Case Study

Sector Perspectives on the Attributes of System Approaches to Water, Sanitation, and Hygiene Service Delivery

Jeffrey P. Walters[1]; Nicholas Valcourt[2]; Amy Javernick-Will, M.ASCE[3]; and Karl Linden, M.ASCE[4]

Abstract: This study sought to coalesce sector knowledge on the use of systems approaches for sustainable water, sanitation, and hygiene (WASH) service delivery in low-income countries. To accomplish this objective, we remotely convened a panel of WASH sector experts within a multiround survey to identify, characterize, and prioritize the necessary attributes of WASH systems approaches. The first survey round asked experts to comment on the conceptual differences between traditional approaches and systems approaches to WASH. Emerging within responses from the first survey round were attributes of WASH systems (i.e., factors, actors) and attributes of WASH systems approaches (i.e., flexible, scalable). A three-round Delphi survey was then administered to reach consensus on these emerging attributes. By the final round of the Delphi, consensus was reached on every attribute of WASH systems, indicating alignment between the experts on the fundamental characteristics and implications of WASH systems. Consensus was also reached on the majority of attributes of WASH systems approaches, including the inherent interconnected and complex nature of WASH systems. However, consensus was not achieved on attributes related to mapping WASH system interconnections, convening stakeholders in collective action, and the appropriateness and feasibility of applying systems approaches at scale. This indicates a need for future research that explores practical and scalable tools and techniques to map and evaluate WASH system interactions, and ways to engage relevant actors in these approaches to collectively apply systems knowledge.
DOI: 10.1061/(ASCE)EE.1943-7870.0002010.

Author keywords: Water, sanitation and hygiene (WASH); Systems approach; Delphi; Sustainability; Developing countries.

Introduction

Sustaining delivery of services remains an enduring challenge in the global water, sanitation, and hygiene (WASH) sector, with studies still showing appreciably low rates of long-term service functionality (Foster et al. 2020; Foster and Hope 2017). Widespread service failures undermine development efforts aimed at providing access to clean water and sanitation for all, United Nations Sustainable Development Goal 6 (SDG6), with a host of direct and indirect impacts on achieving other SDGs (Le Blanc 2015). These failures have historically taken place within traditional, project-oriented WASH service delivery schemes, which have focused on siloed inputs including hand pumps, latrines, water committee training, and hygiene messaging, rather than interventions that target whole systems change (Parris and Kates 2003; Sparkman and Sturzenegger 2017; Thomas 2016). Accordingly, the WASH sector is becoming increasingly aware of the complex interplay of interdependent factors (e.g. funding, management, institutional capacity) and actors (e.g., local government, service providers, private sector) that simultaneously and dynamically influence the functionality and sustainability of WASH services (Casella et al. 2015; Cronk and Bartram 2017; IRC 2019; Neely 2015; Walters and Javernick-Will 2015). The complex nature of WASH service sustainability points to the need for 'systems approaches' for planning and managing services (Lockwood et al. 2016). Such approaches seek to gain actionable information on the interlinked and multidimensional components within the local 'WASH system' in order to inform more complexity-aware strategies for the planning, operation, management, and financing of WASH services (Schouten and Moriarty 2013).

The WASH sector has recently seen a marked increase in the number and scope of literature reporting on the use of systems approaches to improve decision making around WASH services [e.g., (Huston and Moriarty 2018; Lee et al. 2016; Neely 2019)]. This points to an acknowledgment within the sector of the need to move away from traditional approaches toward approaches that better consider and even embrace complexity. While there have been numerous applications of systems approaches to WASH, a recent systematic literature review by Valcourt et al. (2020a) of WASH systems approaches applied over the last two decades showed limited sector agreement on common or key attributes of WASH systems or systems approaches. These attributes include the role of factor interactions, scale of the system boundary, processes for identifying leverage points in the system, and the consideration of feedback mechanisms (Valcourt et al. 2020a). They found that while sector actors were employing a diverse suite of methods and approaches, there was little alignment on which factors and dimensions of WASH systems were most relevant to sustainability and how they should best be explored, identified, and addressed. We posit that this lack of common understanding of the attributes of WASH systems and of actionable and effective

[1]Assistant Professor, School of Engineering and Technology, Univ. of Washington Tacoma, Tacoma, WA 98402 (corresponding author). ORCID: https://orcid.org/0000-0001-6655-1233. Email: jpwalt@uw.edu

[2]Co-founder and Director, Open Water Systems LLC., 1001 SE Water Ave. #205, Portland, OR 97214. ORCID: https://orcid.org/0000-0002-0114-7472

[3]Professor, Dept. of Civil, Environmental, and Architectural Engineering, Univ. of Colorado Boulder, Boulder, CO 80309.

[4]Professor, Dept. of Civil, Environmental, and Architectural Engineering, Univ. of Colorado Boulder, Boulder, CO 80309.

Note. This manuscript was submitted on November 23, 2021; approved on February 8, 2022; published online on March 18, 2022. Discussion period open until August 18, 2022; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Environmental Engineering***, © ASCE, ISSN 0733-9372.

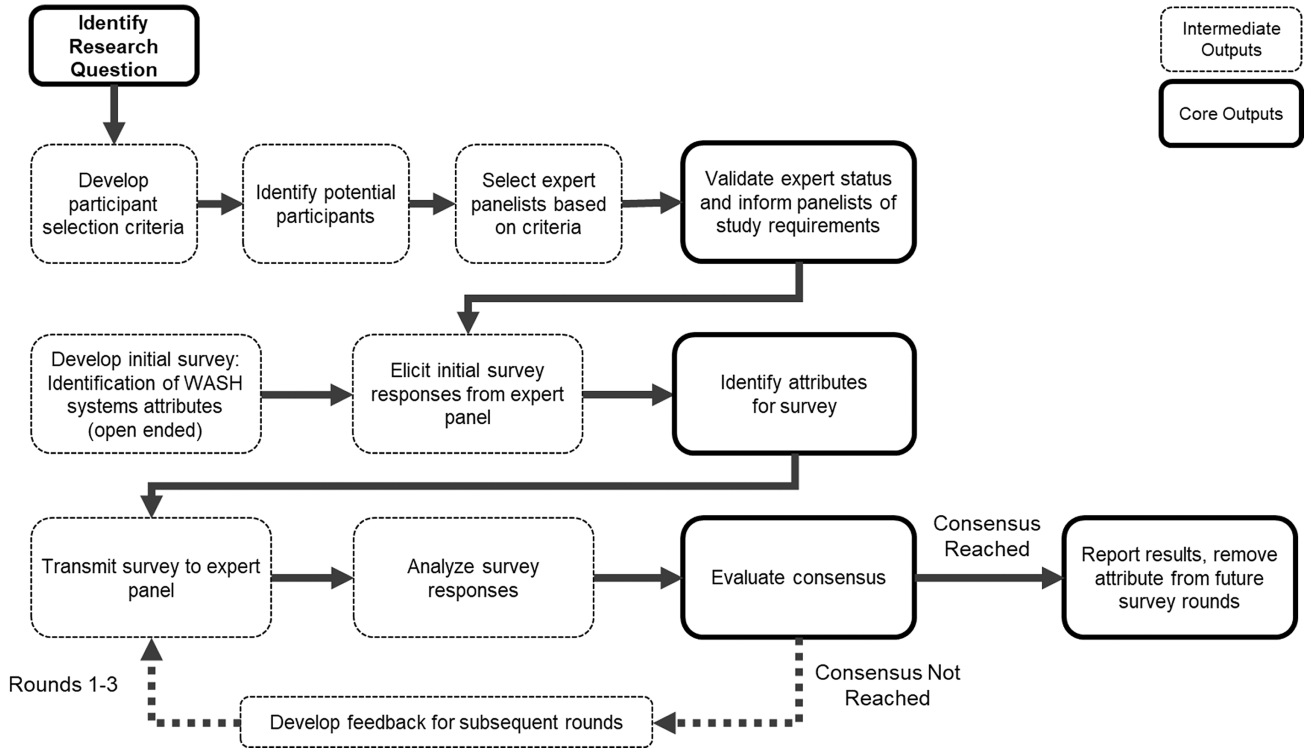

Fig. 1. Summary of the multistep research process.

approaches for understanding and intervening in these systems hinders the sector's ability to advance the use of systems approaches. To address this issue, our study sought to inform and refine sector knowledge, practice, and future research by evaluating consensus—and nonconsensus—on the key attributes of systems approaches in WASH. The overarching research question of this study is: What are the key attributes of systems approaches compared to traditional approaches to WASH service delivery?

To accomplish this objective, we remotely convened a Delphi panel of WASH sector systems experts who were asked to elaborate on the key attributes of WASH systems and systems approaches and to then rate and comment on the importance of each attribute. The study then coalesced the panelists' responses into attribute categories and iteratively asked them to rate and comment on these categories in multiple survey rounds until consensus or nonconsensus was achieved. Based on the study findings, we present recommendations for future research and practice in this emerging field within the WASH sector.

