in association with handwashing.
	Norms	The extent to which an individual's handwashing practice is influenced by their perception of normative setting-specific rules. This includes an individual's perception of whether handwashing is commonly practiced in their community (descriptive norm); whether handwashing is part of their role and their normal behaviour (personal norm); whether handwashing is socially approved of (injunctive norm); and whether handwashing is practiced by their 'valued others' (subjective norm).
Environment	Physical environment	Factors in the physical or built environment including climate and geography.
	Biological Environment	Factors associated with an individual's interaction within their biological environment.
	Social Environment	The structure of an individual's social environment, including how they interact with it and perceive themselves within it.
External context	Political and historical context	The historical and cultural events that have shaped current perceptions and practices of handwashing. The extent to which handwashing-related policies or local and national leadership on handwashing issues, shape handwashing perceptions and practices at the individual level.

Figure 2 maps the BCD framework on the simplified results-chain. This framework is used to map the content of the monitoring tools. The various elements under the different parts of the results-chain are based on the analysis presented above in A simplified results-chain for hygiene programming that will be used to map the content of the 12 tools and to conduct the gap analysisTable 2 Figure 1

Figure 2. A representation of the combination of the simplified results-chain with the BCD framework.



A key characteristic of tools is the measurement technique. Existing reviews on hand hygiene have provided an overview of the strengths and limitations of these different measurement techniques (see for example (Ram, 2013; Rutter *et al.*, 2019; GHP, 2020)). **Erreur ! Référence non valide pour un signet.** presents an overview of these strengths and limitations of the techniques used by the 12 tools.

Table 3. Overview of the strength and limitations of the measurement techniques used in public spaces.
Source: adapted from (Rutter et al., 2019; GHP, 2020)

Measurement technique	Descirption	Strengths	Limitations
Self-reported data	This is normally measured via a survey or interview. There are a range of ways people can self-report aspects of handwashing behaviour. Questions can measure frequency, handwashing at critical times, knowledge, product use, and intention #There are other forms of self-reported data which do not relate to hygiene behaviours (inquiring about the performance by an individual of specific tesks that an individual, for example)	Self-reported information is normally quick and easy to obtain. Useful to understand knowledge about behaviour. Can be used to complement other behavioural outcome measures	Does not provide a reliable understanding of behaviour due to social desirability bias (people are likely to say they wash their hands more than they actually do) and recall bias (people find it hard to remember handwashing frequency accurately) #self-reports concerning tasks can also be subject social desirability bias
Spot-checks (Proxy measures, sometiles called "rapid observations")*	Proxy measures include assessing an indirect measure of handwashing, such as availability and use or depletion of handwashing materials. The Joint Monitoring Programme hygiene indicator, which uses a spot-check assessment to see whether there is a handwashing facility with soap and water present, is an example.	Quick to collect data (much quicker than a survey). Provides a realistic estimate of behaviour (much better than self-reported surveys). Compares program to other handwashing programs around the globe and is used by most national governments	Does not reflect actual behaviour (including frequency and timing of handwashing) but it gives an indication of what behaviour is likely to be. To do this, it uses an assumption: if soap and water and a handwashing facility are not present, then hands are not being washed, as it would be too difficult and inconvenient to do on a regular basis. Even when these things are present, they do not guarantee handwashing, but their presence does create the right enabling conditions, indicating that individuals could easily practice handwashing if they wanted
Observation of practices, structured*	Data collectors spend an extended period of time, e.g. 3 hours or more in each household, school, or workplace observing behaviour and noting down whether or not hands are washed at	Measures actual behaviour rather than reported or proxy measures.	Time consuming and hard to do at scale. Requires staff to be well trained.

Measurement technique	Descirption	Strengths	Limitations
	critical occasions. To minimize bias, participants are not told that handwashing is being observed, but rather that the data collectors are learning about daily routines.	Can be useful to learn about behaviour in context and within daily routines. Considered the most reliable way of measuring handwashing behaviouro	Behaviour may be affected by the presence of observers In some settings, observation may be unacceptable
Observations of practices, limited	Short observations are used in Indonesia. Instead of recording details on each observed individual, the observer is asked to record the behaviour of just 10 people at once, without recording much detail on the handwashing event (duration of hand washing, use of proper technique). In Indonesia, the instructions for the observer are phrased as follows *"Think about 10 other people you saw in this public place. How many did you observed washing their hands with soap or hand sanitizer before entering?"* In Zimbabwe/Lebanon (tool 2), observers do not record details on the hand washing event, but observers do have the opportunity to observe more than 10 people	Less time consuming than structured observations Can be performed by observers that have received less training (like in Indonesia) Less costly / can be deployed at a scale	Reporting in batches of 10 people can result in less reliable information Reporting for 10 people can increase the potential for selection bias Reporting on handwashing events without recording details on the technique reduces the relevance of the observation
Indirect-reports* **	Here the data is provided by individuals who are in charge of a hand hygiene station and external to the institution conducting the survey (see: Rutter et al 2019). In Myanmar data is supplied by the caretaker of the hand hygiene station. The caretaker reports data which in an enumerator-led survey would mostly be collected through spot-checks. In context where record keeping is rather systematic, this may be referred to as administrative data (i.e. in schools).	Considerable time saving if data already exists. Data may be available over long time periods/more frequently	Externally supplied data may not be detailed nor accurate

* Source = (GHP, 2020)

** Source = (Rutter *et al.*, 2019)

*** Based on information from the tools

2.2 RESULTS: MAPPING AND GAP ANALYSIS

Based on the framework presented in 2.1, Table 4 presents the characteristics of the 12 tools. The following observations stand out:

- The data collection methodologies vary. The data collection method, scale of deployment, measurement technique, respondent and the scope of monitoring of many tools differ.
- Most tools use a data collection method with official surveyors who physically inspect sites. There are also examples of remote methods: an online survey, public mobile app used by volunteers/caretakers and the aggregation of NGO activity data.
- Most tools are designed for a generic set of public spaces while others are specifically used for public toilets, communal toilets, markets, and transport hubs.
- Most tools have been developed for the COVID-19 response
- Information on the sampling strategy is available for 3 tools
- Data is publicly available for 3 tools

Table 4. Overview of the 12 monitoring tools designed to monitoring hand hygiene in public spaces.

Country	Resource name	Year	Source	Data collection method	Main measurement technique	Respondent	Scale	Sampling strategy	Main part of the results-chain	Setting details	Critical hand hygiene moment monitored	Data publicly availability
Nigeria	WASH NORM Survey	2018	Gov	**In-person** Official enumerators	Spot-check Indirect-report	Surveyor Committee	National	Stratification and random sampling + purposeful sampling (quasi representative)	**Output**: Availability of the HHFs Accountability for the HHFs	Markets and Transport hubs	After using a public toilet	Yes
Zimbabwe / Lebanon 1	Monitoring HHF in public spaces: rapid observations	2020	OXFAM/ ACF/LSHTM	**In-person** Official enumerators	Observations of practices, structured	Surveyor	Project	Not specified	**Output:** Accessibility to the HHFs Operation and maintenance of the HHFs **Outcome**: Hand hygiene practices of people present at the public space	Public Spaces	Before entering a public space	No
Zimbabwe / Lebanon 2	Monitoring HHF in public spaces: structured observations	2020	OXFAM/ ACF/LSHTM	**In-person** Official enumerators	Observations of practices, limited	Surveyor	Project	Not specified	**Outcome**: Hand hygiene practices of people present at the public space	Public Spaces	Before entering a public space	No
Kenya 1	Monitoring HHF in public spaces: spot-checks	2020	OXFAM/ ACF/LSHTM	**In-person** Official enumerators	Spot-check	Surveyor	Project	Not specified	**Output**: Availability of the HHFs Accessibility to the HHFs Functionality of the HHFs	Public Spaces	Before entering a public space	No

Country	Resource name	Year	Source	Data collection method	Main measurement technique	Respondent	Scale	Sampling strategy	Main part of the results-chain	Setting details	Critical hand hygiene moment monitored	Data publicly availability
Kenya 2	Activity Tracking Dashboard	2020	WASH Cluster	**Remote** Aggregation of NGO activity data	Indirect-report	NGO	Project	Not applicable	**Input**: NGO Activities related to hand hygiene in public spaces	Public Spaces	"public spaces"	Yes
India 1	Swachh Survekshan (urban)	2019	Gov	**In-person** Official enumerators	Spot-check	Surveyor	National	Purposeful sampling	**Output**: Availability of the HHFs	Public/Communal Toilets	After using a public toilet	No
India 2	Hygiene Index (Varanasi)	2019	USAID / EY	**In-person** Official enumerators	Spot-check (+ indirect-report + self-report)	Surveyor / Caretaker	National	Not specified	**Output**: Availability of the HHFs Operation and maintenance of the HHFs	Public/Communal Toilets	After using a public toilet	No
Indonesia	3M monitoring	2020	UNICEF Indonesia	**Remote** Public mobile app-based survey with volunteers/citizens	Observations of practices, limited	Volunteers/ crowdsourcing	National	Not specified	**Output:** Availability of the HHFs Accountability for the HHF **Outcome:** Hand hygiene practices of people present at the public space	Public Spaces	Before entering a public space	No
Myanmar	Population Based Monitoring of HHF	2020	UNICEF	**Remote** Public mobile app-based survey with HHF caretakers	Indirect-report	Caretakers	National	Not specified	**Output:** Availability of HHFs	Public Spaces	Multiple	No
USA.	Survey on intention	2020	CDC	**Remote**: Online-survey	Self-report	Population	National	Random sampling by mail through	**Outputs**: The intention of people to wash	Public Toilets	After using a public toilet/Before	Yes