Methods

We convened leading WASH experts, including practitioners, academics, donors, and government officials, in a remote, multiround Delphi survey administered using Qualtrics. The Delphi method offers a structured and systematic means to remotely evaluate consensus between experts on a complex and nuanced topic (Linstone and Turoff 1975). The remote nature of the approach enables a more structured form of consensus between panelists whose responses are kept anonymous, eliminating the effects of one panelist's responses on others (Hallowell and Gambatese 2009). To promote consensus within the Delphi survey process, participants are provided with aggregated feedback of responses from all of the survey responses and are then asked if they would like to adjust their responses based on the compiled feedback. This process is repeated within multiple survey rounds until consensus or nonconsensus is achieved. This typically consists of between three and four rounds of iterative rating and comment responses. The Delphi method has been applied to explore consensus perspectives in many different industries including construction management (Hallowell and Gambatese 2009), humanitarian shelter and settlements (Opdyke et al. 2021), education (Dole et al. 2003; Finelli 2019), medicine and health sciences (Diamond et al. 2014), business and economics (Evans and Volery 2001; Whitehead et al. 2005), as well as the field of WASH (Kaminsky and Javernick-Will 2013; Nelson-Nuñez et al. 2019; Pugel et al. 2020; Walters and Javernick-Will 2015).

The overall research process used for this study, which included an initial survey and three follow-up Delphi rounds, is shown in Fig. 1 and described in detail in the following sections.

Panelist Selection

Following recommendations from Hallowell and Gambatese (2009), a numeric scoring criterion was used to guide our selection of panel experts (Table 1). We used a set of criteria to identify individuals who had substantial experience working within the WASH sector on the design, application, and/or study of systems approaches. Example criteria included the publication of WASH systems literature, membership of a recognized WASH sector committee, and field experience implementing WASH programming, each with weighted criteria values. A summed score of four or above was required for participation in the study.

There were 43 sector experts considered for participation in the study, of which 39 met the selection criteria and were invited to complete the initial survey. Recruitment of experts was based on a thorough review of the WASH systems literature, conference presentations on WASH systems, and the authors' robust knowledge of the key players within the WASH systems sector. Of these participants, 17 'panelists' completed all of the survey rounds. Of the final 17 panelists, 59% were WASH practitioners, 23% were

Table 1. Criteria for Delphi panel expert selection

Criteria	Points
Primary or secondary writer of scholarly journal articles that focuses on systems approaches to WASH.	1 per article up to 3
Primary or secondary writer of grey literature/ journal articles on systems approaches to WASH.	1 per article up to 3
Member or chair of a nationally (any country) recognized committee focused on WASH service planning, implementation, and management.	1
At least 10 years of professional experience with WASH service planning, implementation, and management.	3
Master or Ph.D. in the field of WASH service planning, implementation, and management.	1

academics, and 18% were either donors or government officials. The recruited experts' WASH experiences spanned nearly every continent of the world, with a particular emphasis on countries in Latin America, Asia, and Africa that are not on track to meet SDG Goal 6: "Ensure access to water and sanitation for all." Overall, the study's panelists had an average of 18 years of experience in the WASH and international development sectors, with a minimum of 7 years of experience.

Initial Survey

The study consisted of an initial survey and three sequential rounds of surveys administered using an online Qualtrics questionnaire. The initial survey asked participants to define a systems approach to WASH within open-ended text responses and to indicate the key attributes that differentiated systems approaches from traditional approaches employed within the sector. Responses were qualitatively coded to identify and group distinct attributes that are essential to systems approaches. For example, one respondent indicated, "The key differences [between traditional and systems approaches to WASH] relate to the holistic perspective and the attention to the interaction between the different elements of the system." From this response, we coded *Holistic* and *Interactions* as attributes of WASH systems approaches—two attributes that later emerged from the panelist responses. Definitions were also developed for each attribute based on the panelists' responses.

Delphi Rounds 1–3

Following the initial survey, three subsequent Delphi rounds were conducted where panelists were given the list of attributes and definitions developed from the previous round(s) and asked to rate their level of agreement on whether the attribute should be considered a key aspect of WASH systems approaches using a six-point Likert scale (1: Strongly disagree, 2: Disagree, 3: Slightly disagree, 4: Slightly agree, 5: Agree, and 6: Strongly agree). Panelists were also asked to provide comments on their rationale for each rating. After each round, we reevaluated the relative rating of each attribute to assess consensus and also qualitatively analyzed panelists' open-ended responses to examine their descriptions of each attribute. This methodology employed an initial survey with three subsequent rounds to seek consensus while minimizing bias based on best practices for Delphi studies (Hallowell and Gambatese 2009; Hsu and Sandford 2007)

Quantitative Assessment: Consensus Measures

Consensus was evaluated using panelists' Likert score ratings (1–6) to determine the median absolute deviation (MAD) and interquartile range (IQR) of responses. In accordance with standard Delphi practice (von der Gracht 2012), we defined consensus being met for each attribute if both the MAD and IQR measures were determined to be less than or equal to 1.0. The mathematical equations for MAD and IQR are shown in Eqs. (1) and (2)

$$Interquartile\ Range(IQR) = Q_3 - Q_1 \tag{1}$$

$$Where \quad Q_3 = \left[3 \times \frac{(n+1)}{4}\right] \quad and \quad Q_1 = \left[\frac{(n+1)}{4}\right]$$

where

$$n = \text{set of Likert scores } X\text{–}\{X_1, X_2, X_3 \ldots . X_n]$$

$$Median\ Absolute\ Deviation\ (MAD) = \frac{\sum |x_i - \bar{x}|}{n} \tag{2}$$

where x_i = observed Likert score; $\bar{x}$ = median Likert score; and n = count of scores.

Delphi Rounds: Rescoring

In Delphi rounds 2 and 3, the distribution of scores was presented to panelists (as a pie chart) for attributes that *did not* reach consensus, along with comments from panelists representing opposing perspectives of each attribute. Panelists were then asked to reevaluate their scores in light of the group's previous round scoring and comments. The panelists were also asked to explain their rationale for maintaining or changing their score in an open-ended comment section for each attribute. The output from this process was a list of attributes that reached consensus, those that did not, and a rich set of comments provided by panelists on the rationale of their scoring. Scoring and comments for attributes that reached consensus were not presented to panelists in subsequent rounds.

Quantitative Assessment: Attribute Agreement

In addition to evaluating measures of consensus, we also evaluated the overall level of *attribute agreement* for each attribute by summing agreement scores (Likert scores 4–6) and disagreement scores (Likert scores 1–3) for all of the panelists' ratings for each attribute. Summing Likert agreement scores in this way is an acceptable approach for nonparametric analysis to assess levels of agreement on Likert-rated dimensions (Jamieson 2004; Murray 2013). The numerical agreement and disagreement scoring for each attribute was based on the distance of the Likert score from neutral. For example, an agreement value of 3 would be assigned for all panelists who rated an attribute as "6—Strongly Agree," whereas a value of 2 would be assigned for all ratings of "5—Agree," and so on. Similarly, a disagreement value of 3, 2, and 1 would be assigned for Likert rating of "1—Strongly disagree," "2—Disagree," and, "3—Slightly disagree," respectively. Thus, attribute agreement and disagreement scores were the summation of these values for each attribute across all panelists' scoring.

Whereas attribute *consensus* was based on the deviation of Likert scores (MAD, and IQR)—representing the overall consistency of scoring across panelists—the *agreement* and *disagreement scores* revealed the relative prioritization that the group of panelists

associated to each attribute as key dimensions of systems approaches to WASH.

Results and Discussion

Panelists' responses from the initial survey—where they were asked to identify and define attributes of WASH systems approaches in open-ended responses—resulted in 22 unique attributes. In reviewing these attributes, two distinct thematic categories clearly emerged from panelists' responses that were used in subsequent rounds to group the attributes: (1) attributes inherent to *WASH systems* themselves (nine attributes), and (2) attributes inherent to *WASH system approaches* (thirteen attributes). Attributes of WASH systems highlighted important characteristics of the local or regional context, whereas attributes of WASH systems approaches pointed to characteristics of an approach or framework through which to understand and navigate WASH systems.

WASH Systems versus Systems Approaches

Notably, some attributes that panelists identified were described as having dimensions of both WASH systems and systems approaches. For example, while the attribute *Interactions* emerged as a component of both WASH systems and systems approaches, the way in which panelists defined it in these two contexts were considerably different. As one panelist indicated: "WASH systems are made up of complex and interlinked systems of people, laws, political and financial institutions, private companies, technologies, markets all constantly interacting, both formally and informally and responding to different sets of incentives, sanctions and influences." Yet, another panelist noted how *Interactions* could also be an attribute of WASH systems approaches, indicating: "The key differences [between traditional and systems approaches] relate to the holistic perspective and the attention to the interaction between the different elements of the system." Similar observations were made for two other sets of factors including *Adaptative* (WASH systems are adaptive and change over time) and *Adapts* (systems approaches are adaptable to changes in the WASH system) as well as *Nonlinear* (WASH systems have nonlinear behavior) and *Nonlinear Progress* (systems approaches can consider the nonlinearity of WASH systems).

Considering the nuances that panelists used to describe the differences between these attributes and others, we considered it important to delineate between a WASH system and WASH systems *approaches* in reporting our findings to accurately represent panelists' conceptualizations of these attributes. Thus, based on panelists' responses, the 22 attributes were assigned to either or both of the categories of "WASH systems" and "WASH systems approach" as presented in Table 2, with definitions provided for each attribute based on an analysis of the wording used by panelists' in their responses in the initial survey round.

Results of the analysis of the Delphi consensus and agreement measures for each attribute are presented in Table 3, including the round in which consensus was achieved, if applicable. Additionally, detailed results of the distribution of panelists scores for each attribute of WASH systems and systems approaches are shown in Figs. 2 and 3, respectively.

Table 2. Attributes of WASH systems and WASH systems approaches identified by panelists in the initial survey round

Attribute	Definition
Attributes of WASH systems	
Actors	Relevant WASH service stakeholders whose actions drive system outcomes (e.g., government, the private sector, households, suppliers, etc.)
Factors	Aspects or components of the WASH system that drive system outcomes (e.g., finances, politics, environmental resources, etc.)
Incentives	Motivations, sanctions, and/or policies that drive actor decisions and actions
Interactions	Interconnections, relationships, and/or interdependencies between system elements
Nonlinear	Unpredictable and disproportionate system outcomes that point to the "whole not being equal to the sum of the parts"
Feedback	Circular causality between system elements that influence system behavior over time
Adaptive	Components and conditions of elements in the WASH system evolve and change over time
Contextual	Elements of each WASH system (both actors and system components) are specific to a given geographical or geopolitical boundary and context
Complexity	The very nature of the WASH system is complex—where understanding individual system elements provides little insight into service delivery outcomes
Attributes of WASH systems approaches	
Monitoring	Provides continual monitoring and evaluation of what is working (and not working) in the WASH system
Adapts	Monitors and responds to shifts or changes in the WASH system
Mapping the system	Provides a process for mapping the components and connections of the system
Holistic	Offers a holistic perspective and attention to all of the different elements of the WASH system, as opposed to focusing only on specific parts of the system
Long-term thinking	Considers long-term outcomes versus short-term outcomes—a de-emphasis on the attribution of donor and NGO activities
Interactions	Seeks to identify and characterize the interactions between WASH system elements
More than technical	Looks beyond technical drivers of sustainable WASH service delivery
Collective action	Fosters collaboration between actors; from a standpoint that a broad spectrum of stakeholders must work together to achieve sustained services
Human systems	Considers and appreciates the influence of human behavior on service delivery outcomes
Leverage points	Enables identification of root causes, structural gaps and programmatic weaknesses
Nonlinear progress	Acknowledges that the way a system evolves over time is not a linear, predictable set of benchmarks but rather nonlinear and dynamic
Scalable	Is relevant and scalable to regional and national levels
Flexibility	Is applicable to a broad range of technical, regional, social, and geopolitical contexts

J. Environ. Eng., 2022, 148(6): 05022002

Table 3. Attributes of WASH systems and WASH systems approaches

Category	Rank[a]	Attribute	Attribute agreement score	Attribute disagreement score	Median	MAD	IQR	Consensus round
WASH systems	1	Actors	48	0	6	0.18	0	Round 1
	2	Interactions	46	0	6	0.29	1	Round 1
	3	Factors	41	3	6	0.65	1	Round 1
	4	Incentives	39	3	6	0.76	1	Round 1
	4	Feedback	39	2	5	0.71	1	Round 1
	6	Nonlinear	38	0	5	0.59	1	Round 1
	6	Contextual	38	2	5	0.76	1	Round 1
	8	Adaptive	31	7	5	0.82	1	Round 3
	9	Complexity	27	1	4	0.71	1	Round 2
WASH systems approaches	1	More than technical	45	0	6	0.35	1	Round 1
	2	Monitoring	41	0	5	0.53	1	Round 2
	3	Adapts	40	0	5	0.47	1	Round 1
	4	Interactions	38	0	5	0.35	1	Round 1
	5	Leverage points	36	1	5	0.59	1	Round 1
	5	Flexibility	36	0	5	0.35	0	Round 1
	7	Long-term thinking	34	5	5	0.94	1	Round 1
	7	Human systems	34	0	5	0.47	0	Round 1
	9	Mapping the system	33	0	5	0.65	2	No consensus
	9	Nonlinear progress	33	4	5	0.65	0	Round 1
	11	Holistic	29	4	5	0.82	1	Round 1
	12	Collective action	22	8	4	1.06	2	No consensus
	13	Scalable	19	10	4	1.12	2	No consensus

[a]Ranked by agreement score with disagreement score, median, median absolute deviation (MAD), and interquartile range (IQR) of Likert scores.

At the completion of the Delphi panel, all nine attributes of WASH systems reached consensus and 10 out of 13 attributes of a WASH systems approach reached consensus. The three attributes of a WASH systems approach that did not achieve consensus were: *Mapping the System*, *Collective Action*, and *Scalable*.

Analysis of the results presented led to the emergence of four salient findings: (1) A majority consensus on the attributes of WASH systems and systems approaches, and nonconsensus on (2) approaches for mapping WASH systems, (3) the role of collective action in a systems approach, and, (4) the ability for WASH systems approaches to be scalable. Panelists reached consensus on all

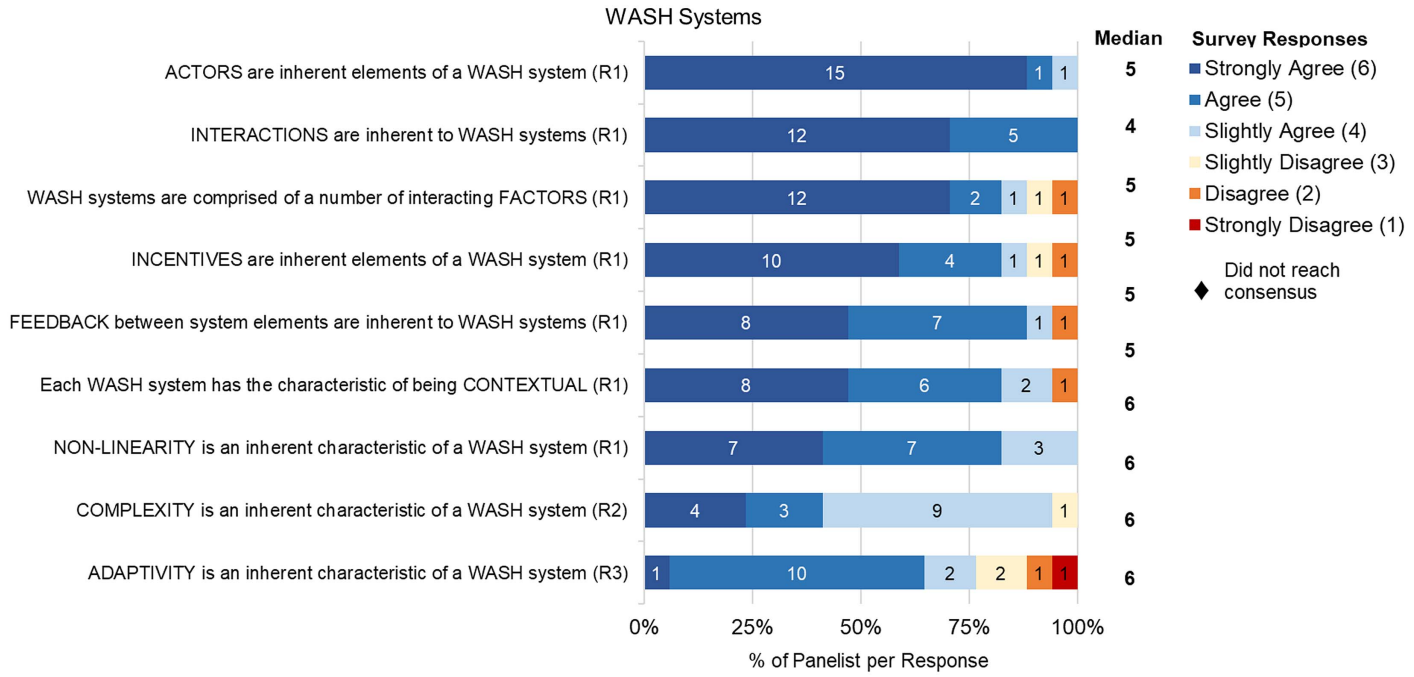

Fig. 2. Scoring distribution for attributes of WASH systems. The total number of panelists per scoring category is indicated on each score bar. The round in which each attribute achieved consensus is noted in parentheses in the attribute caption. Attributes that did not achieve consensus are denoted with a diamond shape. The median value of rating in the final round of scoring is shown on the righthand side of the chart.