Country	Resource name	Year	Source	Data collection method	Main measurement technique	Respondent	Scale	Sampling strategy	Main part of the results-chain	Setting details	Critical hand hygiene moment monitored	Data publicly availability
	to wash hands							probability, address-based sampling	hands in public spaces		eating in a restaurant	
Zambia	COVID-19 results framework	2020	World Vision	**Remote** Aggregation of NGO activity data	Self-report	NGO	Project	Not applicable	**Input**: NGO Activities related to hand hygiene in public spaces	Public Spaces	public spaces	No
Global	Operation and Maintenance monitoring of HHF in public spaces	2020	WaterAid	**In-person** Official enumerators	Spot-check / self-report	Surveyor	Not specified	Not specified	**Output:** Availability of the HHFs Acceptability of the HHFs Operation and maintenance of the HHFs	Public Spaces	N/A	No

Table 5 indicates how often each measurement technique is used in the 12 tools. It shows that among these tools, the spot-check is by far the most commonly used technique followed by indirect-reports and structured observations.

Table 5. The number of times each measurement techniques is used in the 12 tools

Measurement technique	Number of measurements
Self-reported	5
Observation, limited	5
Observation, structured	11
Indirect-report	23
Spot-check	48
Total	**92**

2.3 RESULTS: THE CONTENT AND THE GAP ANALYSIS

Figure 3 maps the tools on the simplified results-chain for hand hygiene programming. The figure visualizes that tools have different objectives and that some tools monitor one area of the results-chain, while other cover multiple areas. The outputs "HHF infrastructure & consumables" and "HHF management" are most often covered. Three tools monitor hand hygiene practices. Only one tool assesses at psycho-social factors. None of the tools are designed to monitor the enabling environment or the impact.

Figure 3. A mapping of the content of the tools on the simplified results-chain

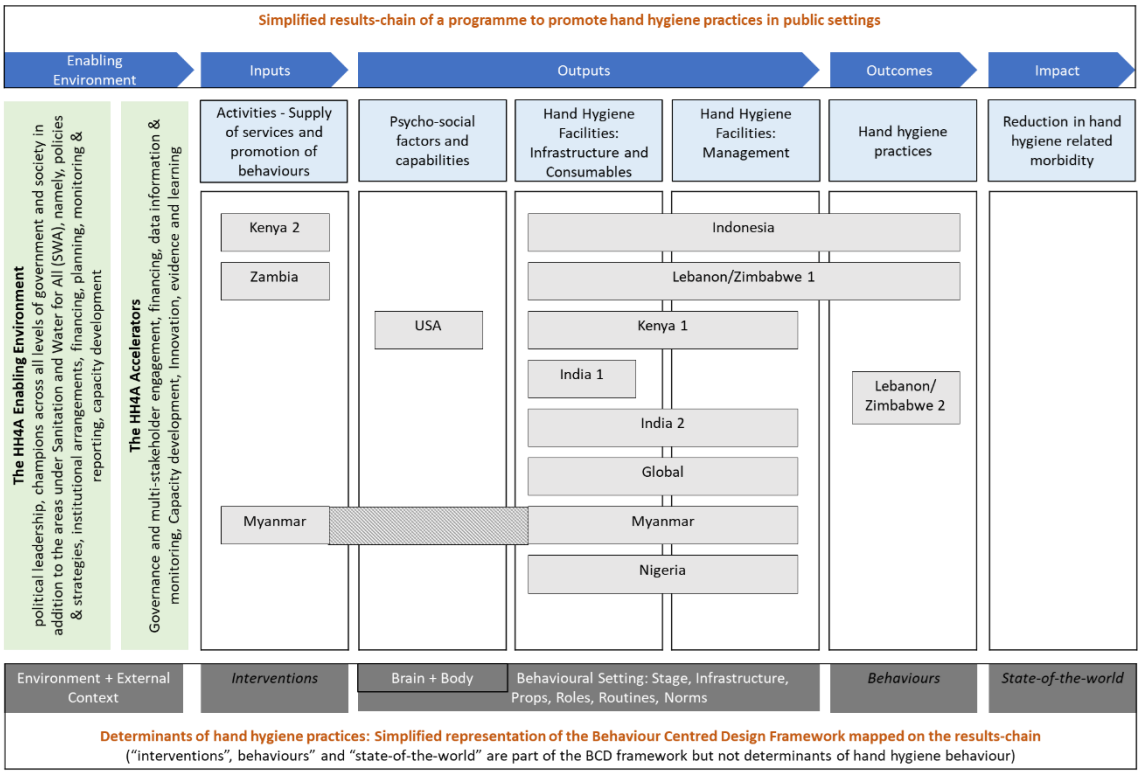


Table 6 provides more details on the content of the tools vis-à-vis the different parts of the results-chain. The table is a frequency table: it lists for each tool the number of measurements (e.g. questions/observations) per element of the results-chain.
The table shows that while most measurements fall under Outputs 2, the majority of those relate to the availability and accessibility of HHFs. Acceptability, functionality, and continuity are less frequently monitored.

The details of the 90 measurements are presented in Annex E which shows that Quality/Acceptability measures in the tools are oriented towards the cleanliness and the hygienic condition of the hand hygiene facility. The following elements of Quality/Acceptability are not covered:

- Attractiveness, the convenience, and ease-of use of the HHF
- The extent to which the HHF facilitates effective hand hygiene

In terms of accessibility, monitoring accessibility for young children and people with reduced mobility is not part of most tools

Concerning Outputs 3, most of the measurements concern the operation and maintenance of HHFs. Accountability, costs, sustainability and adherence to other COVID-19 measures are less frequently monitored or not monitored.

Table 6. The elements of the results-chain mapped against each individual tool
(showing how often each tool measures each element of the results-chain)

Elements of the results-chain \ The tools	Global	India 1	India 2	Indonesia	Kenya 1	Kenya 2	Lebanon/Zimbabwe 1	Lebanon/Zimbabwe 2	Myanmar	Nigeria	USA	Zambia	Total
Enabling environment													
HH4A enabling environment													0
HH4A accelerators													0
Inputs													
Activities related to hand hygiene in public spaces						6			2			1	7
Outputs 1: psycho-social factors and skills													
Psycho-social factors											2		2
Hand hygiene skills													0
Outputs 2: Hand hygiene infrastructure and consumables													
Type/location of public space				1	1				1	2			4
Availability of HHF	2	2	3	1	12				5	2			29
Accessibility of HHF	1		1		6		4		1	2			13
Quality/Acceptability of HHF	4												4
Functionality of HHF	1		1		2								4
Continuity of HHF			1						1				2
Outputs 3: Management of hand hygiene facility													
Supervision of the HHF				1									1
Accountability for the HHF									1	3			4
Financing of HHF management													0
Operation and Maintenance	3		2		1		2						8
Adherence to other COVID measures at the HHF													0
Sustainability of the hand hygiene service									1				1
Outcome: Hand hygiene practices of populations who visit public spaces													
Hand hygiene practices				1			5	3		1			9
Long-term adherence													0
Impact: Health													
Hand hygiene related morbidity													**0**
Total	11	2	8	4	22	6	11	3	12	10	2	1	**92**

Comparing the results of Table 6 with the BCD framework (Figure 1) demonstrates that the elements of Outputs 2 and Outputs 3 of the results-chain correspond closely to the "behaviour setting" of the BCD framework. In particular, almost all measures overlap with the determinant "stage", "infrastructure" and "props". That makes the "behaviour setting" practically the only category of determinants that is covered by the tools. The tools are not designed to cover "Body", "Environment" and "External Context". There are two measurements in the tools that fall under the "Brain" category (Outputs 1).

2.4 RESULTS: MAPPING OF THE MONITORING METHODOLOGIES

Figure 3 above has shown that each of the tools are designed for its own specific purpose.

Figure 4 organises the tools by data collection method and measurement technique. The figure also shows the focus of the tool (in brackets) and whether the tool is currently deployed at scale (tools in black). It underlines the variety of approaches that exist among the available tools.

The colours reflect the relative costs for collecting the data (green = lowest, yellow = medium, orange = higher, red = highest). The colour scheme is indicative and based on information on the cost-effectiveness of data collection summarized in existing reviews, such as (Ram, 2013; Rutter *et al.*, 2019; GHP, 2020). No actual information is available on the costs of collecting data with these tools. In addition, there is little information on the costs of some of the measurement techniques (i.e. indirect reports and observations by citizens using a mobile app).