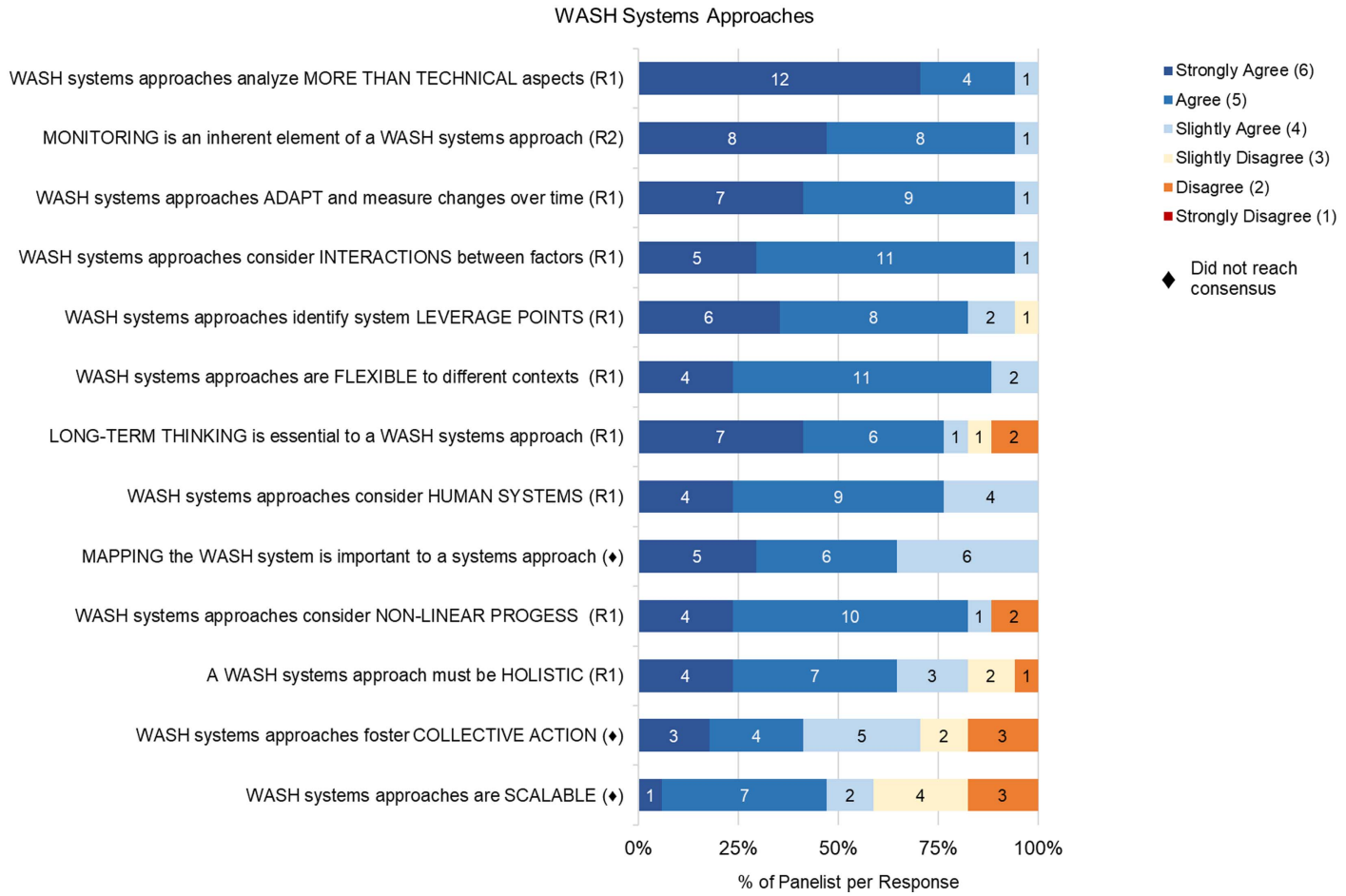

Fig. 3. Scoring distribution for attributes of WASH systems approaches. The total number of panelists per scoring category is indicated on each score bar. The round in which each attribute achieved consensus is noted in parentheses in the attribute caption. Attributes that did not achieve consensus are denoted with a diamond shape. The median value of rating in the final round of scoring is shown on the righthand side of the chart.

nine attributes of WASH systems that emerged within the study and 10 of the 13 attributes of WASH systems approaches. Of the 19 attributes that did reach consensus, 16 did so in the first round, with one that did so in the second round, and two in the third round. While this appears to infer a higher level of alignment on these 19 attributes than noted in the literature cited above, attribute agreement—and disagreement—scores across these attributes varied considerably, with attribute agreement scores varying between 27 and 48, and disagreement scores varying between 0 and 7. This illustrates the nuance of the panelists' perspectives on WASH systems and WASH systems approaches. In the following sections, we discuss these findings in detail using panelist comments and quantitative measures of the Delphi survey. To draw a connection between the findings in Table 3 in this discussion, we present, in brackets, the properties of each attribute [agreement score, rank (n of 9 for attributes of WASH systems and n of 13 for attributes of WASH systems approaches), and the round (R1, R2, or R3) when consensus was achieved (if applicable) or "NC" for no consensus].

Consensus on the Attributes of WASH Systems

All attributes of WASH systems achieved consensus in the first round except for *Complexity* [27, 9 of 9, R2] and *Adaptive* [31, 8 of 9, R3]. However, the agreement scores for these attributes varied from 27 to 48 out of a possible score of 51 (17 panelists x level 3 agreement), indicating important differences in how essential of an attribute of the WASH system panelists deemed them to be.

Consensus on the importance of the highest agreement score of any of the attributes, *Actors* [48, 1 of 9, R1] was reached quickly in the first round; however, panelists responses indicated that the attribute was not always conceptualized as the noun "actors" but also as the verb "acting," as well as actors' roles and influence on the system. As one participant indicated, "what really matters is the role(s) [actors] have assumed and how well they fulfill them." This concept of roles was noted by participants when they indicated that the relationships (or interconnections) between actors were as important as the actors themselves. This suggests that while panelists firmly agreed that *Actors* are a key part of the system, the roles they play, the relationships between them, and how these are conceptualized may still be a topic of study and debate.

While panelists also reached consensus that *Factors* [41, 3 of 10, R1] are an essential attribute of WASH systems, many responses consider the attribute to be too obvious or broad to be actionable. For example, one panelist noted that "As soon as you use the term WASH system, it assumes a complex adaptive system and then the idea of 'factors' is inherent anyway." Regarding the breadth of the use of the term 'factors', another panelist noted that "Factors is a catch-all phrase, unlike actors which is fairly straightforward . . . using 'factors' risks becoming jargon." Other panelists thought that issue was one of semantics and that the term was interchangeable with terms such as elements, components, resources,

roles, and rules. Notably, the latter three terms are considered part of the "5Rs" of a local systems approach (USAID 2014), along with relationships and results; terms that were also used by panelists in responses to other attributes. Overall, this appears to indicate that while *Factors* are clearly considered a part of WASH systems, issues of lexicon and terminology affect how experts actualize the concept. These discrepancies could lead to misunderstandings in communicating complex concepts around systems approaches to WASH as well as hinder the ability of sector practitioners to share lessons learned around this emerging field of practice. This highlights the importance of the need for thoughtful consideration of factor terminology applied within, and across, any WASH systems approach.

Panelists also readily agreed on the role of *Interactions* [46, 2 of 9, R1] as a fundamental attribute of the WASH system with the second-highest agreement score of all attributes. Importantly, multiple panelists noted that interactions determine system performance and outcomes and thus are a key part of WASH systems. For example, one panelist responded that *Interaction/interconnectedness is a core characteristic of the system, so as to say that the WASH system is an interconnected set of parts*." Yet, similar to *Factors*, some panelists thought that the term was too broad or arbitrary to have a succinct definition. This again implies a potential challenge to conceptualizing certain key terms of WASH systems, even when it enjoys broad consensus.

The attribute *Incentives* [36, 5 of 9, R1] also came to consensus in the first round as multiple panelists agreed that incentives influence and drive the system and thus determine how systems change. A thoughtful consideration of incentives was recognized as "essential to understanding the functionality of the system." However, as one panelist noted, "incentives and motivators are factors that cannot be easily measured or tracked," and thus, they can be difficult to conceptualize. Two panelists did not agree that *Incentives* were directly part of the system, but noted that they do influence the system shaping the actions and interactions of actors. This highlights the importance of considering attributes that influence WASH systems but may not necessarily be seen as being inside of that system boundary.

Most panelists agreed that *Feedback* [39, 4 of 9, R1] was an intrinsic attribute of WASH systems, with one panelist stating, "Feedback loops are characteristic of any complex system," and "whether they are negative or positive, whether or not we want to acknowledge them, they exist." These comments are notable considering the lower median rating on *Complexity* (4–see below), and the panelists qualifications that the degree of complexity, and thus conceivably the degree of *Feedback*, varies with the context of the WASH system.

Closely related to *Complexity* and *Feedback*, *Nonlinear* behavior [38, 6 of 10, R1] was one of the attributes that received no disagreement scores and reached consensus in the first round. Panelists saw this attribute as being self-evident; "this is a characteristic we expect and see because of the very nature of the system itself," with another respondent stating, "Why would we ever assume that it was linear?!." However, others cautioned that some elements can exhibit linear correlations. Panelists also expounded on issues of putting concepts of nonlinearity into action with a particular focus on how to monitor it (see *Nonlinear Progress*).