The measurement technique and the data collection method impact the transaction costs of collecting data[7]. For example, the difficulty to collect data with trained enumerators using spot-checks on the availability of soap is greater than aggregating NGO-reports on the deployment of HHFs. But the purpose of these two approaches differs as well, and in both cases the data might be relevant. The potential to collect valid data specifically on hand hygiene practices however increases as the costs increase. There is no universally applicable method for measuring handwashing behaviour that is valid, relevant, affordable, and logistically feasible for the various settings in which behaviour might need to be measured (Ram, 2013). The feasibility of a monitoring approach is a function of the purpose, the scale, the capacity of field-staff, the available time and budget. (NB in particular contexts, such as the COVID19 pandemic, the safety of the field-staff, movement restrictions, local viral load prevalence, community perceptions can also impact the feasibility of a monitoring methodology. Box A presents more information on the safety of field-staff).

[7] See the available reviews, notably (Ram, 2013; Rutter *et al.*, 2019; GHP, 2020)

Figure 4. Mapping of the tools by data collection method and main measurement technique

Data collection method: / Main measurement technique:	Remote Online survey	Remote Aggregation of NGO activity data	Remote Public mobile app-based survey	In-person Official surveyors
Self-report	USA (brain)	Zambia (activities)		
Indirect-report		Kenya 2 (activities)	Myanmar (activities, behaviour setting)	
Spot-checks				Global, Kenya 1 Nigeria, India 1&2 (behaviour setting)
Limited observations of practices			Indonesia (behaviour setting, hygiene practices)	Zimbabwe/Lebanon 2 (hygiene practices)
Structured Observations of practices				Zimbabwe/Lebanon 1 (behaviour setting, hygiene practices)
	Tool in black = tool currently deployed at scale			

Along this line, the transaction costs of data collection increases. The possibility to collect valid data on hand hygiene practices also increases.

Box A. The appropriateness and safety of conducting structured observations and spot checks at public facilities in the context of COVID-19.

In-person data collection in public spaces may not always be feasible in the context of COVID19. The Hygiene Hub provides various summary reports with information on e.g. how to protect field-staff[8] and how to decide between different measurement techniques considering the constraints imposed by COVID-19[9]. For in-person data collection in public spaces, it the following recommendations can be adopted [10]:

- Avoid handshaking. Avoid all forms of physical contact.
- Observe physical distancing by standing 1 to 2 meters away from others (depending on national guidelines)
- Wash their hands frequently, preferably on arrival to each house. Where possible give enumerators alcohol-based hand rub to take with them. Otherwise, enumerators should wash hands with soap at public facilities.
- Enumerators should not enter homes and should avoid using spaces that are enclosed, poorly ventilated or indoors.
- If this cannot be maintained at sufficient frequency, then consider doing remote forms data collection.

If in-person data collection is not feasible, remote data collection methods can be used. The hygiene hub has developed a summary report on remote data collection methods[11].

[8] https://resources.hygienehub.info/en/articles/3859424-what-protection-measures-can-our-hygiene-promoters-take-to-stay-safe-when-working-in-communities

[9] https://resources.hygienehub.info/en/articles/4154861-summary-report-on-adapting-hygiene-project-outcome-measures-for-covid-19-response

[10] These recommendations are for designed for hygiene promotors and have been adapted for enumerators

[11] https://resources.hygienehub.info/en/articles/4165116-summary-report-on-remote-data-collection

2.5 RESULTS: SAMPLING AND SELECTION OF PUBLIC SPACES

The sampling and selection of public spaces can considerably impact the reliability of the collected data. Three tools include information on how public spaces are selected: The CDC study in the USA 2020, WASH NORM 2018 in Nigeria and the Swachh Survekshan 2019 in India. Not all tools are expected to require a sampling strategy, like the ones in which activity data of NGOs is reported (Zambia WVI and Kenya WASH Cluster). Sampling is more relevant when looking at outputs or outcomes.

Sampling public spaces for monitoring is a challenging aspect, as sample frames of public spaces may not be commonly available, and the classification of such spaces for stratification or random sampling may not yet exist (type, size, location, size of floating population, etc). Estimating coverage of facilities in front of public buldings can be particularly challenging. A short description of the selection process for three surveys is provided below:

- In the **NORM Survey 2018** in Nigeria, 1600 population enumeration areas have been selected using stratification and random sampling. All WASH facilities located within a 500 m radius were selected. The facilities included public and private WASH facilities, and those owned or being used by institutions and the community. This sampling strategy has resulted in the inclusion of 218 public spaces (150 markets and 68 transport hubs). The WASH NORM report does not indicate what approach was used to identify all the markets and transport hubs within 500 m of each catchment or if specific inclusion criteria were used (e.g. the size or the usage of that space). The report does not indicate to what extent the 218 public spaces constitute a representative sample for the markets and transport hubs in the country.
- The **Swachh Survekshan 2019** is a pan Indian survey and covers a total of 4237 Urban Local Bodies and Cantonment Boards from all states and Union Territories[12]. As a part of direct observations, survey teams needed to visit public and communal toilets. The field teams used GPS coordinates and landmarks provided by an "assessor monitoring cell" to reach the randomised sampled locations and recorded photographic evidence of their observations. The selection of public/communal toilets was based on "claims" of local authorities or through citizens reports. Besides, under the programme, public toilets must be visible on google maps (as "SBM toilets"). The required number of sites to visit depends on the size of the urban areas (see Table 7). The Swachh Survekshan 2019 survey report does not indicate how many public/communal toilets have actually been included and to what extent this was representative.

Table 7. Number of public/communal toilets to visit by size of the urban area

Inhabitants	<100.000	100.000 – 300.000	3000.000 – 1.000.000	>1.000.000
Locations to visit per Urban Local Body	12	24	40	60

- The CDC study in the USA is based on data from two surveys conducted during October 2019 and June 2020 by an online market research panel (sample size of both surveys was around 4000). This panel is designed to be representative of the noninstitutionalized U.S. population, and panel members are recruited randomly by mail through probability, address-based sampling[13]. The samples from each year were weighted to match the U.S. population across eight characteristics: sex, age, annual household income, race/ethnicity, household size, education, U.S. Census division, and residence in a metropolitan area. Sampling weights were applied to all analyses (source: (Haston, 2020).

[12] Except for the state of West Bengal

[13] Respondents receive points for participating in the panel, which can be used to redeem cash and prizes.

Concerning the coverage within a single public space, the tools do not include information on how to calculate the adequacy of the number of hand hygiene facilities in a public space (two tools look at the number of people that use a station, although this does not generate information on queuing time or crowding).

2.6 CONCLUSIONS

These are the conclusions of the content and gap analysis:

- The results-chain
  - The tools are largely designed to measure outputs (HHF infrastructure & consumables and HHF management) and outcomes (hand hygiene practices). Among the outputs, the availability and accessibility of HHF is measured the most. The tools are not designed to monitor the enabling environment or the impact.
  - With the exception of one tool, the quality and acceptability of the hand hygiene facilities are not monitored (attractiveness, convenience and ease-of-use, facilitation of effective hand hygiene) and the adherence to other COVID-19 hygiene measures at the public space (such as physical distancing (between taps and between people in a queue), inter-user contamination by using the station (fomites)).
  - There are various other elements of the result-chain that are not or barely monitored: supervision, functionality, costs, sustainability, hand hygiene skills.

- Determinants of hand hygiene practices
  - In terms of the five categories of the BCD framework for hand hygiene, although there are two measurements in the "Brain" category, the "Behaviour Setting" is practically the only determinant category that is covered by the tools.

As the tools are designed to capture the situation on the ground, it is not surprising that the enabling environment is not included. In addition, other elements that are not captured might be best collected through separate tools (costs, health impact, psychosocial factors, for example). Compared to a tool that looks at infrastructure and consumables, monitoring changes in elements such as Brain, Body or the External context can require different data collection mechanisms, different respondents and measurement techniques. RANAS or WASH'Em for example, which analyse psychosocial factors, include various qualitative methods. In addition, assessing the enabling environment – like under SWA – uses key informants at the institutional level as respondents.

The following observations stand out concerning the mapping of the monitoring methodologies of the tools:

- The monitoring methodologies vary. The purpose, the data collection method, scale of deployment, measurement technique, respondent and the scope of monitoring of many tools differ. The choice of monitoring approach will depend on the purpose, the scale and the available time and budget.
- Most tools use a data collection method with official surveyors who physically inspect sites.
- The spot-check is the most common measurement technique
- There is a lack of information on sampling and selection strategies employed by the tools. Based on the available information, most tools cannot estimate the coverage of hand hygiene services and rates of hand hygiene practices.
- There is limited information on the strengths and limitations of indirect-reports (Myanmar) and "limited" observations of practices (Indonesia)

3 THE QUALITY OF THE MEASUREMENTS

The tools contain a total of 90 measurements related to hand hygiene in public spaces. This chapter looks at the quality of these measurements. The aim of this chapter is to assess the strength of the measurements and learn from existing tools about what characteristics of a measurement impact its reliability and validity.

First the methods are introduced after which the results are presented. Examples are provided in this chapter and in Annex B. The rating of each of the 90 measurements is presented in Annex E.

3.1 METHODS

The measurements are assessed using two criteria: reliability, validity. These criteria are commonly part of review frameworks that assess the strength of instruments that evaluate compliance with infection control practices (see for example (Valim *et al.*, 2014)). The definitions of reliability and validity that are applied to this assessment are to a large degree based on a review of these concepts by (Bannigan and Watson, 2009).
The review also indicates cases where there is a potential for information bias (reactivity to observations and social desirability bias in questions). The definitions of the criteria and the scoring system are presented in Table 8 and

Table 9.