Many of the panelists agreed that regional context (*Contextual* [38, 6 of 9, R1]) significantly influences the behavior of WASH systems, as respondents indicated: "I cannot think of an engineered system that is not contextually dependent - WASH is no exception"; "of course systems are contextual"; and "this is so basic, everything sits within a context." However, other respondents argued, "some aspects [of WASH systems] are very similar to what I've seen in other countries." The respondent who disagreed with this element indicated that the "drive to recognize the local context can have the perverse effect of over-designing and thereby limiting scale and interoperability between systems." As a result, a respondent indicated, "context matters when considering how to engage, but basic principles of keeping [WASH] systems going are probably common." These insights show how panelists conceptualized multidimensional interpretations of even a relatively basic concept such as context and borders.

The WASH systems attribute *Adaptive* was the only one to reach consensus in the third round with multiple panelists changing their scores based on the scoring and comments from others presented to them in each Delphi round. It also had the highest disagreement score (7) of any of the WASH systems attributes because while many of the panelists indicated that adaptability was a desirable element of WASH system, 4 of the 17 panelists disagreed that it was a fundamental one. As one panelist noted, "To be successful, a system has to be able to evolve and react to different types of stimuli . . . there are many examples of WASH systems that are not adaptive and hence struggle." Another respondent who strongly disagreed shared comments that inspired a lively discussion in the survey responses and resulted in multiple panelists changing their scores. This panelist suggested that "Rigidity may be appropriate in stable contexts, where adaptability may be critical in instances of fluid environments." Another panelist added; "there is significant potential for a system with redundancies to resist change over a decent period of time." These responses suggest that there are delicate balances and tradeoffs between WASH systems that are fluid and responsive to change and one that is strong, redundant, and resilient to change.

While all but one panelist agreed that WASH systems are *Complex*, and thus reached consensus on this attribute in the second round, multiple panelists noted that the degree to which WASH systems could be considered complex may depend on the context or environment in which the system exists. Some panelists also took issue with the definition that was provided in the survey (see Table 2), commenting that they "strongly disagree that understanding single elements is not useful. Both are needed . . . . changing/improving (individual) components can still lead to positive systems change." Others, however, pointed to the issue of practicality, noting that donors rarely give the time and resources to address multiple components of the WASH system and that, given these constraints, focusing on individual components "may offer sufficient value in understanding service delivery outcomes." These qualifications about the role of *Complexity* were reflected in its median score of 4, the only attribute with a low median Likert score to reach consensus (all others had a median Likert score of 5 or 6). The comments show the dynamic and issue-specific nature of systems concepts, such as complexity and adaptability and their application to WASH systems planning and management, highlighting the nuance of understanding of these concepts within the WASH sector.

Summary of Consensus: WASH Systems

While all of the attributes grouped under WASH systems reached consensus, the comments by panelists combined with dissenting ratings belie the apparent agreement on these issues. These insights show that while the expert panelists agreed that these were necessary attributes of systems approaches, more gaps in knowledge and practice still need to be bridged with particular focus on the use of terminology and approaches for practically applying these concepts, a perspective that we explore in the following section on attributes of WASH systems approaches.

Consensus on Attributes of WASH Systems Approaches

For the attributes that were grouped under WASH systems approaches, 10 of 13 reached consensus with all but one (*Monitoring*) doing so in the first round. This includes all of the attributes that had a corresponding attribute in the WASH systems grouping (*Interactions*, *Nonlinear*, *Adapts/Adaptative*) that all reached consensus in the first round under the systems grouping as well. Additionally, a number of these attributes also had no disagreement score (*More than Technical, Monitoring, Adapts, Interactions, Flexibility, Human Systems, Mapping the System*), while some in particular had very few comments submitted by the panelists [*Flexibility* (2); *Interactions* (3); *More than Technical* (4); *Human Systems*] in comparison to the average of 16 comments per attribute. Overall, this suggests that there exists a high level of alignment among the expert panelists regarding what they see as essential components of WASH systems approaches.

The systems approach attribute with the highest agreement score (45) was *More Than Technical* solutions [45, 1 of 13, R1] that one panelist described as "arguably, these are the more important levers." The respondent with the lowest score felt that this attribute was so basic that it "could equally be a statement about what constitutes a 'system'." Others provided examples of nontechnical drivers relating to attributes of the WASH system, "especially perverse incentives created by the political climate and aid programs." Thus, it is clear that the panelists considered this attribute one of the most self-evident in the systems approaches group.

Of all of the attributes of WASH systems approaches that achieved consensus, the only one that did not do so in the first Delphi round was *Monitoring* [41, 2 of 13, R2] with many respondents increasing their level of agreement ratings in the second round. In their responses, panelists discussed issues with the current state of monitoring in the WASH sector including the lack of investments in improving analytical capability despite a large array of indicators on which to draw. This points to the need for better monitoring tools, but also a more systems-focused approach as well; "monitoring is essential, but it must be ... focused on measuring system change and whether the system is getting stronger and therefore more sustainable."

With closely related agreements scores and zero disagreement, panelists also reached consensus that systems approaches *Adapts* [40, 3 of 13, R1] and consider *Interactions* [38, 4 of 13, R1]. These two attributes shared many of the same reflections in panelist commentary as their correlated WASH system attributes of *Adaptive* and *Interactions* that also reached censuses.

Experts readably agree that identifying *Leverage Points* [36, 5 of 13, R1], or key areas to intervene in a system, is a primary reason why systems approaches are employed. As one respondent noted, "this is a fundamental goal of the method - if we can model a system of course we want to know how to best influence it toward the outcomes we want. Why else model it?" Other experts noted the importance of the process of identifying *Leverage Points* themselves stating "we need to get better at figuring out how to identify these within a system." The near-universal agreement of this attribute (16 of 17 panelists) combined with the comments about how to actualize *Leverage Points* are particularly enlightening in consideration of the *Mapping the System* attribute not achieving consensus through three Delphi rounds (see below).

Similar to the *Adapts* and *Interactions* attributes of WASH systems approaches, *Flexibility* [36, 5 of 13, R1] also achieved consensus in the first round without any disagreement Likert ratings. Alike *Interactions*, this attribute also received very few comments (4) compared with an average of 16 per attribute, with some panelists commenting that it was too closely aligned with the *Adapts* attribute. This too may infer issues of duplicative terminology within the WASH systems lexicon.

While panelists indicated that *Long-Term Thinking* [34, 7 of 13, R1] is not necessarily inherent to systems approaches, they recognized the importance and difficulty of obtaining sustainability in WASH services. Other panelists noted that long-term and short-term gains are not mutually exclusive because the latter are necessary to proceed to longer-term results. This on sustainability naturally leads to focusing on contributions to the greater long-term goal; "The difference is less about short- versus Long-Term Thinking and more about focusing on sustainability. One implication of a systems approach is giving up on the idea of attribution and settling for contribution." Some panelists mentioned that a systems approach is inherently necessary to achieve long-term gains, and thus must have the ability to identify *Leverage Points* that promote long-term solutions. Still, other panelists pointed to the inherent difficulty of balancing conflicting expectations with donors, who they noted tend to focus on short-term outcomes; "most of us in this line of work are tied to donors ... which makes the funding political, time-bound and accountable." Despite this consensus on *Long-Term Thinking*, respondents noted that locally-focused short-term improvements could still be equally impactful, highlighting the multidimensional nature of this attribute.

As noted above, the *Human Systems* attribute [34, 7 of 13, R1] generated some of the fewest comments from panelists, achieving consensus in the first round. Similar to the responses regarding *Factors*, one panelist noted that *Human Systems* may be "a more simplistic way of talking about interactions." While others likened the factor to the concept of 'political economy' (Krause 2009), they commented that the attribute may be just a different way of referring to the consideration of qualitative factors in the WASH systems approach. Thus, even though this attribute, like others, achieved consensus in the first round with a moderate agreement score (34), confusion about the application and meaning of the attribute were presented in the panelists' comments.

Regarding attributes that had similar connotations under both WASH systems and systems approaches, panelists noted that *Nonlinear* behavior [38, 4 of 9, R1] should be expected and, correspondingly, that WASH systems approaches must be able to consider *Nonlinear Progress* [33, 9 of 13, R1]. However, one panelist indicated that, in some cases, linear assumptions may provide satisfactory estimates of system behavior, "it is likely that linear approximations will often be good enough and that some relationships may actually be best described by a linear function." Others noted that some linear assumptions, for example changes in the number of households accessing clean water, are desirable because incorporating nonlinear thinking into assessments of WASH systems is difficult to contend with; "data is just not available, and resources not dedicated, to evaluate this with regularity." Panelists also indicated that one can implement a systems approach without necessarily evaluating nonlinearity. This may explain why *Nonlinear Progress* had the second-lowest agreement rating (33) out of all the factors that achieved consensus.