This assessment has certain **limitations**. The measurements are assessed based solely on the content of the tools, without using feedback from survey teams who have used the tools in the context for which they have been designed. In addition, for most tools limited supporting documents are available to the MWG. The lack of context makes it difficult to judge the relevance of individual measurements. Although information gathered by some measurements might not seem actionable to an outsider, they may nonetheless have relevance to programme managers. Therefore, this assessment does not include relevance as a review criterion.

Table 8. Description of the review criteria for measurements: reliability, validity and bias

Criterion	Description
Reliability	Reliability or consistency of the measurement is essentially concerned with 'error in measurement'. It refers to the extent to which a measure is the same each time it is performed and by whoever performs it. • Each Time: **test-retest:** A measurement scale's stability is the extent to which the same results are obtained on repeated administrations of the instrument. The estimation of reliability here focuses on the instrument's susceptibility to extraneous factors from one administration to the next. For this it is important that the parameters that are to be measured are relatively stable over time. For example, in a specific location, the presence of a handwashing facility will be stable over time. In the same location, the presence of a bar of soap can be much less stable (Ram, 2013). • By whoever: **inter-tester**: this refers to the notion that the outcome of a measurement should not be influenced by who takes the measurement; in other words that a measurement should give exactly the same result irrespective of who did the measuring. For this it is important that the parameters that are to be measured are specific, clear, not liable to multiple interpretations and that surveyors are properly trained. For example, rating the employment of a specific handwashing technique can vary between surveyors. In contrast, measuring if a bar of soap is present will likely yield the same answer between surveyors.
Validity	Validity is the degree to which the measurements give useful data about what is intended to be measured. A measurement can be reliable but not valid. A key concept is **content validity**, which rates if measurements include all relevant elements of what it intends to measure. It should also exclude irrelevant information. The following is an example of how the validity of a questions can be lowered because of having too many elements in one question: "Handwashing basins are functional without any blockage and free of dust, stain and litter?". When responding "yes" to this question, it remains unclear what the problem is. To measure content validity, the intention of the measurement should be clear. For example, the measurement of the size of a water reservoir of a hand hygiene facility can be reliable, but it may not be valid. It depends on the intention of the measurement. If the intention is to take stock of the sizes of reservoirs for procurement reasons, the measurement can be valid. If the intention is to say something about the continuity of the water supply of a handwashing facility, measuring the size of a reservoir does not suffice. But the intention of a measurement may not be explicit or clear. "Are posters containing instructions on the correct hand washing technique visible on the water reservoir of a hand washing facility?". Such a question can have multiple intentions. 1) accountability (has somebody done his/her job), 2) confirm that people have access to such information, 3) assuming that access to such information improves hand washing techniques of the users. Concerning the latter, this assumption ("or theory") about the link between a poster and the actual handwashing technique would have to be justified (**the construct validity**). In the WASH sector, there are many implicit assumptions underlying the reason for monitoring specific

	parameters. For the purposes of this review, when the intention is ambiguous or not clear, the validity of the measurement cannot receive a high rating[14]. Another straightforward way to rate the validity of a measurement is to compare it to the gold standard (**concurrent validity**). Notwithstanding challenges in conducting structured observations (see (Jeanes *et al.*, 2019)), the validity of measurers that intend to predict hand washing practices are often compared with that measurement technique (see (GHP, 2020)). For many parameters in the tools, however, a gold standard is not available.
Potential for bias	Bias doesn't make a measurement useless, but it does require caution in the interpretation of information. Certain sources of bias cannot be analysed in this review due to the lack of information. For instance, concerning structured observations, multiple sources of information-bias and selection-bias cannot be assessed. The use of internal or external observers (observer bias due to e.g. allegiance), lack of observer training, sampling bias (time of observation, number of sites sampled) etc. cannot be assessed (see (Jeanes et al., 2019)). The potential for bias of the different measurement techniques has already been reviewed (see notably (Ram, 2013) and (Rutter *et al.*, 2019)). Considering the limitations, this analysis only indicates when there is a potential for two types of information bias, and without rating that potential: **Bias in self-reports and in in-direct reports** • **Social desirability bias** may be prompted by questions which relate to tasks for which the respondent is responsible in public settings (e.g. cleanliness of a HHF, the presence of soap). **Reactivity to an observations of hand hygiene practices** • Behaviour may be affected by the presence of observers which can result in over-reporting (also known as the **Hawthorne effect**) (Rutter *et al.*, 2019).

[14] It should be noted that it is possible that while to an external reviewer the intention is ambiguous, at the project/programme level the intention may be evident to the stakeholder involved. This specific shortcoming of this analysis relates to the limitations explained at the start of the paragraph.

Table 9. The rating system of the individual measurements

Criteria	Explanation of the rating system
Reliability	- Low reliability (1) the measurement faces considerable inter-rater and/or test-retest issues - Medium reliability (3) the measurement faces moderate inter-rater and/or test-retest issues - High reliability (5) the measurement does not face inter-rater or test-retest issues
Validity	- Low validity (1): considerable validity issues: the intention of the measure is not clear - Medium reliability (3): moderate validity issues: the intention is clear, but the content does not include all the relevant elements - High reliability (5): no validity issues: the intention is clear and the content includes all relevant elements
Potential for bias	This assessment only indicates when there is a potential for two types of information bias, and without rating that potential: bias in indirect and self-reports and reactivity to an observation of hand hygiene practices

3.2 RESULTS: QUALITY OF THE MEASUREMENTS

The full rating of the 90 measurements is presented in Annex E. The data in this chapter are based on that annex.

Figure 5 and Table 10 show the average scores for both reliability and validity by measurement technique. General findings are:

- The quality of the individual measurements varies considerably.
- Within most tools the quality of measurements varies as well.
- The average score of each tool is moderate to low.
- The challenges related to the reliability and validity of the measurements are equally great.
- None of the 92 measurements are identical.
- There are two tools of which the average score exceeds 6 out of 10 (Nigeria NORM, Zimbabwe/Lebanon 1). The NORM survey contains several straightforward spot-checks and questions on accountability. The Zimbabwe/Lebanon1 contains several straightforward observations of hand hygiene practices. Within these two tools the quality of individual measurements varies as well.

Figure 5. Ratings for reliability and validity of the 90 measurements by measurement technique

(the numbers between brackets represent the total number of times that measurement technique has been used in the 12 tools)

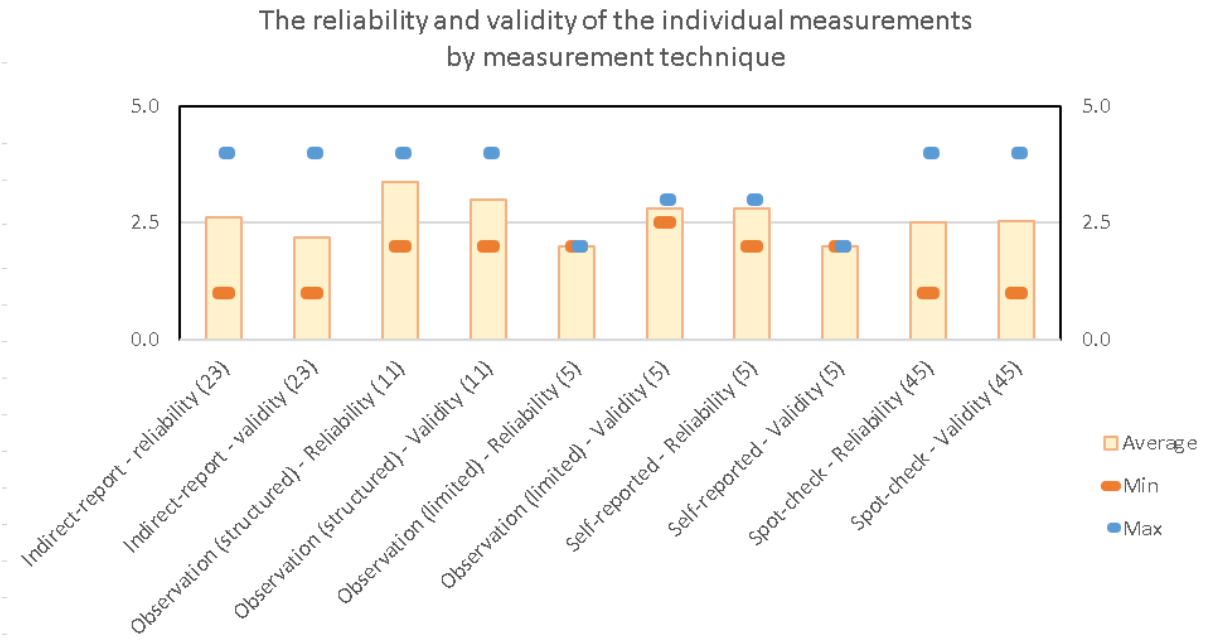


Table 10. Summary of the scores per tool

(average score and cumulative average score)