A general opinion of panelists was that the extent to which a systems approach is *Holistic* [29, 11 of 13, R1] is a function of where the boundary is drawn around the system—a necessary part of any systems approach (Meadows 2009). One respondent stated, "Understanding will never be completely holistic. While a systems approach does require looking across multiple factors and considering interactions, setting boundaries is important." Multiple panelists pointed out that a truly *Holistic* approach is unrealistic, as "it is often impractical or impossible to take a holistic approach, particularly if working at any scale. Perhaps a targeted intervention

informed by systems analysis is more appropriate than a holistic solution?" One expert pointed back to the utility of systems mapping for the task of gaining a holistic perspective from which to act, explaining, "at least initially, you need to figure out all the key elements of the system for the diagnosis or mapping. (And yes, that means drawing a boundary somewhere, but that's the nature of the beast.) After that, you can figure out which elements of the system you need to address." These responses demonstrate the need to balance concepts of systems thinking with the practical implications of incorporating these attributes into a systems approach for WASH.

Summary of Consensus: WASH Systems Approaches

The commentary of boundary and mapping concepts around the *Holistic* attribute closely resembles comments regarding the *Contextual*, *Leverage Points*, *Nonlinear Progress Interactions* attributes—which also reach consensus. This is not surprising in light of a growing awareness of WASH system complexity, promulgated through international WASH conferences, including the Kampala WASH Symposium (Lockwood et al. 2016) and IRC WASH Systems Symposium (IRC 2019), and work emerging from research collaboratives such as the United States Agency for International Development (USAID) Sustainable WASH Systems (SWS) Learning Partnership (e.g., 33–39). Indeed, research focusing on systems thinking and systems modeling posits that system behavior, whether favorable or unfavorable, is a result of interconnected system components (Bossel 2007; Meadows 2009; Sterman 2000). Overall, the consensus on a majority of the attributes of WASH systems approaches showcase favorable trends toward shared understanding within the sector while also highlighting important discrepancies in the use of specific terms and how they can be incorporated into a systems approach; two key outcomes for future research, policy, and practice in the WASH sector.

Nonconsensus on Mapping Approaches

Of the three attributes that failed to reach consensus, only *Mapping the System* [33, 9 of 13, NC] did so with a zero-disagreement score (i.e., all panelists rated the attribute a 4 or above). Thus, while all panelists essentially agreed that this attribute was part of WASH systems approaches, the degree to which they thought it should be prioritized varied notably. For example, while one panelist said that "Mapping is an important step in any systems analysis. Without, it is too easy to miss essential links between factors or actors," another noted that "some diagnostic(s) to understand the system is important. Does it need be a proper "systems map"? I'm not sure. If it is, the outputs need to be digestible and actionable for local audiences." Many panelists indicated that mapping the system is not required for a systems approach but that it can be "helpful to first map and understand and then act." A few panelists mentioned that systems mapping is just one of a host of tools that could be used to understand the system, and some of which could be less 'static' and yet equally, if not more, practical for use by WASH practitioners. Regarding the nondynamic nature of available mapping approaches, one panelist responded that "this effort will quickly become moot once the system starts to evolve and actors enter. It is an ongoing process, and there are lots of different ways to map the system."

A key theme of skepticism regarding the use of systems mapping appeared to be focused on the perceived (lack of) utility and accessibility of such approaches to local stakeholders in general, and government officials in particular. As panelists pointed out, "Governments will never do this - they struggle to monitor outputs and outcomes;" "academics spend way too much time and money on creating complex maps that none of the system actors can understand or utilize;" and "Mapping on its own has little impact; it's the application of insights that matters."

Indeed, historically, approaches that map interactions between WASH system factors and actors tend to have higher analytical complexity, such as Bayesian networks (Carriger et al. 2016; Cronk and Bartram 2017), social network analysis (Shongwe and Dlamini 2021; Sommerville et al. 2015, 2017; Walters 2016), and system dynamics modeling (Cannon et al. 2022; Libey et al. 2022; Liddle and Fenner 2017; Neely and Walters 2016; Valcourt et al. 2020b; Walters and Javernick-Will 2015). Expert panelists' responses further highlight a need to make some of the more intricate systems mapping approaches more practical and actionable for local stakeholders, including government officials.

Despite the apparent dissatisfaction with existing sector approaches for mapping WASH systems, panelists still noted that it was essential for understanding *Complexity*, *Feedback*, and *Nonlinear* dimensions of WASH systems, all attributes that achieved consensus. The goal of mapping, panelists noted, was to "devote time and effort to generate an understanding of a system and its dynamics before proposing interventions," and thus, "mapping is inherent to understanding the system." These insights show that while all panelists agreed that *Mapping the System* was part of a systems approach for WASH, they differed substantially as to what degree it was important and voiced serious concerns about the availability of tools for mapping, the insights that these tools generate, and thus, the ultimate value that mapping approaches bring to the WASH sector.

Nonconsensus on Collective Action

The *Collective Action* attribute [22, 12 of 13, NC]—representing a strategy that intentionally fosters collaboration between actors—did not reach consensus and had the second-lowest agreement score and second-highest disagreement score. Panelists' responses revealed an overall agreement that some form of collaboration among actors was necessary, but that this did not necessarily imply that the use of collective action within a systems approach was the way to achieve collaboration.

Despite dissenting Likert scores from 5 of the 17 panelists, responses generally supported the focus of *collaboration* within a systems approach: "if all the key parties don't coordinate and collaborate with each other, it's impossible to have a well-functioning, sustainable system." However, panelists appeared to differ on whether collective action meant that everyone was connected and the importance of this connectivity. For instance, one panelist noted; "Effective collaboration doesn't necessarily mean connecting everyone to each other. It can also be about creating key links in specific parts of a network, or clearing space for a particular actor to perform its role without interference." Still, other panelists pointed out that collaboration is not always the solution to issues of a lack of alignment among actors; "The sheer misalignment of funds means there is a need for actors to work together and align better, but this does not need to happen through collaboration . . . it could happen from strict regulation." Even in instances where improved collaboration may be required to improve WASH service sustainability, panelists cautioned that "collective action is not necessary or appropriate in every context. It's a tool in the toolbox that may or may not be right for the challenge." For example, one panelist noted that "a change in regulations may have a large impact on how water services are provided, and this may be done by a small set of actors without the collective." In another example cited by multiple panelists; market-based approaches to WASH (UNICEF 2020) require that some actors act in healthy competition with one another, an

approach that would appear somewhat at odds with collective action approaches. This appears to suggest that the panelists believe it is more important to understand how actors are connected and identified important linkages between them than to focus on getting all actors working in the same direction.

These multifaceted perspectives are also reflected in recent work from the SWS Learning Partnership on collective action within WASH systems approaches (Pugel et al. 2022) to WASH where a case study of 11 WASH coalitions observed that while there were multiple pathways through which collective action could help to make progress on complex WASH issues, "no cut-and-paste strategies exist for collective action approaches due to their complexity and sensitivity to local conditions" (Pugel et al. 2022). The authors also noted that the effectiveness of collective action is affected by many factors outside the control of local actors and requires timelines beyond typical WASH project durations; highlighting dimensions of both *Holistic* and *Long-Term Thinking* attributes.

Nonconsensus on Scalability

Lastly, there was a clear subgrouping of panelists who either agreed or disagreed with *Scalable* [19, 13 of 13, NC] as an attribute of WASH systems approaches, causing it to not reach consensus and have the lowest attribute agreement score and highest disagreement score any of the attributes in the survey. Moreover, 11 of the 17 panelists changed their scores throughout the second and third Delphi survey rounds.

Dissenting panelists voiced that systems approaches by virtue of the context-specific nature of WASH systems are not *Scalable* but that different approaches must be used for different system boundaries (i.e., national versus regional versus community level). This is not surprising, as panelists generally agreed that *Context* was an attribute of WASH systems, where scaling up the same approach across boundaries would be presumed to be at odds with the context-specific nature of WASH systems. As one panelist mentioned, "WASH systems approaches can work at different levels of organization—from community to whole of government. But I don't see this as scalable. Scaling up almost invariably neglects the contextuality of WASH systems." Experts with high agreement scores believed that scalability should be an intrinsic part of a systems approach since a systems approach would ideally have a cascading effect from national to subnational levels, or vice versa. A notable change in opinion toward an appreciation for the unique nuance of national and subnational systems was seen in later rounds. As one panelist indicated, "To work with systems is to work at scale. One can work at a community scale or a national scale and anything in between. What is important is that the dynamics at the national level cannot be assumed to be the same at the community level, so just saying that we replicate at the national scale something that worked at the community scale is a fundamental fallacy of systems thinking."

In contrast to the lack of consensus on the *Scalable* attribute, panelists did agree that a systems approach must be *Adaptable* to changes in the local context and offer *Flexibility* in how it is applied. As one expert noted "Taking a systems approach will necessarily mean that people are ready for nonlinearity and adaptation," and that, "given a system's emergent properties, it is important to be able to adapt the ways systems strengthening support is provided." Related to this, another panelist added "what's the point of monitoring if you can't adapt based on what you learn?" These comments indicate that there is clearly an awareness among the panelists for how systems approaches must stay relevant to the local system context but at the same time a concern for how to manage the tradeoff between approaches being contextually aware and still scalable to other contexts.

Sector Implications

The study findings point to a high level of alignment across the panelists in regards to the attributes of WASH systems and how systems approaches can be used as a lens to build a better understanding of these attributes of WASH service sustainability. Overall, these findings are encouraging and highlight the substantial paradigm shift that has occurred within the sector over the past decade in how thought leaders view challenges and opportunities to improve service sustainability. This shift in understanding aligns closely with major sector initiatives to improve knowledge around WASH systems (IRC 2019), as well as an enhanced focus on sustaining services beyond project timelines—a larger trend observed in multiple dimensions of human development motivated by the United Nations Sustainability Development Goals (United Nations 2015).

However, the study findings also highlight notable and compelling knowledge gaps that clearly require future research and discourse. First, while we found an aligned awareness (i.e., consensus) of WASH system attributes of interconnectivity, feedback, and complexity, we found a lack of alignment on the importance of and approaches for practically and actionably mapping WASH systems. These findings seem to be at odds with one another, as these important attributes of a complex system—on which the panelists enthusiastically agreed were important—arise from its structure (Richardson 2011). Thus, without mapping, diagraming, or explicating the structure of local WASH systems in one way or another, it appears difficult if not impossible to gain a functional understanding of the dynamic interplay of a multitude of context-specific factors and interconnections. Thus, there is a clear need for more research to understand the barriers to the adoption of systems mapping approaches within the WASH sector, despite a large suite of past methods from which sector actors can draw upon (Valcourt et al. 2020a).

Second, while we found a unanimous and vehement agreement on the impact of actors, and related attributes (e.g., incentives and human systems) on WASH system outcomes, we found a lack of alignment on the impetus for collective action approaches. Recent research into sector actors' perspectives of collective action identified multiple factors influencing such approaches (Pugel et al. 2020), but there are limited studies on the implementation of collective action in these contexts. The findings of the panel are notable in their contrast to the building enthusiasm and evidence for collective action and collaborative approaches in development at large and potentially highlight the need for future research on how to improve the efficacy of collective action approaches within WASH.

Finally, while we found panelists aligned on the need for WASH systems approaches to be flexible and adaptive to the inherently nonlinear and adaptive nature of WASH systems, we found a lack of alignment on the ability of WASH systems approaches to be applied across geopolitical scales. This finding presents a quandary for WASH systems advocates; if systems approaches require building a rich understanding of individual contexts, then what potential exists for building best practices and lessons learned that can be applied to other contexts? Essentially, if systems approaches are limited by their scalability, then what large-scale impact can they have on a sector that seeks to affect nearly half the world's population (JMP 2021). While some work has been conducted to address this issue (Ramalingam et al. 2014), we believe these findings point to a substantial knowledge gap that can impede the uptake,

application, and impact of systems approaches for WASH. Overall, the discussion on the merits for—and practice of—applying systems approaches to WASH would greatly benefit from longitudinal evidence on the true impact systems approaches have on WASH service delivery; evidence that is currently lacking in the sector of systems approaches to WASH (Valcourt et al. 2020a)

Study Limitations

There are limitations with the Delphi approach in particular, and research of this nature in general, that could have impacted the accuracy and generalizability of the study findings. First, by design, a Delphi study encourages participants to join a trend in group responses toward systematically arriving at consensus. However, there exists a possibility that this could lead to a 'bandwagon effect' of unthoughtful consideration of the subject matter (Hallowell and Gambatese 2009). However, we believe that within the bounds of this study there was minimal bias of this form, as many panelists cited thoughtful reasons for why they did or did not change their responses in light of shifts in aggregate Likert scores. Second, a key limitation of the Delphi approach is that participants lack the opportunity to interact with other participants in person and in real-time, again by design. This could have a detrimental effect on participants' understanding of terminology and sentiment, where they do not have the ability to ask clarifying questions and discuss topic implications as a group. For this reason, we sought to promote a 'quasi-dialogue' between panelists by presenting them with panelist comments from previous rounds in order to demonstrate their rationale for their scoring and provide added significance to the discussion. A third limitation relates to the sample size of participants needed to make generalizable claims on the attributes of WASH systems or WASH systems approaches. We would contend that using a participant selection criterion focused specifically on individuals with expertise with WASH systems approaches provided a solid foundation for vetting and including a sufficient quantity and quality of individuals to engage in this study. We also supplemented the Delphi results with extensive qualitative information across the three Delphi rounds, totaling 351 comments and over 12,000 words of verbatim text. As such, we believe the findings are sufficiently robust for the claims made in this paper. Finally, due to the nature of data collection, potentially long time lapses between rounds means participants may have forgotten their responses and the responses from the other panelists. Additionally, participants may have engaged with other information at this time that may have swayed their perspective on the subject. While we tried to keep the time between Delphi rounds at a maximum of two weeks (Hsu and Sandford 2007), there were challenges with a few panelists needing more than a month to respond. We sought to address this long time gap by sufficiently summarizing key consensus outcomes from the previous round via a consensus summary sent to participants before asking them to participate in the subsequent round.

Conclusions

This study sought to refine sector knowledge, practice, and future research surrounding the use of systems approaches to service delivery by evaluating sector alignment on the key attributes of a systems approach. To facilitate discussion on the attributes of WASH system approaches, we engaged 17 WASH systems experts in a multiround Delphi survey to identify, rate, and evaluate consensus on these attributes. Expert panel ratings, comments, and consensus on attributes indicated that WASH systems approaches consider nontechnical factors, can be applied to a wide range of contexts, require monitoring to evaluate shifts in the system, and consider system interactions that point to leverage points for improving WASH service sustainability.

While the participants reached consensus on all of the attributes of WASH systems (8 of 8) and all but three attributes of WASH systems approaches (11 of 13), there remained a lack of consensus regarding the importance of mapping WASH system interconnectivity, the merit for and means to engage stakeholders in system-level discussion and decision making, and the feasibility of systems approaches to be relevant and scalable to regional and national levels. To address this gap, further research and practices that focus on developing practical and accessible tools to elucidate meaningful and actionable insights based on system interactions are needed to examine how to scale these approaches to other contexts. Additionally, future longitudinal studies are needed in order to better evaluate the true impact of systems approaches on WASH service delivery outcomes.

Data Availability Statement

All data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

This work was completed with financial support from the Sustainable WASH Systems Learning Partnership through USAID under the terms of the Cooperative Agreement AID-OAA-A-16-00075 to the University of Colorado Boulder. The contents are the responsibility of the University of Colorado Boulder Sustainable WASH Systems Learning Partnership and do not necessarily reflect the views of USAID or the United States Government. For more information, visit www.globalwaters.org/SWS. Data collection methods used in this research were reviewed by George Fox University under IRB Protocol # 2182003. We would like to sincerely thank the expert panelists for dedicating substantial time and thought while participating in the Delphi rounds.

References

Bossel, H. 2007. *Systems and models: Complexity, dynamics, evolution, sustainability.* Norderstedt, Germany: Books on Demand.

Cannon, R. A., J. R. Mihelcic, K. Ghebremichael, and Q. Zhang. 2022. "Strategies to improve performance of community-managed water systems with system dynamics modeling." *J. Environ. Eng.* 148 (2): 04021079. https://doi.org/10.1061/(ASCE)EE.1943-7870.0001958.

Carriger, J. F., M. G. Barron, and M. C. Newman. 2016. "Bayesian networks improve causal environmental assessments for evidence-based policy." *Environ. Sci. Technol.* 50 (24): 13195–13205. https://doi.org/10.1021/acs.est.6b03220.

Casella, D., S. Tongeren, and I. Nikolic. 2015. *Change in complex adaptive systems: A review of concepts, theory and approaches for tackling 'wicked' problems in achieving sustainable rural water services*. The Hague, Netherlands: IRC.

Cronk, R., and J. Bartram. 2017. "Factors influencing water system functionality in Nigeria and Tanzania: A regression and Bayesian network analysis." *Environ. Sci. Technol.* 51 (19): 11336–11345. https://doi.org/10.1021/acs.est.7b03287.

Diamond, I. R., R. C. Grant, B. M. Feldman, P. B. Pencharz, S. C. Ling, A. M. Moore, and P. W. Wales. 2014. "Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies." *J. Clin. Epidemiol.* 67 (4): 401–409. https://doi.org/10.1016/j.jclinepi.2013.12.002.

Dole, R. L., K. Arvidson, E. Byrne, J. Robbins, and B. Schasberger. 2003. "Consensus among experts in pediatric occupational and physical

therapy on elements of individualized education programs." *Pediatr. Phys. Therapy* 15 (3): 159–166. https://doi.org/10.1097/01.PEP.0000083046.14713.EF.

Evans, D., and T. Volery. 2001. "Online business development services for entrepreneurs: An exploratory study." *Entrepreneurship Reg. Dev.* 13 (4): 333–350. https://doi.org/10.1080/08985620110052274.