Monitoring tools	Number of measurements	Average score for reliability	Average score for Validity	Accumulative score
Zambia	1	2.0	1.0	3.0
India 1	2	2.0	2.5	4.5
USA	2	3.0	2.0	5.0
Zimbabwe/Lebanon 2	3	2.0	3.0	5.0
Indonesia	4	3.0	2.8	5.8
Kenya 2	6	3.0	2.0	5.0
India 2	8	2.5	2.3	4.8
Nigeria	10	3.6	3.1	6.7
Global	11	2.5	3.1	5.6
Zimbabwe/Lebanon 1	11	3.4	3.0	6.4
Myanmar	12	2.2	2.1	4.3
Kenya 1	22	2.1	2.0	4.1
Total general	**92**	**2.62**	**2.48**	**5.1**

RELIABILITY

Stability/test-retest

Stability of the measurements varies considerably (e.g. availability of soap, vs size of a tank). Spot-checks are the most common measurement (albeit in slightly different forms) and the spot-check on the availability of soap and water is for example part of the JMP core questions for households and institutions. Concerning household settings, there is evidence that it's important to evaluate the stability of this spot-check (see (Ram, 2013)). Compared to household settings, the stability of this measure could be lower in public settings.
The stability of structured observations is probably impacted mostly by elements that cannot be inferred from the information presented on the questionnaire (see section on bias below).
The indirect-reports mostly concern the status of infrastructure and consumables. Aside from bias, such measurements could be less subject to stability because the answer is provided by an individual with knowledge of the facility.
Aside from problems related to bias, the self-reports concerning the actions of the respondents are unlikely to have large stability issues.

Inter-rater

Inter-rater issues are common. Often questions are open to multiple interpretations. The most common shortcoming relates to the inclusion of multiple concepts in a single question or the use of unclear concepts which impact the specificity and the measurability of the measurement. There are also examples where multiple response options for a single question affect the inter-rater reliability. Compared to test-retest issues, inter-rater issues are in theory easy to address.

Examples of low rand high reliability

Question / Spot-check / Observation	Response Options	Respondent	Measurement technique
Handwashing basins are functional without any blockage and free of dust, stain and litter	• Yes • No	Surveyor	Spot-check
Reliability rating: 1 (low) – Test-retest: probably limited stability in a public setting – Inter-rater: Considerable inter-rater issues with terms that are not clearly defined and open for interpretation. Multiple concepts need to be captured in one answer.			
Please indicate if there is soap available for use	• Bar soap is available • Liquid soap is available • Soapy water is available • Soap is not available	Surveyor	Spot-check
Reliability rating: 4 (high) – Test-retest: possibly limited stability in public spaces. – Inter-rater: Question and response options are rather clear, no confusion for surveyors			

VALIDITY

Validity issues are common. Often there is a lack of clarity concerning the intention of a spot-check. There are various examples where the state of the infrastructure is assessed, but the question or the response options do not generate information that is clearly linked to a determinant of hand hygiene. In many cases the relationship is ambiguous. In addition, at times the response options render the measurement less valid by not generating enough information or by confounded information.

However, a key shortcoming of this assessment concerns the lack of background information on the tools. Most monitoring tools are not accompanied by a monitoring framework that describes purpose of measurements. In addition, it is possible that the intention is obvious for those who have developed the tool. For example, a question might relate to very specific targets of a particular programme. In a specific context particular questions can be more valid than they appear to be. To analyse such cases, the assumptions underlying the link between a measurement and the parameter under analysis must be known.

Validity issues can be resolved by improving the questions/responses, but primarily by making sure the intention is clear.

Examples of low and high validity

Question / Spot-check / Observation	Response Options	Respondent	Measurement technique
Handwashing basins are functional without any blockage and free of dust, stain and litter	• Yes • No	Surveyor	Spot-check
Validity rating: 1 (low) – Impossible to obtain clear information on either the functionality or cleanliness of the basin independently, due to the inclusion of both elements in a single question. Responding "yes" to this question does not provide clear information on the problem. This renders the intention of this spot-check ambiguous.			
Use of station: Used soap to wash hands	Yes / No	Surveyor	Structured observation
Validity rating: 4 (high) – Intention clear. This observation measurers if the observed individual uses soap to wash hands			

POTENTIAL FOR BIAS

Indirect-reports

There are various indirect-reports which inquire about the state of the infrastructure, consumables or concerning accountability. Those questions are addressed to a caretaker of the hand hygiene station:

- Is soap available at the handwashing station?
- Is there a dedicated WASH committee?
- Is there 24*7 uninterrupted water supply availability?

Externally supplied data may not be accurate (Rutter *et al.*, 2019). To what extent bias plays a role cannot be determined in this assessment.

Self-reports

Instead of inquiring about hand hygiene behaviour, the self-reports in the available monitoring tools for public spaces ask about caretakers' behaviour vis-à-vis their responsibilities. Examples:

- When was the handwashing facility last cleaned?
- Are leakages fixed/checked at regular intervals?
- Are taps and basins regularly disinfected?

Bias can play a role when providing responses which are desirable by the supervisor of caretakers an/or by providing socially desirable responses.

Observations

In two of the monitoring tools the role of the individual supervising the use of the hand hygiene station is observed:

- Was handwashing supervised/observed at the entrance?
- Did the station manager provide assistance?

In this case the individual supervising the hand hygiene facility can change behaviour because he/she is being observed. This is an example of a type of bias that is specific to public spaces

Spot-checks

No examples have been identified of spot-checks in which there is reason to believe that the potential for bias is different from other settings

Examples of potential for bias (both are rated "yes")

Question / Spot-check / Observation	Response Options	Respondent	Measurement technique
4.4. When was the handwashing facility last cleaned?	Today / Yesterday / Within last week / More than 1 week / Don't know	Caretaker	Self-reported
Potential for bias: Yes – Caretaker can be motivated to provide a response which is socially desirable. – Other general sources for bias (which cannot be assessed) o Selection bias: sampling strategy, selection of public space, timing of observation, number of observations			
Use of station: Washed hands using proper technique	Yes / No	Surveyor	Structured observation
Potential for bias: Yes – Hawthorne effect due to being observed while washing hands – Other general sources for bias (which cannot be assessed) o information bias: quality of training of enumerators o Selection bias: sampling strategy, selection of public space, timing of observation, number of observations			

RESPONSE OPTIONS

Table 11 organizes the response options used in the tools by scale of measurement: nominal, ordinal, interval and open questions. Most questions have a nominal response type (of which most is binary: yes/no). Ordinal response options are much less common. Interestingly, there are various spot-checks that use an ordinal response option that offer options that contain a lot of detail. The options clearly suggest an order which is typically analogous to "good", "average" and "poor". Table 12 shows examples of such response options. The average reliability and validity score of these spot-checks is lower than questions with a nominal response option (data not shown). It would be interesting to investigate to what extent these ordinal response options which contain many details should be avoided, or if there are ways in which they can be strengthened.

Table 11. Frequency of the type of response options used in the tools

Type of questions	#
Open question	2
Ordinal	11
Interval	16
Nominal	63
Total	**92**

Table 12. Examples of spot-checks that have lengthy response options

Spot-check: instruction for surveyor	Response Options
Please tick the statement that applies about visibility	• The handwashing station is easily visible in the vicinity of its setting e.g. from anywhere in the market or bus terminal • The handwashing station is visible if you know where it is located • The handwashing station was difficult to find, not visible at all
Would the facility be accessible for a person using a wheelchair?	• Yes • They can physically reach the station, but could not use it • They can physically use the station (reach tap etc), but could not reach it e.g. area is too muddy or uneven • No
Please indicate if there is water in the handwashing facility	• No water • There is very little water in the tank • The water level is over half of the tank's capacity • The tank is at full capacity
Please indicate if there is any visible branding on the soap	• Soap is visibly branded • Soap is not visibly branded but there are tell-tale characteristics about the brand • The soap is not visibly branded and there are no tell-tale characteristics
Did you find a handwashing station or hand sanitizer available?	"Yes, fully available = 1 (HWS with soap & water or hand sanitizer observed) Yes, partially available=2 (a lack of water, soaps or both observed) No, not available at all=3 (both HWS and hand sanitizer did not exist)"

More information in the annexes

More examples of how the ratings have been applied can be found in Annex B. The ratings for each of the 90 measurements are presented in Annex F.
Annex C presents an overview of the parameters that are assessed in each of the 90 measurements.

3.3 CONCLUSIONS

- The quality of the measurements varies widely (also within tools), with a low average score per measurement technique and per tool.
- None of the 92 measurements are the same.
- There are two tools of which the average score exceeds 6 out of 10 (Nigeria NORM, Zimbabwe/Lebanon 1). The NORM survey contains several straightforward spot-checks and questions on accountability. The Zimbabwe/Lebanon1 contains several straightforward observations of hand hygiene practices. Within these two tools the quality of individual measurements varies as well.
- In the studied examples, the quality of the monitoring data is not primarily determined by the choice of measurement technique (for example, a spot-check vs. an indirect-report), but mostly by the quality of that measurement.
    - Most reliability issues can be addressed by evaluating the stability of spot-checks in public setting (on water and soap) and by using more precise questions and response options
    - Validity can be improved by addressing the lack of clarity around the intention of measurements
    - It would be interesting to investigate to what extent ordinal response options which contain many details should be avoided, or if there are ways in which they can be strengthened.