Finelli, C. J. 2019. "Improving student learning in undergraduate engineering education by improving teaching and assessment." *Adv. Eng. Educ.* 7 (2): 30.

Foster, T., S. Furey, B. Banks, and J. Willetts. 2020. "Functionality of handpump water supplies: A review of data from sub-Saharan Africa and the Asia-Pacific region." *Int. J. Water Resour. Dev.* 36 (5): 855–869. https://doi.org/10.1080/07900627.2018.1543117.

Foster, T., and R. Hope. 2017. "Evaluating waterpoint sustainability and access implications of revenue collection approaches in rural Kenya: WATERPOINT SUSTAINABILITY AND USE IN KENYA." *Water Resour. Res.* 53 (2): 1473–1490. https://doi.org/10.1002/2016WR019634.

Hallowell, M. R., and J. A. Gambatese. 2009. "Qualitative research: Application of the Delphi method to CEM research." *J. Constr. Eng. Manage.* 136 (1): 99–107. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000137.

Hsu, C.-C., and B. A. Sandford. 2007. "The Delphi technique: Making sense of consensus." *Pract. Assess. Res. Eval.* 12 (10): 1–8.

Huston, A., and P. Moriarty. 2018. *Understanding the WASH system and its building blocks*.

IRC. 2019. *All systems go! Symposium Proceedings*. The Hague, Netherlands: IRC.

Jamieson, S. 2004. "Likert scales: How to (ab)use them." *Med. Educ.* 38 (12): 1217–1218. https://doi.org/10.1111/j.1365-2929.2004.02012.x.

JMP (Joint Monitoring Program). 2021. "Progress on household drinking water, sanitation and hygiene: 2000-2020: Five years into the SDGs." In *WHO/UNICEF joint monitoring program for water, supply and hygiene*. Geneva: World Health Organization.

Kaminsky, J., and A. Javernick-Will. 2013. "Contested factors for sustainability: Construction and management of household on-site wastewater treatment systems." *J. Constr. Eng. Manage.* 139 (12): A4013004. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000757.

Krause, M. 2009. *The political economy of water and sanitation. Routledge studies in development and society.* New York: Routledge.

Le Blanc, D. 2015. "Towards integration at last? The sustainable development goals as a network of targets: The sustainable development goals as a network of targets." *Sustainable Development* 23 (3): 176–187. https://doi.org/10.1002/sd.1582.

Lee, B., P. Sommerville, S. Farley, and R. Mayega. 2016. *SPACES MERL: Systems and complexity white paper.* Washington, DC: United States Agency for International Development.

Libey, A., P. Chintalapati, S. Kathuni, B. Amadei, and E. Thomas. 2022. "Turn up the dial: System dynamics modeling of resource allocations toward rural water supply maintenance in East Africa." *J. Environ. Eng.* 148 (4): 04022006. https://doi.org/10.1061/(ASCE)EE.1943-7870.0001982.

Liddle, E. S., and R. Fenner. 2017. "Water point failure in sub-Saharan Africa: The value of a systems thinking approach." *Waterlines* 36 (2): 140–166. https://doi.org/10.3362/1756-3488.16-00022.

Linstone, H. A., and M. Turoff. 1975. *The Delphi method, techniques and applications*. Reading, MA: Addison-Wesley.

Lockwood, H., P. Moriarty, and H. Skilling. 2016. "Kampala WASH symposium background note." In *Proc., Kampala WASH Symp*. Wivenhoe, UK: Aguaconsult.

Meadows, D. H. 2009. *Thinking in systems: A primer.* Chelsea, Vermont: Chelsea Green Publishing.

Murray, J. 2013. "Likert Data: What to use, parametric or non-parametric?" *Int. J. Business Soc. Sci.* 4 (3): 258–264.

Neely, K. 2015. "Complex adaptive systems as a valid framework for understanding community level development." *Dev. Pract.* 25 (6): 785–797. https://doi.org/10.1080/09614524.2015.1060949.

Neely, K. 2019. *Systems thinking and WASH*. Warwickshire, UK: Practical Action Publishing.

Neely, K., and J. Walters. 2016. "Using causal loop diagramming to explore the drivers of the sustained functionality of rural water services in Timor-Leste." *Sustainability* 8 (1): 57. https://doi.org/10.3390/su8010057.

Nelson-Nuñez, J., J. P. Walters, and D. Charpentier. 2019. "Exploring the challenges to sustainable rural drinking water services in Chile." *Water Policy* 21 (6): 1251–1265. https://doi.org/10.2166/wp.2019.120.

Opdyke, A., B. Goldwyn, and A. Javernick-Will. 2021. "Defining a humanitarian shelter and settlements research agenda." *Int. J. Disaster Risk Reduct.* 52 (Jan): 101950. https://doi.org/10.1016/j.ijdrr.2020.101950.

Parris, T. M., and R. W. Kates. 2003. "Characterizing a sustainability transition: Goals, targets, trends, and driving forces." *Proc. Natl. Acad. Sci.* 100 (14): 8068–8073. https://doi.org/10.1073/pnas.1231336100.

Pugel, K., et al. 2022. "Pathways for collaboratively strengthening water and sanitation systems." *Sci. Total Environ.* 802 (Jan): 149854. https://doi.org/10.1016/j.scitotenv.2021.149854.

Pugel, K., A. Javernick-Will, M. Koschmann, S. Peabody, and K. Linden. 2020. "Adapting collaborative approaches for service provision to low-income countries: Expert panel results." *Sustainability* 12 (7): 2612. https://doi.org/10.3390/su12072612.

Ramalingam, B., M. Laric, and J. Primrose. 2014. *From best practice to best fit*. London: Overseas Development Institute.

Richardson, G. P. 2011. "Reflections on the foundations of system dynamics: Foundations of system dynamics." *Syst. Dyn. Rev.* 27 (3): 219–243. https://doi.org/10.1002/sdr.462.

Schouten, T., and P. Moriarty. 2013. "The Triple-S theory of change." In *Water services that last*. The Hague, Netherlands: IRC.

Shongwe, M. I., and S. Dlamini. 2021. "A systems approach to investigating non-functionality in rural water schemes: A case study." *Sustainable Water Resour. Manage.* 7 (3): 24. https://doi.org/10.1007/s40899-021-00509-9.

Sommerville, P., R. Former, A. Bunger, and C. Hempflin. 2015. *Organizational network analysis: Nicaragua workforce development system.* Washington, DC: LINC Local LLC.

Sommerville, P., R. Former, A. Bunger, C. Hempflin, and P. Hunt. 2017. *Organizational network analysis: Public service governance network, Goma, democratic Republic of the Congo.* Washington, DC: LINC Local LLC.

Sparkman, D., and G. Sturzenegger. 2017. *Why business as usual will not achieve SDG6 in LAC: The promise of wastewater reuse, green infrastructure and small business around WASH: Conclusions from World Water Week 2016*. Washington, DC: Inter-American Development Bank. https://doi.org/10.18235/0000884.

Sterman, J. 2000. *Business dynamics: Systems thinking and modeling for a complex world.* New York: McGraw-Hill Higher Education. https://doi.org/10.18235/0000884.

Thomas, E. A. 2016. *Broken pumps and promises*. Cham, Switzerland: Springer.

UNICEF (United Nations Children's Fund). 2020. *Guidance on market-based sanitation.* New York: UNICEF.

United Nations. 2015. *Transforming our world: The 2030 agenda for sustainable development.* New York: United Nations.

USAID (United States Agency for International Development). 2014. *Local systems: A framework for supporting sustained development*. Washington, DC: USAID.

Valcourt, N., A. Javernick-Will, J. Walters, and K. Linden. 2020a. "System approaches to water, sanitation, and hygiene: A systematic literature review." *Int. J. Environ. Res. Public Health* 17 (3): 702. https://doi.org/10.3390/ijerph17030702.

Valcourt, N., J. Walters, A. Javernick-Will, K. Linden, and B. Hailegiorgis. 2020b. "Understanding rural water services as a complex system: An assessment of key factors as potential leverage points for improved service sustainability." *Sustainability* 12 (3): 1243. https://doi.org/10.3390/su12031243.

von der Gracht, H. A. 2012. "Consensus measurement in Delphi studies." *Technol. Forecasting Soc. Change* 79 (8): 1525–1536. https://doi.org/10.1016/j.techfore.2012.04.013.

Walters, J., and A. Javernick-Will. 2015. "Long-term functionality of rural water services in developing countries: A system dynamics approach to understanding the dynamic interaction of factors." *Environ. Sci. Technol.* 49 (8): 5035–5043. https://doi.org/10.1021/es505975h.

Walters, J. P. 2016. "Exploring the use of social network analysis to inform exit strategies for rural water and sanitation NGOs." *Eng. Project Organ. J.* 6 (2–4): 92–103. https://doi.org/10.1080/21573727.2016.1241243.

Whitehead, T., J. Preston, and T. Holvad. 2005. "The whole-life impacts of transport-charging interventions on business performance: A time-marching framework." *Environ. Plann. A: Econ. Space* 37 (5): 877–894. https://doi.org/10.1068/a35166.