4 DATA AND INDICATORS

Table 13 show that data from four tools were available for review. For two more tools data will reportedly become available in the near future.

Table 13. The availability of data for each tool

Monitoring tool	Data available ?
Nigeria	Yes, shared
Kenya 2	Yes, shared
USA	Yes, shared
Zambia	Yes, shared
Zimbabwe/Lebanon 1 and 2	No
Kenya 1	No
Global (WaterAid)	Verify
India 2	Verify
India 1	Verify
Indonesia	Yes, not shared
Myanmar	Yes, not shared

NIGERIA

The WASH NORM Survey 2018 has adapted the JMP WASH service ladder for WASH services in schools for use in public spaces:

- Basic: Handwashing facilities with water and soap available by the toilet/latrine
- Limited: Handwashing facilities with water but not soap available by the toilet/latrine at the public space
- No service: No handwashing facilities or no water available by the toilet/latrine

Indicators	%
Type of public place	
Markets	68.8
Motor parks	31.2
Facility « class » - toilets are accessible:	
Daily	55.0
Week calendar	11.5
Weekly	29.4
Fortnightly	4.1
Accountability	
Proportion of markets/motor parks with a management committee available	80.7
Proportion of markets/motor parks with a market or motor park association available	74.3
Proportion of markets with dedicated committee for management of WASH facilities available	29.8
Availability of hand hygiene facilities	
Proportion of Markets/Motor parks with basic hygiene service (with handwashing facilities with water and soap available by the toilet/latrine at the public place)	4.6
Proportion of Markets/Motor parks with limited hygiene services (with handwashing facilities with water but no soap available by the toilet/latrine at the public space)	3.2

Challenge:
What is the validity of the accountability indicators? What do they intent to indicate?

USA

Study conducted by the CDC using data of online market study which includes questions on hand hygiene (comparing data from before (2019) with during the COVID-19 pandemic (2020). In the report the data are disaggregated and presented by sex, age group, ethnicity, health status, region, income, work status, education, urban/rural, household size, marital status.

Indicators	2019	2020 (pandemic)
Percentage of respondents who reported remembering to wash their hands before eating before eating at a restaurant	55.2	70.6
Percentage of respondents who reported remembering to wash their hands after using the bathroom in public	95.5	94.8

Challenge:
How valid are the indicators in showing actual intention?
How valid are the indicators in predicting practices?

KENYA 2

The Ministry of Health and the WASH Cluster in Kenya ask implementing partners to report on their **activities** as part of the response to the COVID-19 outbreak. A dashboard synthesizes that information. There are various parameters relate to hand hygiene and various water supply parameters related to hand hygiene facilities.

Indicators	total	% of total in public spaces (bus stations, markets, etc)
The number of handwashing stations established	26,316	9%
The number of water storage tanks installed	1370	4%

The number of public spaces where IEC materials have been distributed	144.271	41%
The number of mass hygiene promotion activities conducted	1,251,647	100%
The number of water facilities rehabilitated	1,097	4%
m3 of water distributed daily	724	9%

Challenge:
Most activity data is not a valid indication of 1) access and 2) coverage

ZAMBIA

The Monitoring, Evaluation, Accountability and Learning (MEAL) toolkit has been developed by the COVID-19 Emergency Response (COVER) MEAL team. It contains a comprehensive set of tools (rapid assessment, onsite monitoring, key information interviews), although there is only 1 parameter on public spaces

Indicator	Target	YTD 2020	Trend	Performance
The number of community-level public handwashing stations established or maintained for COVID-19 prevention	2,986	1,774	59%	59%

Challenges:
Mixing « maintained » and « established » in a single indicator reduces the validity of the indicator
Most activity data is not a valid indication of 1) access and 2) coverage

4.1 FINDINGS

Main finding is that only 4 datasets are available. Additional datasets should be requested.

Findings based on the available datasets:

- There are no comparable indicators
- The shortcomings of the measurements can be reflected in the indicators
- The validity of the indicators is not always evident.
- Only one out of 4 data sets present information on the actual availability of services

5 OVERVIEW OF THE MAIN FINDINGS

The main findings of the assessment are:

- A wide range of approaches exist to monitor hand hygiene. The available tools show that this is also the case for public spaces. In addition to the challenge of lacking cost-effective approaches to monitor hand hygiene practices at scale, tools for public spaces also lack the capacity to estimate coverage (among and within spaces). This is primarily because of insufficient information on sampling strategies and metrics to estimate the number of required hand hygiene facilities needed per site.
- The tools focus heavily on the behaviour setting (Outputs) and practices (Outcomes). The tools are not designed to monitor the enabling environment and psychosocial factors are not monitored. The quality and acceptability of hand hygiene services are also largely neglected by the available tools. In terms of equity, monitoring accessibility for young children and people with reduced mobility is not part of most tools.

- The quality of the measurements varies widely (also within tools), with a low average score per measurement technique and per tool. Aside from the known limitations of some measurement techniques, improving the quality of the collected data can be achieved by improving the reliability and validity of the measurements. The lack of a clear intention of measurements often has a major impact on the validity.
- There are no harmonized measurements between the tools. Even the most common measurement – the spot-check on the availability of soap and water – is never the same.
- It would be interesting to assess the stability of some measurements in public spaces (compared to other settings). In addition, it can be interesting to see how questions using ordinal response options with a lot of detail can be strengthened.
- Not enough datasets are available to draw lessons from. The shortcomings of the measurements are usually reflected in the indicators. Only one dataset reports on the actual presence of hand hygiene facilities.

5.1 NEXT STEPS

1) Develop recommendations for public spaces: To what extent can and should the following elements be part of the recommendations:
   a) Guidance on the prioritization of elements to monitor based on the scale and purpose of the monitoring activity
   b) Core questions and indicators for selected criteria
   c) Guidance on sampling and selection of sites that help to obtain representative data on coverage

ANNEXES

A. SYNTHESIS OF ELEMENTS USED IN RECENT REVIEWS OF HAND HYGIENE

The framework proposed in this report can be compared with the approaches used in other reviews of measurement tools. Seven relevant review have been identified. The main elements of the review frameworks are listed in the table below. The table indicates to what extent these elements can be assessed for the available monitoring tools and if these respective elements have been incorporated in this review.

Review	**Main elements of review framework**	**Does the available information permit the analysis of the respective element?**
(Van Remoortel *et al.*, 2017)	Data type: • binary (compliance yes/no) • continuous data (compliance over time)	No
	Timing of assessment: • Uptake (during implementation of campaign or programme) • Adherence (within one year after end of implementation • Long-term use (>1 year after end of implementation)	No
	Critical moment At what key time(s)?	Yes (some)
(Valim *et al.*, 2014)	Type of measurement (response types) • Likert scale • Binary	Yes
	Content of instruments • Content and dimensions	Yes
	Psychometric characteristics • Validity • Reliability	Yes
	• Sample characteristics	Partially
(Rutter *et al.*, 2019)	Measurement technique • Consumption • detection test • observation • self-report • in-direct-report	Yes
	What parameter are measured?	Yes
	Strengths weakness of measurement technique	Yes
	When to use measurement technique. Scale, efficiency, bias	Partially
(Jeanes *et al.*, 2019)	Types of bias in structured observations • Information bias • Selection bias	Partially
	• Observer/respondent • Reliability • Potential for information bias • Potential for selection bias • Scale	Partially
(Vindigni, Riley and Jhung, 2011)	Type of setting	Yes
	Measurement technique • Proxy • Self report	Yes

	• Observation	
	Timing of assessment • Uptake/adherence/long term	No
(Oliveira and Paula, 2011)	Measurement technique • Self report (questionnaire) • Observation (quantitative, qualitative) • Proxy (consumption)	Yes
(Ram, 2013)	• Measurement technique	Yes
	• Potential for bias	Partially
	• Efficiency	Partially
	• Validity	Yes
(De Buck *et al.*, 2017)	• Methods of data collection • Data collection method • Scale • Respondent	Yes

B. EXAMPLES OF RELIABILITY AND VALIDITY RATINGS

Kenya 1. Monitoring HHF in public spaces: spot-checks. Selected measurements and justification for rating.

Question / Spot-check / Observation	Response Options	Respondent	Measurement technique	Reliability	Validity
1. Type of facility. Please tick the one that applies:	• 20 litre facility • 60 litre facility • 100 litre facility	Surveyor	Spot-check	3	1
Reliability: Test-retest: stable. Inter-rater: potential interrater issues in case the reservoir size does not correspond to the 3 response options Validity: The intention of the question is not clear?					
2. Status of Tank – Please indicate if there are any cracks or breakages on the handwashing tank:	• Yes • No	Surveyor	Spot-check	2	2
Reliability: Test-retest: stable. Inter-rater: challenging due to unclear defined terms in the question Validity: The intention of the question is not clear. Is it leaking ? Should it be replaced ?					
3. Status of Tap – Check tap. Please indicate if the tap is faulty or leaking	• Faulty – not working • Working but leaking • Working without leaking	Surveyor	Spot-check	2	2
Reliability: Test-retest: possibly limited stability (need water in tank to evaluate this + when high tank is filled up pressure increases which can affect leaking rate. Inter-rater: the phrase "leaking" is open to interpretation Validity: Response option "faulty" is valid. But the intention of "Working but leaking" is not (are taps that leak a little bit hampering hand hygiene ? does it create a pool of water / reduces the attractiveness ? requires considerably more tank refills ?): intention not clear.					
4. Please indicate if there is water in the handwashing facility	• No water • There is very little water in the tank • The water level is over half of the tank's capacity • The tank is at full capacity	Surveyor	Spot-check	2	2
Reliability: Test-retest: possibly limited stability in public spaces. Inter-rater: Non-linear scale in the response options (from "little" to "over half") Validity: Intention is clear (availability of water), although response options may provide more info than needed and the intention is not clear					
5. Please indicate if there is soap available for use	• Bar soap is available • Liquid soap is available • Soapy water is available • Soap is not available	Surveyor	Spot-check	4	4
Reliability: Test-retest: possibly limited stability in public spaces. Inter-rater. Clear question, coherent response options. Validity: Intention is clear, and response options are complete					
7. If bar soap is available, please comment on the appearance of the soap	• Soap looks dirty • Soap looks clean	Surveyor	Spot-check	2	3
Reliability: Test-retest. Possibly limited stability in public spaces. Inter-rater: rather subjective Validity: Intention is not clear enough					

Indonesia. 3M monitoring: guidance note for reporters. Selected measurements and justification for rating.

Question / Spot-check / Observation	Response Options	Respondent	Measurement technique	Reliability	Validity
Did you find a handwashing station or hand sanitiser available?	Yes, fully available = 1 (H.W.S. with soap & water or hand sanitiser observed) Yes, partially available=2 (a lack of water, soaps or both observed) No, not available at all=3 (both H.W.S. and hand sanitiser did not exist)	Volunteers / crowdsourcing	Spot-check	4	2
Reliability: Test-retest: possibly limited stability in public spaces. Inter-rater: no issues Validity: Intention is clear, information is complete, but option 2 does not provide the necessary detail					
Think about 10 other people you saw in this public place. How many did you observed washing their hands with soap or hand sanitiser before entering?	Enter number 1-10 people	Volunteers / crowdsourcing	Observation	2	3
Reliability: Test-retest: with 10 people, limited stability. Inter-rater: structured observations normally use forms to record each individual person. Possibly large inter-rater issues with this question. Validity: Intention is clear, although no info on proper handwashing technique **Potential for BIAS**: training, timing, allegiance, Hawthorne effect					
Was handwashing supervised/observed at the entrance?	Yes / No	Volunteers / crowdsourcing	Observation	3	4
Reliability: Test-retest: limited stability. Inter-rater: question concerns multiple people (10), but response option is binary. Does it concern supervision for all 10? Validity: What is the difference between observed and supervised. The intention is not entirely clear. **Potential for BIAS**: training, timing, allegiance; Hawthorne effect					

India 2. Hygiene Index. Selected measurements and justification for rating.

Question / Spot-check / Observation	Response Options	Respondent	Measurement technique	Reliability	Validity
24*7 uninterrupted water supply availability (Y/N)	Yes / No	Caretaker	Indirect-report	2	2
Reliability: Test-retest: rather stable. Inter-rater: does question relate to the moment of the survey, week, month, year? Validity: Question is incomplete. Binary response, which makes the intention unclear (JMP core questions include a time-frame, for example)					
Provision of adequate handwashing basins (with respect to number of toilet seats)	Yes / No	Surveyor	Spot-check	4	3
Reliability: Test-retest: stable. Inter-rater: If the ratio is known, this question is straightforward and clear (although clearly stating that it involves a ratio of the number of functional seats/basins would be more reliable) Validity: Question simply investigates a standard. Assuming that the national standard refers to accessible or functional toilet seats/stalls, such specifications would increase the validity of the question.					
Handwashing basins are functional without any blockage and free of dust, stain and litter	Yes / No	Surveyor	Spot-check	1	1
Reliability: Test-retest: not stable. Inter-rater: Considerable inter-rater issues with terms that are not clearly defined / open for interpretation Validity: Impossible to obtain clear information on the functionality of the basin due to the inclusion of other criteria.					
Provision of handwashing soap/bar	Yes / No	Surveyor	Spot-check	2	2
Reliability: Inter-rater: Does provision mean “available”, or “distributed in the past”? Possibly confusing for surveyors Validity: Intention not completely clear rendering the information less valid.					

Lebanon/Zimbabwe 1. Selected measurements and justification for rating.

Question / Spot-check / Observation	Response Options	Respondent	Measurement technique	Reliability	Validity
2a. Use of station: Used soap to wash hands	Yes / No	Surveyor	Structured observation	4	4
Reliability: Test-retest: could depend on number of people + time of day. Inter-rater: depends on training Validity: High validity (in combination with questions 2b, 2c 2d) **Potential for BIAS**: training, timing, number of observations, Hawthorne effect					
2b. Use of station: Washed hands using proper technique	Yes / No	Surveyor	Structured observation	3	4
Reliability: Test-retest: could depend on number of people + time of day. Inter-rater: depends on training (rating handwashing technique is more subject to interpretation than counting 20 seconds) Validity: High validity (in combination with questions 2a, 2c 2d) **Potential for BIAS**: training, timing, number of observations, Hawthorne effect					
2c. Use of station: Washed hands for at least 20 seconds	Yes / No	Surveyor	Structured observation	4	3
Reliability: Test-retest: could depend on number of people + time of day. Inter-rater: depends on training (Validity: High validity (in combination with questions 2a, 2b 2d), although the binary response limits the information that this question generates ((what to do with cases of 18seconds?) **Potential for BIAS**: training, timing, number of observations, Hawthorne effect					
2d. Use of station: Rinsed properly	Yes / No	Surveyor	Structured observation	3	3
Reliability: Test-retest: could be influenced by the length of the queue? Inter-rater: depends on training (rating rinsing technique is more subject to interpretation than counting 20 seconds) Validity: High validity (in combination with questions 2a, 2b 2c), although the binary response limits the information that this question generates **Potential for BIAS**: training, timing, number of observations, Hawthorne effect					

Nigeria. Selected measurements and justification for rating.

Question / Spot-check / Observation	Response Options	Respondent	Measurement technique	Reliability	Validity
B19. Are the handwashing facilities accessible to those with limited mobility or vision?	Yes / No	Surveyor	Spot-check	3	3
Reliability: Test-retest: medium (may be a function of crowding/availability of an agent at the facility, . Inter-rater: medium, may depend on training and interpretation, especially given there are two concepts are rated in one question/answer Validity: Two concepts are rated in one question/answer, reducing the validity of the information to analyse the problem					
B21. Are both soap and water currently available at the handwashing facilities?	Yes, water and soap = 1 Water only = 2 Soap only = 3 Neither water or soap = 4	Surveyor	Spot-check	4	4
Reliability: Test-retest: possibly limited stability in public spaces. inter-rater: question and response options clear Validity: Question and response options are concise and complete to understand the respective issue (in a context where soap is the chosen cleansing agent)					
B22. Is there evidence of use of the Handwashing Facilities?	Yes, there is dampness suggesting use = 1 Yes, other evidence = 2 No Evidence = 3;	Surveyor	Spot-check	3	2
Reliability: Test-retest: low stability Inter-rater: potentially considerable inter-rater issues because of response option “dampness” and “other evidence”. Validity: Low validity. Not clear how to interpret “evidence of use”. Not clear how to interpret “other evidence”.					
B29. Is there a dedicated committee/ caretaker for management of the WASH facilities?	Yes=1, No=2, NA=3	Caretaker	In-direct report	4	3
Reliability: Test-retest: high stability. Inter-rater: high inter-rater Validity: Having both “committee” and “caretaker” in one question reduced the validity of this question. In addition, this question clearly attempts to investigate if an entity is responsible for the facilities, but providing the answer “yes” to this question may not give a valid measure of actual accountability for the facilities					

C. OVERVIEW OF THE MEASURED PARAMETERS BY TOOL

Country	Criteria	Parameter	#
Global	**Acceptability**	Basin	1
		Cleanliness	1
		Drainage	1
		Water quality	1
	Accessibility	Path	1
	Availability	Soap	1
		Water	1
	Functionality	HWF	1
	O+M	Cleanliness	1
		General	1
		HWF	1
			11
India 1	**Availability**	Soap	1
		Water	1
			2
India 2	**Accessibility**	HWF ratio	1
	Availability	Drying, Air dryer, paper towels	1
		Soap	2
	Continuity	Water	1
	Functionality	Basin	1
	O+M	HWF	1
		Tap	1
			8
Indonesia	**Availability**	HWF / Soap / Water / Rub	1
	Location	Type of space	1
	Practice	Soap or rub used	1
	Supervision	Supervision of HHF use	1
			4
Kenya 1	**Accessibility**	Barriers	1
		Crowding	1
		for Children	1
		HWF Height	1
		Visibility	1
		Wheelchair	1
	Availability	Drainage	1
		IEC	3
		Soap	1
		Soap branding	1
		Soap cleanliness	1
		Soap dispension	1
		Soap placement	1
		Tap type	1
		Water reservoir size	1
		Water volume	1

	Functionality	Tap	2
	Location	Type of space	1
	O+M	Water reservoir	1
			22
Kenya 2	**Activity/Intervention**	HH Promotion	1
		HWF installation	1
		IEC distribution	1
		Water distribution	1
		Water storage tanks installation	1
		Water supply rehabilitation	1
			6
Lebanon / Zimbabwe 1	**Accessibility**	Autonomy of HWF use	4
	O+M	Soap	1
		Water	1
	Practice	Age	1
		Duration	1
		Rinsing	1
		Soap used	1
		Technique	1
			11
Lebanon / Zimbabwe 2	**Practice**	Number of people	1
		Soap used	1
		Water only	1
			3
Myanmar	**Accessibility**	Crowding	1
	Accountability	O+M	1
	Activity/Intervention	IEC distribution	1
		Soap distribution	1
	Availability	Drainage	1
		HWF	2
		IEC	1
		Soap	1
	Continuity	Water	1
	Location	Type of space	1
	Longevity	Date of construction	1
			12
Nigeria	**Accessibility**	Distance	1
		Limited mobility / limited vision	1
	Accountability	General	3
	Availability	HWF	1
		Water / Soap	1
	Location	Type of space	2
	Practice	Evidence	1
			10
USA	**Psychosocial**	Planning - Intention to wash hands	2
Zambia	**Activity/Intervention**	HWF installation	1
Total			**92**

D. THE ELEMENTS OF EXISTING FRAMEWORKS USED FOR THE RESULTS-CHAIN

The elements covered under the available monitoring tools can be compared with other common hand hygiene and water supply frameworks. Here six[15] frameworks have been selected that together constitute a comprehensive list of criteria:

- The WHO recommendations on hand hygiene in public spaces (WHO, 2020)
- The criteria under the human right to water and sanitation (OHCHR, 2014)
- The criteria under the hand hygiene surveillance in schools (WHO EUROPE, 2019)
- The Rural Water Metrics Global Framework (World Bank Group, 2017)
- Technical Guide for handwashing facilities in public places and buildings (WaterAID, 2020)
- The criteria of the HH4A results-chain ((HH4A, 2020))

Table 14. List of the elements under the selected hand hygiene and water supply frameworks.

WHO recommendations on hand hygiene in public spaces (WHO 2020)	Human Rights Framework for water and sanitation (OHCHR 2014)	Hand Hygiene surveillance in Schools (WHO 2019)	Water Supply Metrics (WSP 2017)	Technical Guide for handwashing facilities in public places and buildings (WaterAID 2020)	HH4A results chain (HH4A 2020)
-Availability -Location of HHF. -Accessibility -Accountability -Supervision -O+M - Obligation/ Regulation - Routine / Long term adherence	-Availability -Accessibility -Affordability -Quality -Acceptability	-Availability -Location of H.W.F. -Accessibility -Acceptability -Management -O+M -Accountability -Hygiene Promotion -Functionality	-Availability -Type of source -Accessibility -Continuity -Quality -Reliability -Affordability -Customer satisfaction -Functionality -Sustainability-Management -Accountability -O+M -Costs	- Attractive, convenient and easy-to-use - Facilitates effective hand hygiene - Sustainable - Adherence to other COVID measures	-Enabling Environment -Accelerators[16] -Supply of products and services -Hand hygiene promotion -Progress towards hand hygiene for all -Reduced morbidity

[15] WHO 2020 has been selected because it forms the reference for the promotion for hand hygiene in public spaces; OHCHR 2014 because of its important for WASH programming; WHO 2019 contains the most comprehensive set of hand hygiene criteria in institutions that is currently part of the catalogue of monitoring tools; WSP 2017 is a synthesis of water supply indicators; HH4A 2020 is a results-chain for the development of country level action plans for achieving hand hygiene for all.

[16] The HH4A accelerators (based on the SDG 6 Global Acceleration Framework). The SDG 6 Global Acceleration Framework is a new, unifying initiative that aims to deliver fast results at an increased scale. It is part of the UN Secretary-General's Decade of Action to deliver the SDGs by 2030. See https://www.unwater.org/publications/the-sdg-6-global-acceleration-framework/

E. RATINGS OF ALL THE INDIVIDUAL MEASUREMENTS

See excel file called “annex E”

G. REFERENCES

Bannigan, K. and Watson, R. (2009) ‘Reliability and validity in a nutshell’, *Journal of Clinical Nursing*, 18(23), pp. 3237–3243. doi: https://doi.org/10.1111/j.1365-2702.2009.02939.x.

De Buck, E. *et al.* (2017) ‘Approaches to promote handwashing and sanitation behaviour change in low-and middle-income countries: a mixed method systematic review’, *Campbell Systematic Reviews*, 13(1), pp. 1–447.

GHP (2020) ‘The Handwashing Handbook’. The Global Handwashing Partnership. Available at: https://globalhandwashing.org/handwashing-handbook/ (Accessed: 14 December 2020).

Haston, J. C. (2020) ‘Characteristics Associated with Adults Remembering to Wash Hands in Multiple Situations Before and During the COVID-19 Pandemic — United States, October 2019 and June 2020’, *MMWR. Morbidity and Mortality Weekly Report*, 69. doi: 10.15585/mmwr.mm6940a2.

HH4A (2020) ‘The results framework of the hand hygiene for all initiative’.

Jeanes, A. *et al.* (2019) ‘Validity of hand hygiene compliance measurement by observation: A systematic review’, *American Journal of Infection Control*, 47(3), pp. 313–322. doi: 10.1016/j.ajic.2018.08.004.

OHCHR (2014) ‘Handbook on realizing the human rights to water and sanitation’. Available at: https://www.ohchr.org/en/issues/waterandsanitation/srwater/pages/handbook.aspx (Accessed: 9 December 2020).

OHCHR (2019) ‘Human rights to water and sanitation in sphere of life beyond household, in particular in public spaces: report’. Available at: https://www.ohchr.org/EN/Issues/WaterAndSanitation/SRWater/Pages/PublicSpaces.aspx (Accessed: 9 December 2020).

Oliveira, A. C. de and Paula, A. O. de (2011) ‘Monitoring adherence to hand hygiene: a literature review’, *Acta Paulista de Enfermagem*, 24(3), pp. 407–413.

Ram, P. (2013) ‘Practical guidance for measuring handwashing behavior: 2013 update’, *Washington, DC: Water and Sanitation Program*.

Rutter, S. *et al.* (2019) ‘Evaluating children’s handwashing in schools: An integrative review of indicative measures and measurement tools’, *International Journal of Environmental Health Research*, pp. 1–19.

UNICEF 2020. Understanding Hygiene promotion in the context of Risk Communication & Community Engagement (RCCE) and Infection Control and Prevention (IPC) for the COVID-19 outbreak. COVID-19 Emergency Response UNICEF Hygiene Programming Guidance Note. Updated - 26 October 2020. https://www.unicef.org/documents/understanding-hygiene-promotion-context-coronavirus-disease-covid-19-outbreak

Valim, M. D. *et al.* (2014) ‘Instruments for evaluating compliance with infection control practices and factors that affect it: an integrative review’, *Journal of Clinical Nursing*, 23(11–12), pp. 1502–1519. doi: https://doi.org/10.1111/jocn.12316.

Van Remoortel *et al.* (2017) ‘Measuring behavioural change outcomes in development aid: a call for …’. Available at: https://fr.slideshare.net/HansVanRemoortel/measuring-behavioural-change-outcomes-in-development-aid-a-call-for-standardization-to-improve-the-evidence-synthesis (Accessed: 9 December 2020).

Vindigni, S. M., Riley, P. L. and Jhung, M. (2011) 'Systematic review: handwashing behaviour in low-to middle-income countries: outcome measures and behaviour maintenance', *Tropical Medicine & International Health*, 16(4), pp. 466–477.

WaterAID (2020) *Technical guide for handwashing facilities in public places and buildings | WASH Matters*. Available at: https://washmatters.wateraid.org/publications/technical-guide-for-handwashing-facilities-in-public-places-and-buildings (Accessed: 9 December 2020).

White, S. *et al.* (2020) 'The determinants of handwashing behaviour in domestic settings: An integrative systematic review', *International Journal of Hygiene and Environmental Health*, 227, p. 113512. doi: 10.1016/j.ijheh.2020.113512.

WHO (2020) 'Recommendations to Member States to improve hand hygiene practices to help prevent the transmission of the COVID-19 virus'. Available at: https://www.who.int/publications-detail-redirect/recommendations-to-member-states-to-improve-hand-hygiene-practices-to-help-prevent-the-transmission-of-the-covid-19-virus (Accessed: 9 December 2020).

WHO / UNICEF (2020) *Water, sanitation, hygiene, and waste management for the COVID-19 virus: interim guidance, 23 April 2020*. World Health Organization.

WHO EUROPE (2019) *Surveillance of water, sanitation and hygiene in schools. A practical tool*. Available at: https://www.euro.who.int/en/publications/abstracts/surveillance-of-water,-sanitation-and-hygiene-in-schools.-a-practical-tool-2019 (Accessed: 9 December 2020).

World Bank Group (2017) 'Toward a Universal Measure of What Works on Rural Water Supply : Rural Water Metrics Global Framework'. Available at: https://openknowledge.worldbank.org/handle/10986/27950 (Accessed: 9 December 2020).